Can Paying Politicians Well Reduce Corruption? The Effects of Wages and Uncertainty on Electoral Competition[*]

Tom-Reiel Heggedal[†] Leif Helland[‡] Rebecca Morton[§]

Abstract

We investigate the effects of wages and uncertainty on political corruption as measured by rent-taking. First, our laboratory data show that contrary to standard theory, rent-taking is not independent of, but decreases with wages in the absence of popularity shocks. Second, the orthodox view that rent-taking is greater in the presence of popularity shocks, given wages, is not necessarily true. Third, we find that in the presence of popularity or ideological shocks rent-taking is increasing in the variance of the shock for given wages, and is decreasing in wages for a given variance of the shock. While our third finding is in line with the directional predictions of the Nash equilibria, the deviation from Nash is large when the variance of the popularity shock is high and wages are low. We show that the deviations can be explained using a Quantal Response Equilibrium approach and taking risk-attitudes into account.

Keywords: electoral competition; laboratory experiment; political rents.

JEL: D72; C91

---

[*] In fond memory of Rebecca Morton and Leif Helland, both astute scholars and warm fellow beings. Tom-Reiel Heggedal would like to acknowledge the helpful comments by the associated editor and three anonymous referees, as well as constructive comments from participants at the ESA conference in Vienna, 2017; the 11th Nordic Conference on Behavioural and Experimental Economics in Oslo, 2016; the 2nd Behavioral and Experimental Public Choice Workshop in Lille, 2018. This research was funded by The Research Council of Norway, grant 250506.

[†] Department of Economics & CESAR, BI Norwegian Business School.

[‡] Department of Economics & CESAR, BI Norwegian Business School.

[§] Department of Politics, NYU NYC & SSEL, NYU Abu Dahbi.

Introduction

Political corruption is a timeless problem that afflicts both developed and developing countries. Although corruption might in some cases be an efficient way to correct pre-existing governmental failures (e.g. Aidt 2003; Besley 2006, pages 55-59), most believe the reduction of corruption is a desirable goal and that political institutions should be designed to make such corruption difficult. In particular, considerable research has argued that competitive elections can inhibit corruption by allowing voters to reward and choose politicians who are less corrupt while sanctioning corrupt ones by ousting them. Research such as Persson et al. (2003) and Rudolph and Däubler (2016) suggests that majoritarian systems which promote greater political competition and alternation in power are less corrupt than societies using proportional representation.

Yet, across majoritarian governments the level of corruption varies. Understanding this variation and thus the extent that majoritarian elections can control corruption is a question that has received much recent focus. Most of this research has concentrated on the effects on politician and voter behavior of providing voters with information about corruption as in Ferraz and Finan (2008). Recent research in this vein finds that voters and politicians will respond to information on corruption when there are competitive alternatives who are perceived as less corrupt. For example, Grossman and Michelitch (2018, page 297) argue that their results "underscore the idea that constituency competitiveness is likely a necessary condition for transparency initiatives to affect politician behavior." And De Vries and Solarz in their recent review of the literature conclude: "Voters can be expected to punish corruption electorally only when viable alternative options are available." (2017, page 404).

Many factors can affect the viability of available alternatives in majoritarian systems such as candidate nomination procedures and the ease in which candidates can enter competition, the use of public campaign financing, the exogenous financial rewards from office that are independent of corruption behavior such as salaries, financial security in retirement, etc., and uncertainty over the extent that voters weigh ideological positions or other characteristics of candidates relative to concerns over corruption. In this paper we focus in particular on the last two of these factors

– the influence of high salaries and wages from holding office and the effects of uncertainty in electoral outcomes due to preference shocks.

We generate our predictions using the canonical probabilistic voting model with rent taking (Polo 1998, Persson and Tabellini 2000:chapter 3). The core trade-off in this model also appears in a number of more complex model environments in which there is some degree of conflict between voters' and candidates' policy objectives.[1] Thus, understanding the extent to which human decision makers recognize and respond to the incentives of the canonical model has the potential to shed light on expected behavior in a set of richer electoral institutions.

In the model corruption is seen as a rent-seeking game where greater rent-taking reduces payoffs to voters by siphoning away tax revenue. Two exogenously given and purely rent motivated candidates simultaneously commit to a level of rent-taking. Voters observe the rent-levels candidates have committed to and vote for one of the candidates. Thus, the model zooms in on the effects of electoral competition among candidates with commonly known reputations for corruption.

In the absence of popularity or ideological shocks electoral competition theoretically drives rational rent-taking to zero, independent of the level of exogenous rewards such as salaries. Things change when a popularity or ideological shock is added to the model. The shock is orthogonal to rents and is resolved after candidates have committed to a rent-level. Seen from the perspective of the candidates, rent-taking now reduces the win-probability smoothly. In equilibrium candidates balance rents and win-probability. Increased variance in the popularity or ideological shock increases rent-taking for given salaries. For a given variance in the shock, increasing politicians' salaries reduces rent-taking. Hence, the theory predicts that with uncertainty higher wages can lead to reduced rent-taking and fewer corrupt choices by elected officials.

As we discuss in the next section, recent empirical research has examined the links between wages and behavior of politicians in order to determine if such a link exists. However, measuring

---

[1] Examples include rent-taking under varying electoral rules (Persson and Tabellini 1999, 2000:chapter 8); the choice of growth policies (Besley et al. 2005, 2010); geographical redistribution (Lindbeck and Weibull 1987, Dixit and Londregan 1996); the effects of valence issues (Ashworth and de Mesquita 2008, Morton and Myerson 2012); the choice of environmental agreements (Battaglini and Harstad 2020); and the effects of mass media (Strømberg 2004), to mention a few.

the effects of rewards and uncertainty on candidate behavior is exceedingly difficult when using field data, particularly if the goal is to establish whether causal relationships exist. Although there has been a notable increase in using experimental approaches in naturally occurring elections to study corruption, it is no surprise that the standard manipulation is focused on providing information or incentives for voters to monitor electoral officials, rather than manipulating the salaries of elected officials or ideological or other characteristics of candidates that might affect uncertainty in elections. Moreover, even if we could conduct such interventions in naturally occurring elections, we would need to control for all the other factors that might affect corruption and confound our ability to discern effects.

In order to be able to make causal inferences by controlling and manipulating these factors and to have a large number of observations of candidate corruption choices, we chose to investigate candidate competition in the laboratory. Though there is a growing experimental literature in political science and political economy, the present study is, to the best of our knowledge, the first to investigate rent-taking in this context.

We have a number of findings. First, rent-taking is not independent of, but decreases with wages in the absence of popularity shocks. Second, in the presence of popularity shocks (which is likely to characterize any election), the directional predictions of the model come through nicely. Rent-taking is increasing in the variance of the shock for given wages, and is decreasing in wages for a given variance of the shock. Third, we observe substantial deviations from the Nash point-predictions, also in the presence of popularity shocks. We argue that the combination of a Quantal Response Equilibrium (which allows for systematic errors in decision-making) and risk averse players generates expectations about deviation patterns. We find that observed deviations match these expectations well.

We do not explicitly test alternative theories of social, psychological or moral motivations. Nevertheless, it is natural to wonder how such other theories relate to our data given the particulars of our design. We discuss these theories briefly at the end of the paper.

Related literature

A number of empirical papers have investigated the relationship between wages and salaries of elected officials and the extent of corruption or other measures of performance. The evidence on the effects of wages is mixed; while effects have been found they seem to not be sizable. On the one hand, Alt and Lassen (2003) find that average US state government relative wages are negatively related to subjective measures of corruption given by a state journalist survey and Coates (1999) finds that there is a positive relationship between salaries of United States legislators and the quantity of legislation passed. However, on the other hand, Coates (1999) also finds no evidence of an effect on economic growth. And in a cross-country study of various corruption indices, Treisman (2000) found no significant relationship between government wages and the measures of corruption. Beylis, Finan, and Mazzocco (2012) estimate a structural model of corruption with data on Brazilian municipalities and consider the extent that a policy experiment of increasing salaries leads to less corruption. They find that doubling wages would have a minor effect on corruption relative to other possible interventions.

This work is largely observational and the relationships observed may not be causal. Some of the more recent empirical work has used exogenous factors that can affect public wages to measure their effects on corruption. Gagliarducci and Nannicini (2013) use a regression discontinuity design to estimate the effects of mayoral wage changes when municipalities in Italy change in size at the 5,000 inhabitant level. They find that mayors with higher wages appear to improve efficiency in government, although they argue that the effect is mainly by a change in selection rather than behavior. Van der Linde et al. (2014) use a similar design with Dutch mayoral wages and find that higher wages lead to increases in debt and budgets. They argue that their results are evidence of better performance of the officials even though the direction of the effects are the opposite from the Italian study. Altindag et al. (2017) find a negative effect of salary increases on the performance of Turkish MPs using a reform in salary structures as a source of identification. The performance measure used is a composite of the number of motions, law drafts, speeches, and words by the MPs. Ferraz and Finnan (2009) study Brazilian municipal governments for which salary caps on elected politicians are conditioned on population size. Using these discontinuities they find that better paid politicians perform better in the sense of producing more legislation and petitions. The effects of higher salaries on public goods production, however, is mixed.

A number of scholars have used the unique opportunity of a large exogenous wage increase for members of the European parliament (Mocan & Altindag 2013, Braendle 2015, and Fisman et al. 2015) to study the effects on politician behavior. That is, beginning in 2009 all members received an identical salary of roughly 90,000 euros when previously salaries had been the same as the lower house of their respective countries and ranged in value from 10,080 to 144,084 euros. These studies provide conflicting and mixed evidence. The most robust result is the evidence that the increase in wages induced more politicians to run for re-election and stay in office longer, but the effects on behavior in office are not robust across authors and specifications.

Model

In this section we set up a parsimonious probabilistic voting model of rent-taking under electoral competition.[2] The environment of this model is used in the experiments where we compare the equilibrium predictions with actual outcomes from the lab.[3] The main feature of the model is that candidates commit to rent extracting policies prior to the election.[4] Platforms are observed by the voters, who vote for the platform offering the lowest rents. The candidate with the most votes wins the election and enjoys rents and the wages from holding office. The intensity of competition varies with the wages of politicians and the level of uncertainty in the electoral mechanism due to popularity or ideological shocks. In the model higher wages for politicians always drive down rents if there is uncertainty in the electoral mechanism. In the absence of such uncertainty there is no rent-taking regardless of the wages offered to the election winner.

---

[2] The model is a simplified version of the "Inefficent Electoral Competition" model of Persson and Tabellini (2000), p70–73. As a minor extension of the original set-up we allow for risk-attitides in the utility functions of the candidates.

[3] The instructions to the experiments were written in a neutral language, without references to politics, voters or rents. These instructions can be found in the Appendix. In a previous experiment on candidate competition, Aragones and Palfrey (2004) find no difference between behavior of subjects in a neutral language game and one that uses political language to describe the game.

[4] Commitment to policies is common in the literature on political competition and ‘pre-election politics’ (as opposed to agency problems), and, in particular, in the literature on probabilistic voting where popularity shocks form the foundation for parties’ rent extraction (Polo 1998; Svensson 1998; Svaleryd & Vlachos 2009; Aldashev 2015; Lind & Rohner 2017).

The model abstracts away from many potentially important sources of corruption, such as e.g., adverse selection, post election moral hazard and search frictions, in order to focus squarely on the effects of electoral competition.[5] The "platforms" and "rent policies" that parties commit to may broadly be interpreted either in terms of strategic selection of party candidates, policies that entail more or less active waste (in the terminology of Bandiera et al. (2009)) by for example changing the checks-and-balances within and across government branches, or outright corruption. So, a party that strategically ends up selecting candidates that can be expected to extract more rents, at the cost of a lower probability of winning the election, is fully consistent with our assumption that parties may commit to high rents (and vice versa for a party that selects candidates which can be expected to extract less rents).

Environment: The economy consists of a continuum of identical voters and two candidates indexed $j \in \{A, B\}$. The mass of citizens is normalized to 1. The only activity in the government is to convert incoming taxes to rents and the wage accruing to the candidate in power.[6] The policy preference of a voter is given by

$$V(r) = y - r - W, \tag{1}$$

where y is the average (total) income of voters, r is the (endogenous) rent, and W is the (exogenous) wage.

Voters also have preferences over some other policy dimension orthogonal to the rent platforms of the candidates. This preference is stochastic and may represent the ideological bias in the

---

[5] There is a small experimental literature on rent-taking in dynamic electoral agency games. The environments investigated in this literature differs in significant ways from the environment studied in our experiment. The literature on electoral agency address issues of moral hazard and adverse selection in a setting were promises are not binding—not the effects of platform competition *per se.* See Helland & Monkerud (2013) for an overview. Woon (2014) study the effects of reputation formation in electoral competition experimentally, departing from an agency perspective. For a recent study of the impact of search frictions on rent taking in electoral competition see Andersen & Heggedal (2019). Four recent empirical studies depart from an agency perspective on electoral competition. Calvacanti et al. (2018) investigate the effect of random audits on candidate selection in Brazilian data. Arias et al. (2019) use Mexican data to investigate the power of social networks in claiming accountability. Chong et al. (2014) implement a field experiment in Mexico to study the relationship between corruption information, turnout, and the vote support of incumbents and challengers. Anduzia et al. (2013) conduct a survey experiment in Spain. Results indicate that corruption perceptions are biased against the respondent's party, but that this bias disappears for high political awareness.

[6] Candidates' incentives to take rents do not change in the model if the provision of a public good is included as an alternative use for tax-revenues.

population or the popularity of a particular candidate. This popularity shock denoted by δ is *not* observed by the candidates prior to the announcement of platforms, though the distribution of δ is common knowledge. We assume that the popularity shock follows a continuous uniform distribution with support $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$, where a positive number means that candidate A is more popular than candidate B. Thus a voter prefers candidate A if

$$V(r_A) + \delta > V(r_B).$$

The candidates have utility over total rewards from holding office, that consist of the wage W and rent-taking r_j. We assume that candidates exhibit constant relative risk aversion. Denote $p_j(r_j, r_{-j}, \delta)$ the probability of winning the election for candidate j given the other candidate's policy and the popularity or ideological shock. A candidate's objective is to set $r_j > 0$ so as to maximize the expected utility

$$E[u(r_j, r_{-j})] = p_j(r_j, r_{-j}, \delta)\frac{(W + r_j)^{1-\alpha}}{1-\alpha}, \tag{2}$$

where α measures the degree of risk aversion.

The candidate with the highest vote share wins the election.[7] The shock measures the extent that voters' choices are likely affected by the rent choices of the candidates. When the shock is potentially large, then the potential effect that an increase in rent-taking has on the probability a candidate wins is less and candidates can potentially get away with more rent-taking since voters care about other candidate characteristics that to a large extent are unaffected by rent-taking. The timing in the model is as follows: 1) The candidates simultaneously and independently set $\{r_A, r_B\}$; 2) The value of δ is realized; 3) The citizens vote; 4) Payoffs are realized.

Characterizing equilibrium: We look for a Nash equilibrium in the game between the candidates, taking voting behavior as given (in the experiment voter behavior is automated). To find the probability of winning the election note that the citizens are indifferent between the two candidates if the popularity shock is such that

---

[7] Note that in the experiments the distribution is discrete (though with a large value set), so there may be ties with positive, albeit tiny, probability.

$$\hat{\delta} = V(r_B) - V(r_A),$$

where $\hat{\delta}$ denotes the value of the shock. If the realized value of the shock is larger than $\hat{\delta}$ candidate A wins, while candidate B wins in the opposite case. Thus we can write the probability of winning for candidate B as

$$\begin{aligned} p_B &= \Pr_B\{\hat{\delta} \leq \delta\} \\ &= \Pr_B\{V(r_B) - V(r_A) \leq \delta\} \\ &= \Pr_B\{r_A - r_B \leq \delta\} \\ &= \begin{cases} 0 & \text{if } r_B > \frac{1}{2\psi} + r_A \\ \psi(r_A - r_B) + \frac{1}{2} & \text{if } -\frac{1}{2\psi} \leq r_A - r_B \leq \frac{1}{2\psi} \\ 1 & \text{if } r_B < -\frac{1}{2\psi} + r_A \end{cases}, \end{aligned}$$

and the probability of winning for candidate A is given by $p_A = 1 - p_B$.

Turning to the equilibrium strategies of the candidates, first note that an equilibrium does not exist where $p_j = 1$, as in this case candidate $-j$ would get zero and thus have an incentive to lower r_{-j}. Then, focusing on the interior of p_B, the first order condition for candidate B is

$$W + r_B = \frac{(1-\alpha)}{\psi} p_B$$

The first order condition for candidate A is symmetric to that of B, thus the unique equilibrium is given by

$$r_A = r_B = r^{eqm} = \max\left[0, \frac{1-\alpha}{2\psi} - W\right].$$

Note that we assume $y \geq \frac{1-\alpha}{2\psi}$ so equilibrium rents are not constrained by available resources. When uncertainty evaporates in the model ($\psi \to \infty$), competition for W stiffens, and in equilibrium rent-taking is competed down to zero. More generally, a larger ψ increases the competitive pressure between the candidates as there is 'less uncertainty' as to the effect of decreasing rent-seeking on the probability of winning election. Note that the density is inversely related to the bound of the support, so r is increasing in the width of the support $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$. Also, note that a higher density lowers the variance, i.e., there is less (more) uncertainty in the society when ψ increases (decreases).[8] With some uncertainty in the electoral mechanism, a larger W increases the competitive pressure as there is a larger 'prize' for the winning candidate. Thus, with uncertainty in the electoral mechanism lower ψ and/or W (i.e., reduced competitive pressure) reduces rent-taking. Due to the lower bound on r there are many combinations of ψ and W that give $r = 0$.[9] Last, a larger α reduces rent-taking as the candidates are less willing to take on the risk of not winning associated with higher rent-taking.

Design and procedures

Design: The experiment was implemented as a series of 50 rent-taking games.[10] The fixed number of games was public knowledge, and subjects played in one treatment only. Voters were automated and language was neutral. At the start of the experiment subjects were randomly allocated to labels A or B and kept their labels through the experiment. At the beginning of each game pairs of opposite labels were randomly formed from fixed blocks of 8 subjects ("typed stranger matching"). After a match had formed subjects simultaneously choose numbers (r_j, r_{-j}) in the range 0 to 100. A random number $\tilde{z}$ was drawn from a uniform distribution with mean $= 0$ and publicly known support on $-Z$ to $+Z$ $\left(Z := \frac{1}{2\psi}\right)$. The payoff to A was $\pi_A = r_A + W$ if $r_A < r_B + \tilde{z}$, and $\pi_A = 0$ if $r_A > r_B + \tilde{z}$. The payoff to B was $\pi_B = r_B + W$ if $r_B + \tilde{z} < r_A$, and $\pi_B = 0$ if $r_B + \tilde{z} > r_A$. If $r_A = r_B + \tilde{z}$ a fair lottery determined the winner, which got $\pi_j = r_j + W$, while the loser got $\pi_{-j} = 0$.

The experiment consisted of five treatments in total. The treatments are depicted in table 1, were W and r^* are measured in experimental currency units (ecu) and r^* denotes risk neutral

---

[8] The variance of the distribution is given by $\sigma^2 = \frac{(2Z)^2}{12}$, with $Z := \frac{1}{2\psi}$.

[9] Note that without a lower bound on r, the endogenous rent would be negative if $\frac{1-\alpha}{2\psi} < W$. In this case the candidates compete to get W by offering to share some of this surplus of winning with the citizens.

[10] A concern is that using convenience samples of students (CS) is not informative of candidate behavior. Unfortunately, laboratory experiments using political elites as subjects are sparse. Hafner-Burton et al. (2014), however, provide evidence indicating broad similarity between CS and a sample of elite politicians with respect to depth of reasoning (measured by a p-beauty contest) and patience (measured by choice of immediate versus delayed gratifications). LeVeck et al. (2014) compare a CS to a sample of elite politicians with respect to bargaining behavior (measured by an ultimatum game). They find elite politicians to be somewhat more generous as proposers, and somewhat less inclined to accept meager offers as responders. We also note that Armantier & Boly (2008) in their study of corruption find that external validity is not violated when moving from a lab experiment in a developed country to a field experiment in a developing country.

($\alpha = 0$) Nash equilibrium rent-taking from the model. This design allows us to cleanly identify the causal effects on rent-taking behavior by manipulating the uncertainty in the electoral mechanism and the level of exogenous rewards facing candidates.

Treatment	Z	W	r^*	# Subj.	# Blocks
$T1$: *Low Wage*	0	1	0	48	6
$T2$: *Medium Wage*	0	10	0	48	6
$T3$: *Low Uncert Low Wage*	10	1	9	48	6
$T4$: *High Uncert Low Wage*	50	1	49	48	6
$T5$: *High Uncert High Wage*	50	40	10	48	6

Table 1: *Treatments of the experiment.*

In our two first treatments there is no uncertainty in the electoral mechanisms. Even though wages increase by ten fold when moving from treatment *Low Wage (T1)* to treatment *Medium Wage (T2)*, predicted rent-taking in the Nash equilibrium remains zero. In our three last treatments the electoral mechanism is noisy. In *Low Uncert Low Wage (T3)* noise is moderate and wages are low. In *High Uncert Low Wage (T4)* noise is high while wages remain low. In *High Uncert High Wage (T5)* both noise and wages are high. In Nash equilibrium we would expect rents to increase linearly (and almost proportionally) with increasing noise as we move from *Low Wage (T1) or Medium Wage (T2)* to *Low Uncert Low Wage (T3)* and on to *High Uncert Low Wage (T4)*. Comparing the two high noise treatments, rent-taking in Nash equilibrium should fall by almost a factor of five as we move from *High Uncert Low Wage (T4)* to *High Uncert High Wage (T5)*. This effect is due to the almost five fold increase in wages. While directional predictions do not change with the level of risk aversion as long as $\frac{1-\alpha}{2\psi} > W$ the level of rent-taking decreases with risk-aversion (α) in noisy treatments. In summary, in addition to the point predictions of rent-taking in Table 1 above the theoretical model predicts the following relationship predictions:

Relationship Prediction 1: *There should be no difference in rent-taking as wages increase when there is no uncertainty: Low Wage (T1) = Medium Wage (T2).*

Relationship Prediction 2: *When uncertainty increases but wages stay constant, rent-taking* *should be higher: High Uncert Low Wage (T4) > Low Uncert Low Wage (T3) > Low Wage (T1).*

Relationship Prediction 3: *When wages increase while uncertainty is positive but constant, rent-taking should be lower: High Uncert Low Wage (T4) > High Uncert High Wage (T5).*

Procedures: All sessions were conducted in the research lab of BI Norwegian Business School using participants recruited from the general student population at the BI Norwegian Business School and the University of Oslo, both located in Oslo, Norway.[11] Recruitment and session management were handled via the ORSEE system (Greiner 2015). No subject participated in more than one session. z-Tree was used to program and conduct the experiment (Fischbacher 2007). Anonymity of subjects was preserved throughout. On arrival, subjects were randomly allocated to cubicles in the lab in order to break up social ties. After being seated, instructions were distributed and read aloud in order to achieve public knowledge of the rules. Instructions and screen shots are provided in online appendixes H to K.

Given that our experimental data is never connected to subject identifying information the project does not require full ethical review according to Norwegian legislation. A pre-study plan was registered on the 24^{th} of May 2016 at the Randomized Controlled Trails Registry of the American Economic Association.[12] The analysis below follows the specifications of this plan. Our pre-study plan included a pilot study. The pilot indicated that a power of 90 percent or better would require 6 blocks per treatment. In accordance with this we used a total of 240 subjects in our experiment: 5 treatments à 6 blocks à 8 subjects per treatment.[13] In total the experiment collected data on $12,000$ individual decisions. A session of the experiment took approximately 45 minutes, for an expected average payment of 250 NOK (approximately 30 USD at the time of the sessions).

---

[11] We define a session as a collection of subjects in the laboratory at a specific date and time. Note that a session in our experiment has more than one block.

[12] `https://www.socialscienceregistry.org/trials/1289`

[13] Power was calculated using the code of Belamare et al. (2016).

Results

Directional predictions

Our *treatment variables* are the support of the popularity shock (Z) and the size of wages (W). Our *treatment measure* is the average difference in the level of rent-taking over treatments. We perform two kinds of tests. Firstly, we report p-values (p^p) from treatment regressions. These regressions have random effects for unique subjects and standard errors clustered at the unique block level. The units of observation are individual decisions.[14] Secondly, we report p-values from non parametric (Wilcoxon rank sum) tests (p^n) using average behavior in a block over all 50 games as units of observation. Details of the tests are contained in the online appendices A and B.

All promises: Figure 1 shows the mean, the Nash prediction and the confidence interval of all rent taking promises by treatment.

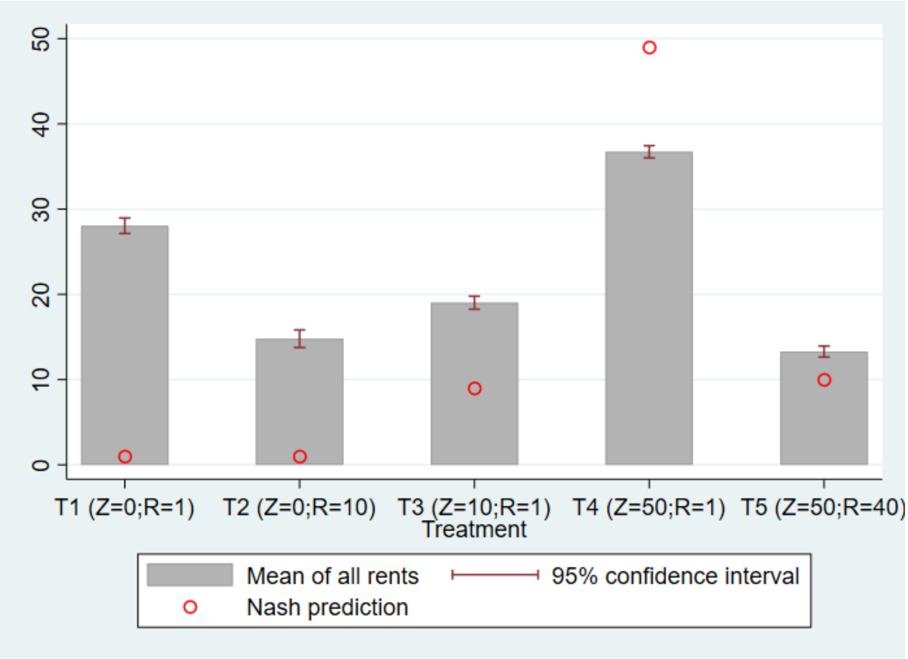


Figure 1: All rent taking promises.

---

[14] To the extent that the parametric and non-parametric tests agree we can rule out that large intra-block variances are masked by the non-parametric rank tests (that does not take intra-block variances into account).

Treatments in general deviate substantially from the Nash-prediction when considering all rent promises. In the absence of noise we see substantially lower rent-taking when wages are increased from 1 to 10 ecu: $(\textit{Low Wage } (T1) - \textit{ Medium Wage } (T2)) = 13.3$ ecu. The difference is highly significant in the parametric test, and almost significant at conventional levels in the non-parametric test $(p^n = .110; p^p = .000)$. As noted, this is contrary to Relationship Prediction 1 above. The orthodox view that rents are competed away in the absence of noise is also not confirmed in the data, contrary to the point predictions of Table 1 and Relationship Prediction 2 above: $(\textit{Low Wage } (T1) - \textit{ Low Uncert Low Wage } (T3)) = 9.0$ with $(p^n = .013; p^p = .000)$.

Consider now the noisy environment with low wages. As noise increases given $W = 1$ ecu, rent-taking in Nash equilibrium should be greater supporting Relationship Prediction 2 in this case. This directional prediction is manifested in the data, and is highly significant: The treatment difference $(\textit{Low Uncert Low Wage } (T3) - \textit{ High Uncert Low Wage } (T4)) = -17.7$ ecu with $(p^n = .003; p^p = .000)$. Lastly, in the high noise environment $(Z = 50)$ a substantial increase in wages results in sharply lower rent taking, and this effect is highly significant: The treatment difference $(\textit{High Uncert Low Wage } (T4) - \textit{ High Uncert High Wage } (T5)) = 23.4$ ecu with $(p^n = .002; p^p = .000)$, supporting Relationship Prediction 3.

Winning promises: Figure 2 shows the mean, the Nash prediction and the confidence interval of winning rent taking promises by treatment.

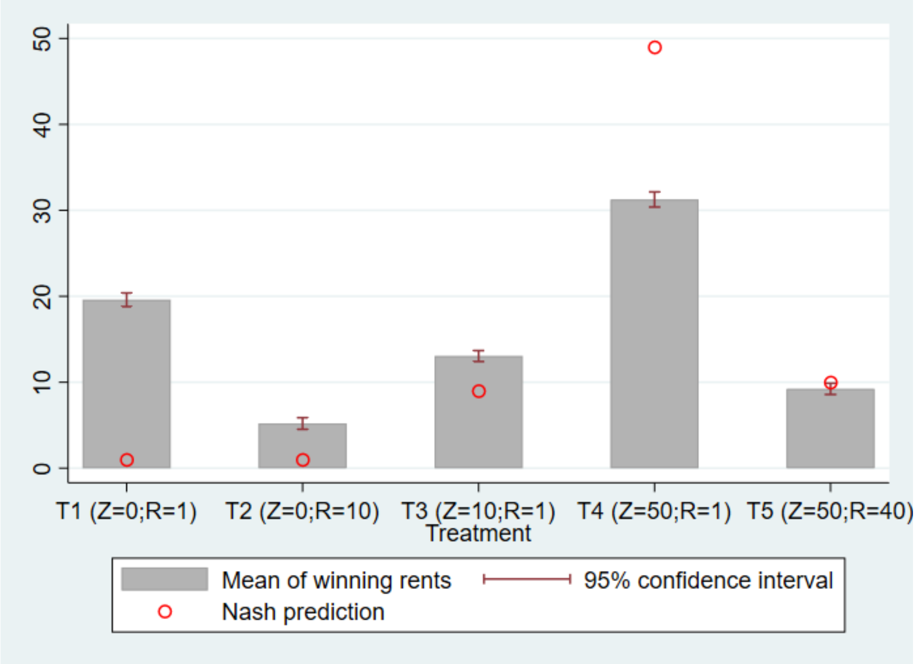


Figure 2: Winning rent taking promises.

Except for the *High Uncert Low Wage* treatment ($T4$), behavior deviates less from the Nash-prediction when looking at winning rent promises only. In the absence of noise we again observe substantially lower rent-taking as wages are increased tenfold: (*Low Wage* ($T1$)$-$ *Medium Wage* ($T2$)) $= -14.4$ ecu. For winning promises the difference is highly significant in both tests ($p^n = .004; p^p = .000$). Again the orthodox view that rents are fully dissipated in the absence of noise is not confirmed in the data: (*Low Wage* ($T1$)$-$ *Low Uncert Low Wage* ($T3$)) $= 6.6$ ecu with ($p^n = .013; p^p = .000$). Consider then the noisy environment with low wages. The Nash directional prediction is manifested in the data, and is highly significant: The treatment difference (*Low Uncert Low Wage* ($T3$)$-$ *High Uncert Low Wage* ($T4$)) $= -18.2$ ecu with ($p^n = .002; p^p = .000$). Lastly, in the high noise environment ($Z = 50$) the four fold increase in wages once again results in sharply lower rent-taking, and the effect is highly significant: The treatment difference (*High Uncert Low Wage* ($T4$)$-$ *High Uncert High Wage* ($T5$)) $= 23.0$ ecu with ($p^n = .002; p^p = .000$).

The patterns in the data remain qualitatively unchanged if we only consider the final 25 periods of the game, after learning has taken place and behavior has settled (see the online appendices C (for all rent promises) and D (for winning rent promises)).

So, to sum up, the model works nicely in terms of directional predictions as long as there is some noise in the electoral mechanism (as is likely to be the case in any election): Rent-taking is increasing in the variance of the shock for given wages, and is decreasing in wages for a given variance of the shock. In the absence of popularity shocks the directional predictions fail: Rent taking is now decreasing in wages.

Observing that rents are lower when wages are higher, one may ask whether higher wages pay for themselves. We propose estimates of the return to increasing wages to shed some light on this question. In the absence of noise, we obtain the return to wages $\frac{\Delta r|Z=0}{\Delta W} = -1,6$ by comparing winning rents in *Low Wage* ($T1$) with winning rents in *Medium Wage* ($T2$) where W is higher. In the presence of noise, we obtain the return to wages $\frac{\Delta r|Z=50}{\Delta W} = -0,57$ by comparing winning rents in *High Uncert Low Wage* ($T4$) with winning rents in *High Uncert High Wage* ($T5$) where W is higher but Z is the same. Thus whether higher wages pay for themselves in our experiment depends on the noisiness of the environment.

Time path of play

Figure 3 shows the time path of play in our five treatments. For each of the 50 games played average rent-taking over the 6 blocks in a treatment is plotted. We plot both for all rent promises and for winning rent promises. Visual inspection suggests that behavior converges from above in *Low Wage* ($T1$) to *Low Uncert Low Wage* ($T3$) and in *High Uncert High Wage* ($T5$), and stabilizes within 10-15 periods. In *High Uncert Low Wage* ($T4$) there seems to be no such pattern of convergence. In online appendix E we perform formal tests of convergence, these tests confirm the visual inspection. Since, the Quantal Response Equilibrium concept we employ is a static one, we use data only from the final 25 periods of the experiment in this analysis, to make sure subjects have been given ample time to learn and that behavior has converged.

Figure 3: Time-paths of play by treatment.

Deviation from Nash

There are three deviations from the Nash predictions that are particularly noteworthy. First, in absence of noise wages seem to lead to lower rent-taking behavior (*Low Wage* ($T1$) vs *Medium Wage* ($T2$)). Second, an environment with some noise seems to induce less rent taking than an otherwise similar environment with no noise (*Low Wage* ($T1$) vs *Low Uncert Low Wage* ($T3$)). Third, in a high noise environment low wages are not fully offset by greater rent-taking (*High Uncert Low Wage* ($T4$) vs *High Uncert High Wage* ($T5$)).

Risk aversion alone cannot jointly explain these observed deviations from Nash equilibrium. Risk aversion has no impact on rent taking in the absence of noise as the lower bound on r binds. In the noisy environment increasing risk aversion reduces equilibrium rent-taking. This improves on the observed deviation from the Nash-prediction in the *High Uncert Low Wage* treatment ($T4$), but increases the deviation from the Nash predicitons in both the *Low Uncert Low Wage* ($T3$) and the *High Uncert High Wage* ($T5$) treatments.

In order to better understand the observed deviations from Nash equilibrium we analyze noisy play using the concept of Quantal Response Equilibrium (QRE) to candidates' rent promises. In our estimations we also augment the QRE with risk averse players. QRE has been successfully applied in the experimental literature to rationalize deviations from Nash outcomes in various games.

The structural approach to QRE, introduced by McKelvey and Palfrey (1995), is based on a random payoff model, where the utility u_j of a candidate j, given the other candidate's cumulative distribution function for rent-taking strategies, F_{-j} , is perturbed by a random error interpreted as trembles: $\hat{u}_j(r, F_{-j}) = E_{r_{-j}} u_j(r, F_{-j}) + \epsilon_{j,r}$. Each player assigns a probability to a given action equal to the probability that this action is a best response given the error. The resulting quantal responses can be interpreted as noisy best responses. In equilibrium players' beliefs about the opponents mixing probabilities are correct. While this equilibrium requirement puts high demands on the rationality of the players if taken literally, the resulting rule for the mixing probabilities is very intuitive: the probability of choosing an action increases with its expected payoff. If a Gumbel distribution for the error is assumed, i.i.d. across actions and players, the quantal response is of the logistical form. Given symmetry, as in our games, the quantal responses are given by the (identical) distribution function over strategies for each candidate, $F^Q(r)$, that solves the following functional fixed point:

$$F^Q(r) = \frac{\int_0^r \exp(\lambda E u_j(r, F^Q(x))dx}{\int_0^{100} \exp(\lambda E u_j(r, F^Q(x))dx} \quad \forall r \in [0, 100],$$

where $\lambda > 0$ is the parameter governing trembles (where $1/\lambda$ is interpreted as trembles or noise). In using a logistic specification we follow the majority of the experimental literature, making our findings comparable. This specific choice of the distributional form puts discipline on the resulting QRE.[15] When λ approaches zero, all rents are equally likely, which can be interpreted as completely noisy strategies. On the other extreme, if λ goes to infinity, the quantal response approaches the best response of the underlying rent-taking game, and behavior converges to the

---

[15] See Haile et al. (2008), Goeree et al. (2005), and Goeree et al. (2016:42-53) for discussions of the issue of falsifiability of QRE.

Nash equilibrium. To evaluate whether QRE can rationalize the data we fit the QRE cumulative distribution functions of all rent-taking promises to the corresponding CDFs in data by choosing a λ to minimize the sum of squared deviations. In our estimations we used a grid of 100 rent levels, and minimized the distance between the data and QRE using a maximum likelihood procedure.[16] Since strategic tensions vary much over the five games underlying our treatments there is no particular reason to believe that players tremble in a comparable way over treatments. In line with this approach we estimated the QRE separately for each treatment.[17]

The QRE logic imposes a number of restrictions on expected behavioral patterns in our experiment. We highlight four of them here. One should expect: i) that in the absence of noise deviations from Nash will be higher with low wages than with high wages (*Low Wage* ($T1$) vs. *Medium Wage* ($T2$)); ii) that upward deviations from Nash will be higher in the absence of noise than in the presence of noise (*Low Wage* ($T1$) vs. *Low Uncert Low Wage* ($T3$)); iii) that in the presence of noise deviations from Nash will be more pronounced in our experiment than has been found in experiments on noisy electoral competition with ideological candidates, and; iv) that with risk averse players observed behavior will undershoot the Nash prediction if noise is high and wages low (i.e., *High Uncert Low Wage* ($T4$)). We elaborate on these expectations more fully as we discuss our results below.

Figure 4 displays the QRE estimates assuming subjects are risk neutral. Risk-aversion can be accounted for in the QRE estimates in various ways.[18] One approach is to estimate the trembles ($1/\lambda$) of the QRE jointly with risk attitudes (α). Such an approach runs the risk of over fitting. Our approach consists in transforming payoffs, using available α estimates in the experimental literature, prior to estimating the QRE. Specifically, we use the midpoint of the two structural estimates in Harrison & Ruström (2008:69-74). That is, prior to estimating the QRE we scale

---

[16] The estimation procedure used is what Goere et al. (2016:154) refers to as the "the equilibrium correspondence approace", and amounts to searching for the λ that solves the fixed point problem of the QRE. Matlab codes are available on request.

[17] However, as discussed in the end of this section, we also carry out robustness tests estimating the decision errors simultaneously for all treatments and jointly estimating risk-attitudes and decision errors.

[18] One might ask whether it would have been a better idea to use the lottery procedure (Roth & Malouf 1979) in order to remove risk aversion in the first place. We decided against this since the empirical evidence on whether the procedure acheives what it sets out to do is at best mixed (Berg et al. 2008).

the material payoffs using a CRRA utility function with $\alpha = \frac{3}{4}$.[19] Figure 5 displays these results. Note that the fit of the QRE improves in every treatment when payoffs are adjusted for risk-aversion, and that the improvement is particularly pronounced in *Medium Wage* ($T2$) and *High Uncert Low Wage* ($T4$).

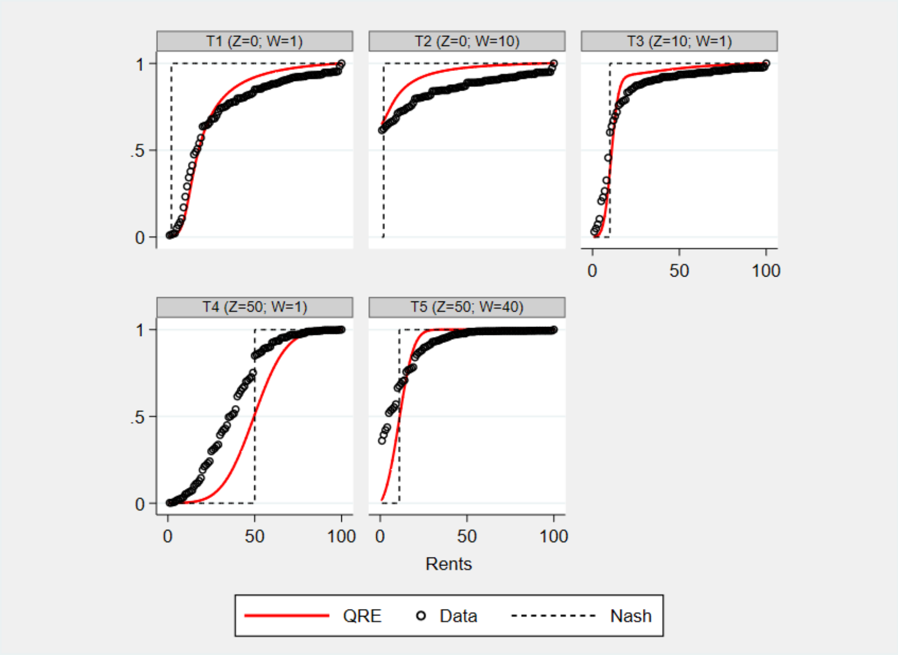


Figure 4: CDFs of data, QRE with $\alpha = 0$, and Nash, by treatment.

---

[19] Using a probit (logit) approach and the data from Hey & Orme (1994), Harrison & Ruström (2008) obtains an estimate of $\hat{\alpha} = 0.66$ ($\hat{\alpha} = 0.80$) with a standard error of 0.04 (0.04).

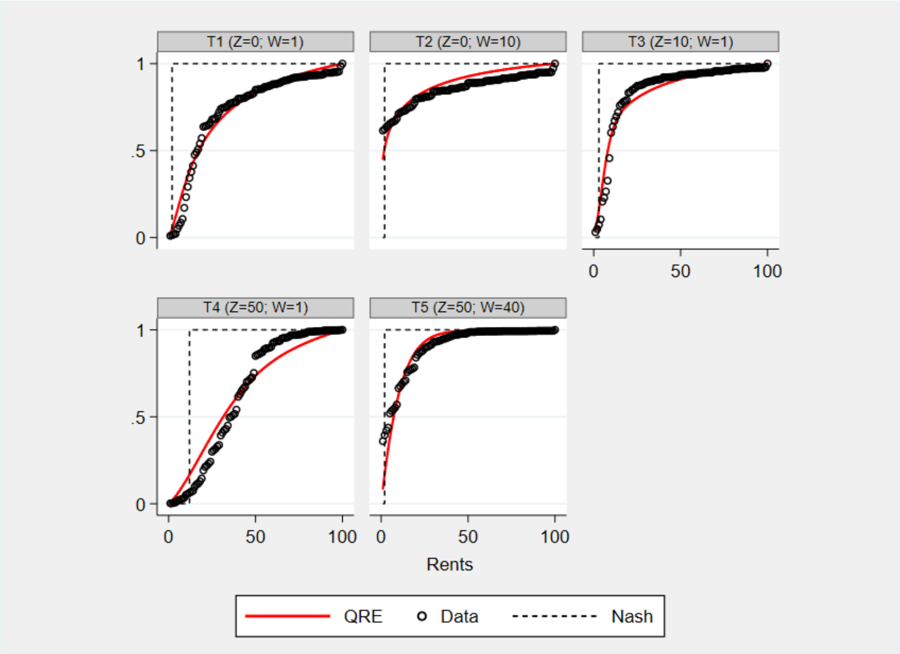


Figure 5: CDFs of data, QRE with $\alpha = \frac{3}{4}$, and Nash, by treatment.

Focus on Figure 5. Consider first the environments without noise ($T1$ and $T2$). The large deviation from Nash observed in *Low Wage* ($T1$) is expected. The game in this treatment resembles a Bertrand price competition with a low price floor (Helland et al. 2017 with references), and incentives to stay in the Nash equilibrium are weak as subjects have (close to) alternative best responses. In this situation some suspicion that the opponent will tremble in his rent-taking suffices to motivate an increase in one's own rent-taking. In QRE this effect induces an upwards spiral in rent-taking.

This upwards spiral in rent-taking is muted when a larger expected payoff from staying in the Nash equilibrium is introduced in *Medium Wage* ($T2$).[20] Thus the smaller deviation from Nash in *Medium Wage* ($T2$) compared to in *Low Wage* ($T1$) is expected from the logic of the QRE.[21] These results are also reflected in the estimated values of the trembles ($1/\lambda$). In *Low Wage* ($T1$)

---

[20] In *Low Wage* ($T1$) expected cost of deviating from the Nash equilibrium is $\frac{1}{2}$ ecu, compared to 5 ecu in *Medium Wage* ($T2$).

[21] In the first 10 periods our results are comparable to those in Dufwenberg et al. (2007), which investigates a 10 period Bertrand duopoly with a price floor of 10 ecu. However, in the last 40 periods behavior stabilizes at a higher level in our experiment.

the estimated trembles are larger than the ones found in *Medium Wage* ($T2$), indicating that decision-making is more noisy when wages are low.[22]

Consider now the noisy environments ($T3$, $T4$ and $T5$). A substantial literature addresses two candidate position games over a single dimension with ideological or popularity shocks and candidates who are policy-motivated (e.g. Morton 1993 with references). In a recent study Drouvelis et al. (2014) demonstrate divergence from the median voter in environments with a popularity shock and (asymmetric) weights on office -and policy motivation. In most of their treatments platforms converge on, or are very close to, the equilibrium prediction. Convergence in our noisy rent-taking games is less impressive.

However, in our rent-taking games losses are linear in rents foregone while losses are quadratic the positioning games. Quadratic losses discipline subjects to stay closer to their Nash best responses in QRE (larger deviations become more costly, and appear with lower probability in QRE). Thus the larger deviations from Nash observed in our noisy rent-taking games compared to noisy positioning games are in line with QRE predictions.

As noted, deviation from Nash is particularly pronounced in the high noise environment with low exogenous rents ($T4$). In this treatment incentives are weak and behavior has fairly little impact on the probability of winning. Comparing Figures 4 and 5 suggests an explanation for the large deviation in *High Uncert Low Wage* ($T4$). The QRE alone does not match the observed data well. The combination of weak incentives and high noise should make us expect rent-taking to be symmetrically distributed around 50 ecu if subjects are risk neutral in QRE. Adding risk aversion to the QRE shifts expected rent taking downwards and skews the CDF of the QRE to the left to fit the observed mean of 36.5 closely.

We also note that the estimated values of the trembles ($1/\lambda$) are larger in *High Uncert Low Wage* ($T4$) than in *Low Uncert Low Wage* ($T3$) and *High Uncert High Wage* ($T5$).[23] That

---

[22] With $\alpha = 0$, the estimated values of lambda are $0,44$ in *Low Wage* ($T1$) and $1,16$ in *Medium Wage* ($T2$). With $\alpha = \frac{3}{4}$, the estimated values of lambda are $0,54$ in *Low Wage* ($T1$) and $1,11$ in *Medium Wage* ($T2$).

[23] With $\alpha = 0$, the estimated values of lambda are $0,95$ in *Low Uncert Low Wage* ($T3$), $0,24$ in *Uncert Low Wage* ($T4$), and $1,11$ in *High Uncert High Wage* ($T5$).

With $\alpha = \frac{3}{4}$, the estimated values of lambda are $0,96$ in *Low Uncert Low Wage* ($T3$), $0,36$ in *Uncert Low Wage* ($T4$), and $1,18$ in *High Uncert High Wage* ($T5$).

decision-making is found to be more noisy in *High Uncert Low Wage* ($T4$) is in line with the QRE predictions as the incentives to stay in the Nash equilibrium are very weak in this treatment compared to in the other treatments.[24]

In online appendix F we perform some robustness tests with respect to the estimation of the QRE. First, we impose a risk aversion parameter of $\alpha = 0.75$ but force decision errors to be identical over treatments. Predictably, this worsens the fit with data somewhat in some treatments (notably in $T3$ and $T4$), but not dramatically. Secondly, we estimate the risk aversion parameter and the identical decision error jointly. This returns a lower risk aversion parameter ($\alpha = 0.35$) than the one imposed for the estimations in Figure 5. Interpreting this difference substantially, however, is hard due to the joint estimation of risk attitudes and the decision error. Of note, the overall fit of the QRE to data in Figure 5 and in this exercise appear highly comparable. We conclude that the overall picture is a QRE that is quite robust to joint estimation of decision errors for all treatments, as well as joint estimation of decision errors and risk attitudes.

Discussion

Our results provide important evidence of the effects of increasing wages on corruption as measured by rent-taking and the mechanisms underlying such effects. The finding that rent-taking is significantly greater than zero without noise demonstrates how attractive corruption through rent-seeking can be even when the consequence of rent-seeking on the probability of winning elections is sizable. Our finding that with noise, increasing wages can reduce corruption as measured by rent-taking implies that higher rewards from winning office likely works by increasing the costs to candidates of engaging in corruption for a given degree of uncertainty in the effects on the probability of winning. So increasing wages can have a negative causal effect on corruption regardless of the degree of uncertainty in the effects of corruption on the probability of winning, and the likely mechanism is to increase the costs borne by candidates if engaging in corruption.

---

[24] We have also estimated lambda using observations from all periods of play (as opposed to the last 25 periods). Comparing estimated trembles treatment by treatment we find that trembles are smaller when using only the last 25 periods. This is in line the QRE as behavior converges over time in our experiment. Results are available upon request.

Many theories of such costs exist. However, aspects of our design renders the more prominent alternatives unlikely candidates to explain our data.

Costs of losing the election: One might suspect that the negative effect of wages on rent-taking is driven by efficiency wages (Becker & Stigler 1974, Akerlof 1982).[25] The argument requires a non-contractible effort stage (i.e. incomplete contracts) in which election winners can be induced to expend costly effort by fear of losing valuable office in future elections. The higher the wages, the more the incumbent stands to lose by misbehaving and the less she will engage in corruption. In the set-up we use there is no effort stage for incumbents, and thus no room for a strict efficiency wage argument. Still, also in our model higher wages reduce rent-taking due to the fear of losing a more valuable office. In outcomes this resembles the efficiency wage argument. However, the underlying mechanism is not disciplining of misbehavior but platform competition.

Costs of not reciprocating kind acts: Could the negative effect of wages on rent-taking be due to a gift exchange relationship (Akerlof 1984, Fehr et al. 1993)? The basic idea is that politicians choose to abstain from misbehaving to reciprocate generosity from voters in the form of high wages. Again, the mechanism requires an effort stage in which incumbents are free to reciprocate. Thus, the absence of such a stage in our set-up shuts down the gift exchange channel. Furthermore, attempts by candidates to reciprocate high wages by low rent-taking would give little meaning in our setting even in the presence of an effort stage, given that wages are exogenously set and voters are computer programs and not humans.

Costs of deviating from a fair division: Potentially, the observed deviations from Nash equilibrium in our experiment could be driven by outcome based fairness preferences (Fehr & Schmidt 1999; Bolton & Ockenfels 2000). However, such preferences are unlikely to have traction in our experiment for at least three reasons. First, the games we use are symmetric and the resulting equilibria are fair in the sense of giving candidates an equal share of resources in expectations. This holds both for the model with and without errors in decision making. Second, and more generally, ample experimental evidence has documented that fairness preferences are not likely to

---

[25] Coates (1999) applies the logic to politics.

impact behavior in highly competitive environments in which contracts are complete (e.g. Smith 1982, Plott 1989). Third, the conflict between voters and candidates is the one that relates most directly to the effect of wages on rent-taking. By reducing rent-taking when wages are high the distribution between voters and candidates becomes more even. But in our set-up there are no relevant distributive conflicts between candidates and voters given that voters are computer programs and not humans.

Costs of inefficiencies: Lastly, is it conceivable that subjects disregard the welfare of the experimenter but are motivated by efficiency concerns, possibly in combination with fairness preferences (Engelmann & Strobel 2004)? If this is the case they would want to extract maximal rents. All else equal, the incentive to deviate from a collusive agreement is stronger the higher is the wage. It is possible that the negative relationship between wages and rent-taking reflects the difficulty of establishing and maintaining collusive agreements for high wages. To succeed in colluding, however, challenging coordination problems must be overcome. Since pairs of candidates are randomly formed in each new game in our experiment such collusion is unlikely to succeed. In online appendix G we substantiate the claim that collusion does not explain our observations by further analysis of data.

Conclusion

In this paper we have investigated in the laboratory a simple rent-taking model of electoral competition in order to study the effects of wages on corruption. We find that higher wages significantly reduce corruption in the experiments, even when unpredicted by the orthodox theory, but in line with QRE predictions. We find that a major factor in how likely electoral competition reduces corruption is also the degree to which there is noise in the effects of corruption choices on the probability of winning. But the effects are not consistent. Holding wages constant, when noise is introduced it actually reduces corruption initially, but for higher levels of noise, more noise increases corruption.

Our results may partially explain some of the mixed results in field studies of the effects of wages on corruption given the inconsistent effects of wages depending on noise levels. Furthermore, we find that in a noisy environment while increases in wages can reduce corruption, the effects are not nearly as strong as the orthodox theory would predict. Thus, it may not be surprising that the effects of wages on corruption can vary significantly across countries and electoral environments.

From our results we appreciate that lower wages cause higher rent-taking both in the presence and absence of popularity shocks. We note an interesting pattern in observational data. While increasing salaries seem to improve politicians performance at the local level (Gagliarducci and Nannicini 2013, Van der Linde et al. 2014, Ferraz and Finnan 2009) there is no clear evidence of such a relationship at the national or supranational level (Altindag et al. 2017, Mocan & Altindag 2013, Braendle 2015, Fishman et al. 2015). This pattern might reflect higher expected costs of misbehaving due to more effective non-electoral sanctions at these at national and supranational levels. In our controlled laboratory environment rent-taking behavior remains fully anonymized and the only sanction for misbehavior is the electoral one. An interesting future avenue might be to introduce third party auditing and sanctioning in the rent-taking model.

References

Aidt, T. (2003): Economic analysis of corruption: a survey. *Economic Journal* 113(491):F632-F652

Akerlof, G. (1982): Labor contracts as partial gift exchange. *Quarterly Journal of Economics* 97(4):543-569.

Akerlof, G. (1984): Gift exchange and efficiency-wage theory: Four views. *American Economic Review* 74(2):79-83.

Aldashev, G. (2015): Voter turnout and political rents. *Journal of Public Economic Theory.* 17(4):528-552.

Alt, J. & D. Lassen (2003): The political economy of institutions and corruption in American states. *Journal of Theoretical Politics* 15(3):341-365.

Altindag, D., S. Filiz & E. Tekin (2017): *Does It Matter How and How Much Politicians Are Paid?* IZA DP No. 10923

Andersen, J. J. & T.-R. Heggedal (2019): Political Rents and Voter Information in Search Equilibrium. *Games and Economic Behavior* 114:146-168.

Anduiza, E., A. Gallego & J. Muñoz (2013): Turning a Blind Eye: Experimental Evidence of Partisan Bias in Attitudes Toward Corruption. *Comparative Political Studies* 46(12):1664-1692.

Aragones, E. & T. Palfrey (2004): The effect of candidate quality on electoral equilibrium: An experimental study. *American Political Science Review* 98(1):77-90.

Arias, E., Balán, P., Larreguy, H., Marshall, J., & Querubín, P. (2019): Information provision, voter coordination, and electoral accountability: Evidence from Mexican social networks. *American Political Science Review*, 113(2), 475-498.

Armantier, O. & A. Boly (2008): *Can corruption be studied in the lab? Comparing a field and a lab experiment*. CIRANO-Scientific Publications: 2008s-26.

Ashworth, S. & B. De Mesquita (2008): Elections with platform and valence competition. *Games and Economic Behavior* 67(1):191-216.

Bandiera, O., A. Prat & T. Valletti (2009): Active and Passive Waste in Government Spending: Evidence from a Policy Experiment. *American Economic Review* 99(4):1278-1308.

Battaglini, M. & B. Harstad (2020): The Political Economy of Weak Treaties. *Journal of Political Economy* 128(2): 544-590.

Becker, G. & G. Stigler (1974): Law Enforcement, Malfeasance and Compensation of Enforcers. *Journal of Legal Studies* 3:1-18.

Bellemare, C., L. Bissonnette & S. Kröger (2016): Simulating power of economic experiments: the powerBBK package. *Journal of the Economic Science Association* 2(2):157–168.

Berg, J., T. Rietz & J. Dickhaut (2008): On the performance of the lottery procedure for controlling risk preferences. In C. Plott & V. Smith (eds.): *Handbook of Experimental Economics Results. Volume 1.* Amsterdam: North-Holland.

Besley, T., T. Persson, D. Sturm (2010): Political competition, policy and growth: theory and evidence from the US. *The Review of Economic Studies* 77(4):1329-1352.

Besley, T. (2006): *Principled Agents? The Political Economy of Good Government.* Oxford: Oxford University Press.

Besley, T., T. Persson & D. Sturm (2005): Political competition and economic performance: Theory and evidence from the United States. National Bureau of Economic Research.(No. w11484)

Beylis, G., F. Finan & M. Mazzocco (2012). *Understanding corruption: Theory and evidence from the audits of local. United States*: UCLA and UC-Berkeley.

Bolton, G. & A. Ockenfels (2000): ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90(1):166-193.

Braendle, T. (2015): Does remuneration affect the discipline and the selection of politicians? Evidence from pay harmonization in the European Parliament. *Public Choice* 162(1-2):1-24.

Cavalcanti, F., G. Daniele & S. Galetta (2018): Popularity shocks and political selection. *Journal of Public Economics* 165:201-216.

Chong, A., A. De La O, D. Karlan & L. Wantchekon (2014): Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice and party identification. *The Journal of Politics* 77(1):55-71.

Coates, D. (1999): *Efficiency wages for politicians: Do better paid elected officials produce better outcomes?* Mimeo: Department of Economics, University of Maryland.

De Vries, C. & H. Solarz (2017): The Electoral Consequences of Corruption. *Annual Review of Political Science* 20:391-408.

A. Dixit & J. Londregan (1996): The determinants of success of special interests in redistributive politics. *The Journal of Politics* 58(4):1132-1155.

Drouvelis, M., A. Saporiti, N. Vriend (2014): Political Motivations and Electoral Competition:Equilibrium Analysis and Experimental Evidence. *Games and Economic Behavior* 83:86-115.

Dufwenberg, M., U. Gneezy, J. Goeree & R. Nagel (2007): Price Floors and Competition. *Economic Theory* 33(1):211-24.

Engelmann, D. & M. Strobel (2004): Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review* 94(4):857-869.

Fehr, E., G. Kirchsteiger & A. Riedl (1993): Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108(2):437-459.

Fehr, E. & K. Schmidt (1999): A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114(3):817-868.

Ferraz, C., & F. Finan (2009): *Motivating Politicians: the Impacts of Monetary Incentives on Quality and Performance.* NBER working paper series, 14906

Ferraz, C., & F. Finan (2008): Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics* 123(2), 703-745.

Fischbacher, U. (2007): z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2):171-78.

Fisman, R., N. Harmon, E. Kamenica & I. Munk (2015) Labor Supply of Politicians. *Journal of the European Economic Association* 13(5):871–905.

Gagliarducci, S. & T. Nannicini (2013): Do better paid politicians perform better? Disentangling incentives from selection. *Journal of the European Economic Association* 11(2):369-398.

Goeree, J., C. Holt & T. Palfrey (2016): *Quantal Response Equilibrium. A Stochastic Theory of Games.* Princeton: Princeton University Press.

Goeree, J., C. Holt & T. Palfrey (2005): Regular Quantal Response Equilibrium. *Experimental Economics* 8(4):347-367.

Greiner, B. (2015): Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association* 1(1):114-125.

Grossman, G. & K. Michelitch (2018): Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda. *American Political Science Review* 112(2): 280–301.

Hafner-Burton, E., B. LeVek, D. Victor & J. Fowler (2014): Decision Maker Preferences for International Legal Cooperation. *International Organization* 68(4):845-876.

Haile, P., A. Hortacsu & G. Kosenok (2008): On the Empirical Content of Quantal Response Equilibrium. *American Economic Review* 98(1):180-200.

Harrison, G. & E. Rutström (2008): Risk aversion in the laboratory. In J. Cox & G. Harrison (eds.) *Risk aversion in experiments*. Bingley UK: Emerald Group.

Helland, L. & L. Monkerud (2013): Electoral agency in the lab: Learning to throw out the rascals. *Journal of Theoretical Politics* 25(2):214-33.

Helland, L., E. Moen & E. Preugschat (2017): Information and coordination frictions in experimental posted offer markets. *Journal of Economic Theory* 167:53-74.

Hey, J. & C. Orme (1994): Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62(6):1291-1326.

LeVeck, B., D. Hughes, J. Fowler, E. Hafner-Burton & D. Victor (2014): The role of self-interest in elite bargaining. *Proceedings of the National Academy of Sciences* 111(52):18536-18541.

Lind, J. T. & Rohner, D. (2017): Knowledge is Power: A Theory of Information, Income and Welfare Spending. *Economica* 84(336):611-646.

Lindbeck, A. & J. Weibull Lindbeck (1987): Balanced-budget redistribution as the outcome of political competition. *Public choice* 52(3):273-297.

McKelvey, R. & T. Palfrey (1995): Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10(1):6-38.

Meirowitz, A. (2008): Electoral contests, incumbency advantages, and campaign finance. *Journal of Politics* 27(3): 681–699

Mocan, N. & D. Altindag (2013). Salaries and work effort: An analysis of the European Union parliamentarians. *Economic Journal* 123(573):1130-1167.

Morton, R. & R. Myerson (2012): Decisiveness of contributors' perceptions in elections. *Economic Theory* 49(3):571-590.

Morton, R. (1993): Incomplete Information and Ideological Explanations of Platform Divergence. *American Political Science Review* 87(2):382-9.2

Persson, T. & G. Tabellini (2000): *Political Economics.* Cambridge Mass.: The MIT-Press.

Persson, T. & G. Tabellini (1999): The size and scope of government: Comparative politics with rational politicians. *European Economic Review* 43(4-6): 699-735.

Persson, T., G. Tabellini & F. Trebbi (2003): Electoral rules and corruption. *Journal of the European Economic Association* 1(4):958-989.

Polo, M. (1998): *Electoral competition and political rents.* Innocenzo Gasparini Institute for Economic Research. Milan.

Plott, C. (1989): An Updated Review of Industrial Organization: Applications of Experimental Methods. In R. Schmalensee & R. Willig (eds): *Handbook of Industrial Organization*, Vol. 2. North Holland:1109-1176.

Roth, A. & M. Malouf (1979): Game theoretic models and the role of bargaining. *Psychological Review* 86:574-594.

Rudolph, L. & T. Däubler (2016): Holding individual representatives accountable: The role of electoral systems. *Journal of politics* 78(3):746-762.

Strømberg, D. (2004): Mass Media Competition, Political Competition, and Public Policy. *Review of Economic Studies* 71(1):265-284.

Smith, V. (1982): Microeconomic Systems as an Experimental Science. *American Economic Review* 72(5):923-955.

Svaleryd, H. & J. Vlachos (2009): Political rents in a non-corrupt democracy. *Journal of Public Economics* 93(3-4):355-372.

Svensson, J. (1998), "Controlling Spending: Electoral Competition, Polarization, and Primary Elections", Mimeo, The World Bank.

Treisman, D. (2000): The causes of corruption: a cross-national study. *Journal of Public Economics* 76(3), 399-457.

van der Linde, D., S. Falcke, I. Koetsier & B. Unger (2014): *Do Wages Affect Politicians' Performance? A regression discontinuity approach for Dutch municipalities.* Discussion Paper Series/Tjalling C. Koopmans Research Institute, 14(15).

Woon, J. (2014): An Experimental Study of Electoral Incentives and Institutional Choice. *Journal of Experimental Political Science* 1(2):81-200.

Supplementary online appendix for "Can Paying Politicians Well Reduce Corruption? The Effects of Wages and Uncertainty on Electoral Competition"

A) Test for treatment differences: all rent promises and all periods

	Treatment				
Block	T1	T2	T3	T4	T5
B1	24.51	7.87	15.79	43.79	16.53
	(1.16)	(0.75)	(0.58)	(0.59)	(0.83)
B2	31.36	31.44	18.65	43.51	8.78
	(0.89)	(1.67)	(0.61)	(0.85)	(0.72)
B3	21.31	15.38	12.91	38.60	14.63
	(0.81)	(1.16)	(0.42)	(0.88)	(0.96)
B4	21.82	23.96	19.83	37.61	17.41
	(0.84)	1.65)	(0.99)	(0.75)	(0.82)
B5	48.30	6.41	19.83	30.16	13.22
	(1.36)	(0.83)	(1.06)	(0.71)	(0.73)
B6	21.00	3.68	27.10	26.66	9.14
	(0.98)	(0.63)	(1.48)	(0.98)	(0.71)
All	28.05	14.79	19.02	36.72	13.29
	(0.46)	(0.53)	(0.39)	(0.37)	(0.28)

Table S1: *Mean of all rent promises by block (standard error of the mean)*

All periods.

	T2	T3	T4	T5
T1	0.110‡	0.013†	0.055†	0.002†
	(1.601)	(2.242)	(-1.601)	(2.882)
T2		0.169†	0.005†	0.374†
		(-0.961)	(-2.562)	(-0.320)
T3			0.003†	0.027†
			(-2.722)	(1.922)
T4				0.002†
				(2.882)

Table S2: *WRS tests of treatment differences: all rent promises, all periods.*

p-values (test-statistics). †*One sided test;* ‡ *Two sided test*

Variable	Coefficient (Standard error)	
$T1$	28.05	***
	(4.04)	
$T2$	-13.26	**
	(5.80)	
$T3$	-9.03	**
	(4.43)	
$T4$	8.68	*
	(4.83)	
$T5$	-14.76	***
	(4.27)	
χ^2-model	67.67	***
R^2	0.15	
N	12.000	

Table S3: *Treatment regression: all rent promises, all periods. Reference category T1.*

Random effects for subjects, robust standard errors clustered on unique blocks.

****1%; **5%; *10%*

	T2	T3	T4	T5
T1	$0.000^{\ddagger}$	$0.000^{\dagger}$	$0.011^{\dagger}$	$0.000^{\dagger}$
	(20.64)	(20.05)	(5.19)	(27.26)
T2		$0.176^{\dagger}$	$0.000^{\dagger}$	$0.366^{\dagger}$
		(0.87)	(19.73)	(0.12)
T3			$0.000^{\dagger}$	$0.006^{\dagger}$
			(30.41)	(6.32)
T4				$0.000^{\dagger}$
				(61.34)

Table S4: χ^2 *tests of treatment differences based on Table S3: all rent promises, all periods.*

p-values (test-statistics)

B) Test for treatment differences: winning rent promises all periods

	Treatment				
Block	T1	T2	T3	T4	T5
B1	16.58	2.28	13.84	38.00	12.28
	(0.84)	(0.33)	(0.61)	(1.05)	(0.84)
B2	23.84	12.39	14.72	38.46	5.25
	(0.85)	(1.10)	(0.70)	(1.06)	(0.61)
B3	14.45	6.80	11.22	32.87	9.72
	(0.67)	(0.83)	(0.44)	(1.11)	(0.94)
B4	15.01	7.83	11.50	32.23	12.42
	(0.56)	(1.31)	(0.75)	(1.02)	(0.95)
B5	33.33	1.34	12.44	26.07	9.49
	(1.32)	(0.23)	(0.64)	(0.90)	(0.75)
B6	14.26	0.56	14.53	19.96	6.22
	(0.72)	(0.18)	(1.30)	(0.83)	(0.62)
All	19.60	5.20	13.04	31.26	9.23
	(0.40)	(0.34)	(0.32)	(0.45)	(0.33)

Table S5: *Mean of winning promises by block (standard error of the mean)*

All periods

	T2	T3	T4	T5
T1	0.004‡	0.013†	0.019†	0.002†
	(2.882)	(2.242)	(-2.082)	(2.882)
T2		0.005†	0.002†	0.075†
		(-2.562)	(-2.882)	(-1.441)
T3			0.002†	0.013†
			(-2.882)	(2.242)
T4				0.002†
				(2.882)

Table S6: *WRS tests of treatment differences: winning rent promises, all periods.*

p-values (test-statistics). †*One sided test;* ‡ *Two sided test*

Variable	Coefficient (Standard error)	
$T1$	20.60	***
	(3.26)	
$T2$	-14.56	***
	(3.89)	
$T3$	-7.05	**
	(3.33)	
$T4$	11.41	***
	(4.31)	
$T5$	-10.92	***
	(3.49)	
χ^2-model	68.89	***
R^2	0.33	
N	6.000	

Table S7: *Treatment regression: winning rent promises, all periods. Reference category T1. Random effects for subjects, robust standard errors clustered on unique blocks. ***1%; **5%; *10%*

	T2	T3	T4	T5
T1	$0.000^{\ddagger}$	$0.000^{\dagger}$	$0.098^{\dagger}$	$0.000^{\dagger}$
	(26.35)	(17.83)	(1.68)	(22.61)
T2		$0.000^{\dagger}$	$0.000^{\dagger}$	$0.070^{\dagger}$
		(11.18)	(53.85)	(2.18)
T3			$0.000^{\dagger}$	$0.004^{\dagger}$
			(40.05)	(7.20)
T4				$0.000^{\dagger}$
				(52.15)

Table S8: χ^2 *tests of treatment differences based on Table S7: winning rent promises, all periods. p-values (test-statistics).* †*One sided test;* ‡ *Two sided test*

C) Test for treatment differences all rent promises last 25 periods

	Treatment				
Block	T1	T2	T3	T4	T5
B1	20.55	11.20	22.82	42.87	17.70
	(0.69)	(1.11)	(0.90)	(1.34)	(1.22)
B2	36.58	26.52	24.63	42.66	9.19
	(1.05)	(2.29)	(0.93)	(1.21)	(1.05)
B3	26.85	16.39	17.22	42.27	19.66
	(1.06)	(1.10)	(0.69)	(1.39)	(1.46)
B4	19.61	19.96	24.09	36.53	21.33
	(1.23)	(2.38)	(1.47)	(1.06)	(1.24)
B5	46.02	8.32	24.19	29.23	16.65
	(2.04)	(0.99)	(1.45)	(0.99)	(1.20)
B6	24.42	4.86	18.50	25.69	12.38
	(0.92)	(0.74)	(1.72)	(1.44)	(1.17)
All	29.00	14.54	21.91	36.54	16.15
	(0.57)	(0.68)	(0.52)	(0.55)	(0.52)

Table S9: *Mean of all rent promises by block (standard error of the mean). Last 25 periods.*

	T2	T3	T4	T5
T1	0.025‡	0.075†	0.100†	0.008†
	(2.242)	(1.441)	(-1.281)	(2.402)
T2		0.055†	0.003†	0.261†
		(-1.601)	(-2.722)	(-0.641)
T3			0.002†	0.019†
			(-2.882)	(2.082)
T4				0.002†
				(2.882)

Table S10: *WRS tests of treatment differences: all rent promises, last 25 periods. p-values (test-statistics).* †*One sided test;* ‡ *Two sided test.*

Variable	Coefficient (Standard error)	
$T1$	29.00	***
	(3.91)	
$T2$	-14.46	***
	(4.95)	
$T3$	-7.10	**
	(4.10)	
$T4$	7.54	*
	(4.84)	
$T5$	-12.86	***
	(4.28)	
χ^2-model	46.50	***
R^2	0.15	
N	6.000	

Table S11: *Treatment regression: all rent promises, last 25 periods. Reference category T1.*

Random effects for subjects, robust standard errors clustered on unique blocks.

****1%; **5%; *10%*

	T2	T3	T4	T5
T1	0.000‡	0.000†	0.005†	0.000†
	(26.83)	(20.78)	(6.65)	(27.29)
T2		0.012†	0.000†	0.323†
		(5.07)	(27.97)	(0.21)
T3			0.000†	0.003†
			(22.32)	(7.38)
T4				0.000†
				(37.44)

Table S12: χ^2 *tests of treatment differences based on Table S11: all rent promises, last 25 periods*

p-values (test-statistics)

D) Test for treatment differences winning rent promises last 25 periods

	Treatment				
Block	T1	T2	T3	T4	T5
B1	17.94	4.05	19.66	36.60	13.72
	(0.70)	(0.53)	(0.87)	(1.65)	(1.23)
B2	30.73	8.26	19.63	38.49	5.30
	(1.07)	(1.25)	(1.06)	(1.58)	(0.88)
B3	20.01	9.90	14.66	35.24	14.34
	(0.88)	(0.95)	(0.66)	(1.76)	(1.60)
B4	13.64	6.40	15.52	31.79	16.82
	(0.89)	(1.73)	(1.25)	(1.43)	(1.52)
B5	30.18	2.68	15.66	25.39	11.52
	(1.77)	(0.41)	(1.08)	(1.23)	(1.23)
B6	19.44	1.12	9.41	18.76	8.76
	(0.89)	(0.36)	(1.44)	(1.18)	(1.05)
All	21.99	5.40	15.76	31.04	11.75
	(0.51)	(0.43)	(0.46)	(0.67)	(0.54)

Table S13: *Mean of winning promises by block (standard error of the mean)*

Last 25 periods.

	T2	T3	T4	T5
T1	0.004‡	0.075†	0.027†	0.008†
	(2.882)	(1.441)	(-1.922)	(2.402)
T2		0.003†	0.002†	0.013†
		(-2.722)	(-2.882)	(-2.242)
T3			0.005†	0.039†
			(-2.562)	(1.761)
T4				0.002†
				(2.882)

Table S13: *WRS tests of treatment differences: winning rent promises, last 25 periods.*

p-values (test-statistics). †*One sided test;* ‡ *Two sided test.*

Variable	Coefficient (Standard error)	
$T1$	22.72	***
	(2.82)	
$T2$	-16.93	***
	(3.13)	
$T3$	-6.14	*
	(3.23)	
$T4$	8.59	**
	(4.03)	
$T5$	-10.80	***
	(3.23)	
χ^2-model	85.82	***
R^2	0.31	
N	3.000	

Table S14: *Treatment regression: winning rent promises, last 25 periods. Reference category T1. Random effects for subjects, robust standard errors clustered on unique blocks.*

****1%; **5%; *10%*

	T2	T3	T4	T5
T1	0.000‡	0.000†	0.013†	0.000†
	(46.59)	(24.25)	(4.96)	(32.70)
T2		0.000†	0.000†	0.002†
		(27.10)	(64.14)	(8.64)
T3			0.000†	0.018†
			(20.19)	(4.42)
T4				0.000†
				(34.88)

Table S15: χ^2 *tests of treatment differences based on Table S14: winning rent promises, last 25 periods. p-values (test-statistics).* †*One sided test;* ‡ *Two sided test.*

E) Dynamic regressions

We formally address the question of convergence running dynamic regressions treatment by treatment (Noussair et al 1995, Noussair et al 1997, Cason & Noussair 2007). The specification employed is:

$$y_{it} = \sum_{i=1}^{6} \beta_{1i} D_i(1/t) + \sum_{i=1}^{6} \beta_{2i} D_i((t-1)/t) + \mu_{it},$$

were y_{it} is all rent taking promises, i indicates block and $t \in [1, T]$ indicates period. The $((t-1)/t)$ terms take the value 0 in period 1, thus β_{1i} provides an estimate of the value of y_{i1} for block i. As t grows the $((t-1)/t)$ terms approach 1 and the $1/t$ terms approach 0, thus β_{2i} is an estimate of the asymptote of y_{iT}. The criteria for convergence are as follows. The process is said to exhibit strong convergence if $H_0 : \beta_{21} = \beta_{22} = \beta_{23} = \beta_{24} = \beta_{25} = \beta_{26}$ cannot be rejected. The process is said to exhibit weak convergence if β_{2i} is closer to the Nash equilibrium value of the treatment than is β_{1i}. Table S16 provides the results. The regressions are estimated with random intercepts for unique subjects, and corrected standard errors for correlation over panels (Prais-Winsten regression).

	T1	T2	T3	T4	T5
β_{11}	37.0	24.8	49.2	45.6	22.6
	(7.0)	(5.7)	(3.6)	(4.5)	(6.3)
β_{12}	35.0	37.6	38.3	26.8	40.1
	(6.4)	(8.2)	(5.2)	(4.4)	(6.3)
β_{13}	54.1	23.7	48.4	33.6	37.8
	(5.6)	(4.8)	(4.1)	(5.2)	(4.2)
β_{14}	49.1	20.0	49.3	18.0	25.7
	(8.7)	(9.4)	(5.4)	(8.1)	(4.6)
β_{15}	32.4	29.3	41.6	54.3	39.4
	(6.0)	(9.2)	(1.8)	(5.7)	(4.7)
β_{16}	50.5	29.9	20.9	38.5	33.3
	(5.3)	(5.0)	(10.0)	(4.0)	(4.8)
β_{21}	23.7	6.1	12.3	43.6	16.2
	(2.1)	(1.6)	(1.1)	(1.1)	(1.5)
β_{22}	20.5	12.8	18.0	30.4	12.1
	(1.9)	(2.3)	(1.5)	(1.1)	(1.5)
β_{23}	29.2	1.6	15.5	44.7	5.9
	(1.7)	(1.3)	(1.2)	(1.3)	(1.0)
β_{24}	48.0	32.6	16.7	27.6	16.5
	(2.6)	(2.6)	(1.6)	(2.1)	(1.1)
β_{25}	20.2	23.6	10.0	36.9	10.7
	(1.8)	(2.6)	(0.5)	(1.4)	(1.2)
β_{26}	18.7	4.3	28.4	37.7	6.7
	(1.6)	(1.4)	(3.0)	(1.0)	(1.2)
H_0 (p-value)	0.000	0.000	0.000	0.000	0.000
$E(r^*)$	0	0	9	49	10

Table S16: *Dependent: all rent promises. Prais-Winsten regressions treatment by treatment, with subject random effects. Coefficients (standard errors).*

We appreciate that none of the treatments converge in a strong sense (the H_0 can not be rejected for any of the treatments). Except for T4, however, the overall picture is clearly one of convergence in the weak sense. This is illustrated in Table S17. In the table a positive sign indicates that β_{2i} is closer to the Nash equilibrium than β_{1i}, the asterisks indicate significance levels of the observed differences (at the ***1%, **5% or *10% levels respectively, using a χ^2 test of differences in coefficients). In T4 equally many blocks move towards and away from Nash, and only two differences are significant (one positive and one negative). In the other treatments five or six blocks move closer to the Nash equilibrium and four of five of these are significant at conventional levels. Negative signs (movements away from Nash) are never significant at conventional levels in these treatments.

Treatment Block	T1		T2		T3		T4		T5	
B1	+	*	+	***	+	***	÷		+	***
B2	+	**	+	***	+	***	+	**	+	***
B3	+	***	+	***	+	***	+		+	***
B4	+		÷		+	***	+		+	*
B5	+	*	+		+	***	÷	***	+	***
B6	+	***	+	***	÷		÷		+	***

Table S17: *Summary of convergence pattern.*

F) QRE robustness results

Figure S1 displays the CDFs of the estimated QRE, the data and the Nash prediction when $\alpha = 0.75$ (exogenously set) while the decision error (λ) is simultaneously estimated and forced to be identical over all treatments. Comparing with Figure 5 in the main text, simultaneously estimating λ mainly leads to a less impressive fit in T4. Overall, the consequences of this estimation choice, nevertheless, does not overturn the statement that the QRE captures the patterns of our data well when risk-attitudes are taken account of.

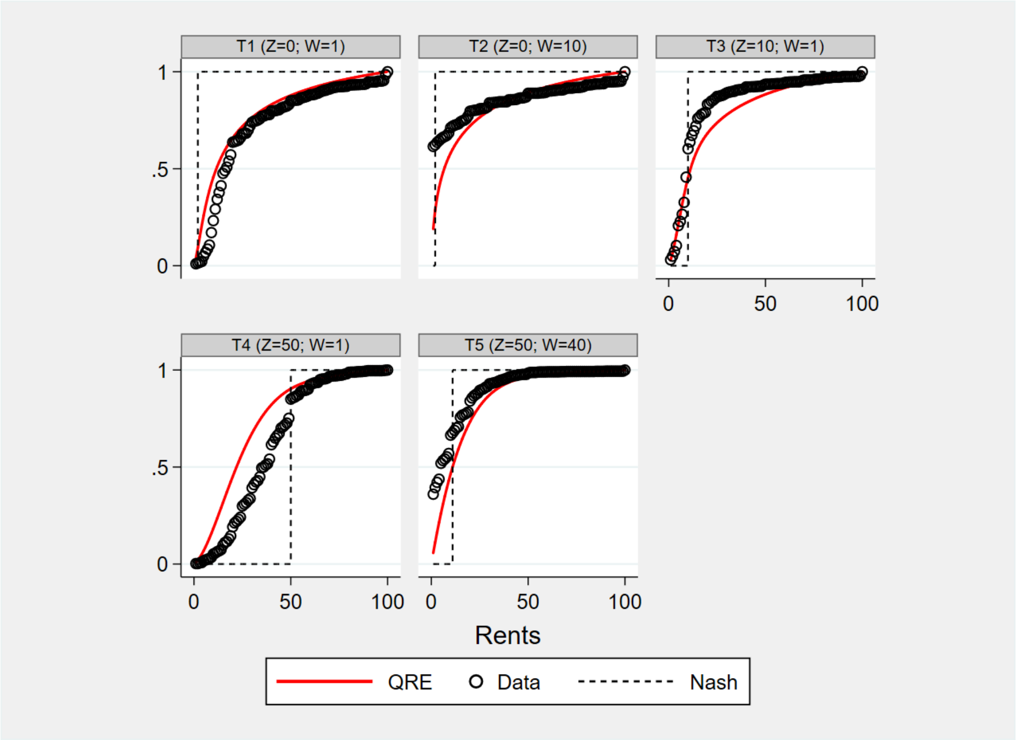


Figure S1: *Simultaneous estimation of decision errors (λ) for all treatments with risk attitudes (α) set to 0.75.*

Figure S2 displays the CDFs of the estimated QRE, the data and the Nash prediction when both risk-attitudes and the identical decision errors are estimated simultaneously. Comparing with Figure 5 in the text there is little difference in results. The estimated risk-aversion parameter is somewhat lower ($\hat{\alpha} = 0.35$) than the exogenous risk parameter used in the estimations underlying Figure 5 in the main text ($\alpha = 0.75$).

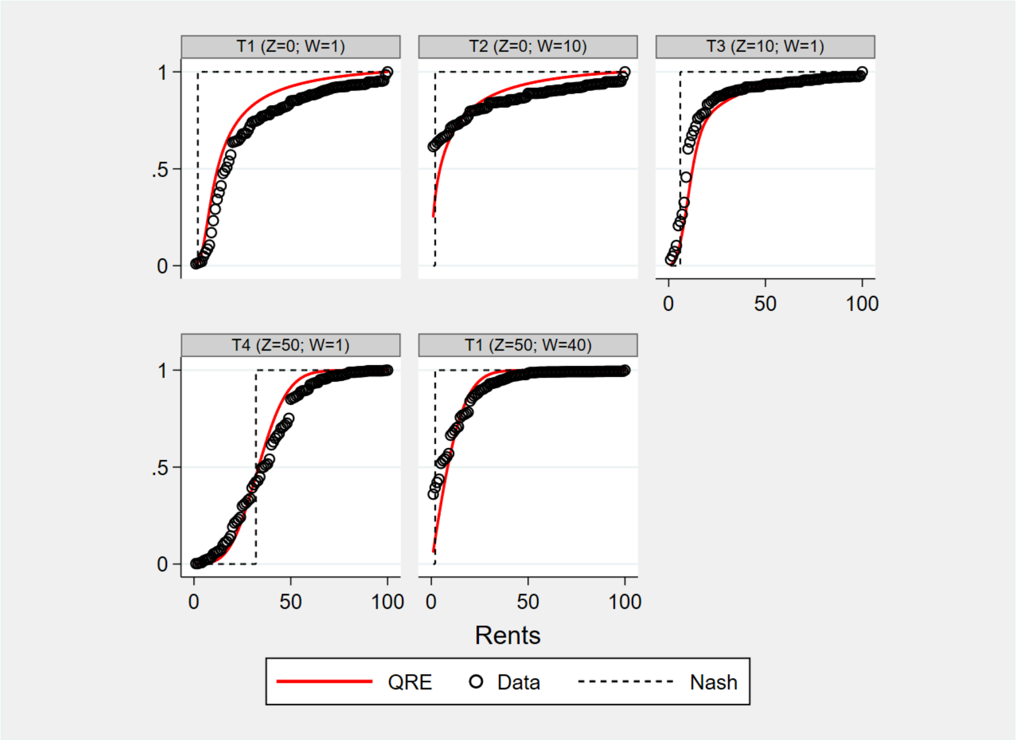


Figure S2: *Joint estimation of risk aversion* (α) *and decision error* (λ) *simultaneously for all treatments.*

G) Data on collusion

We agree with Kruse et al. (1994) and Potters and Seutens (2009) that feasible schemes of tacit collusion should not go beyond simple patterns such as constant rent-levels or simple rotation schemes. In line with this we focus on two statistics. First, we look for stable rent-levels above equilibrium as a sign of collusive outcomes. Second, following the literature on cooperation in the absence of common knowledge rationality (e.g. Kreps et al. 1982), we conjecture that collusive arrangements—if established—will unravel in the final periods of the experiment. Thus, we check for significant price decreases in the final periods of the experiment.

Figure S3 follows Friedman et al. (2015), it displays the stability of rents by plotting the probability of a rent proposal being changed in the current period as a function of the rent proposed in the previous period. Each point in the figures represents the average probability of a rent change in the current period given a proposed rent in the previous period from bin $[0-10)$, $[10-20)$, ..., $[90-100]$. The radius of the circles are proportional to the number of sellers having proposed a rent in the relevant bin in the previous period. The figure aggregates over the last 25 games of the experiment (when subjects presumably have had ample time to learn how to collude).

The figure makes clear that in none of the treatments do rents stabilize above equilibrium levels (remember that risk neutral Nash rent-taking is 0 in T1 and T2, 9 in T3, 49 in T4 and 10 in T5). The mass of rent-promises are found in the vincinity of the Nash equilibrium, and generally rent-levels in the vincinity of the Nash equilibrium are more stable than other rent-levels.

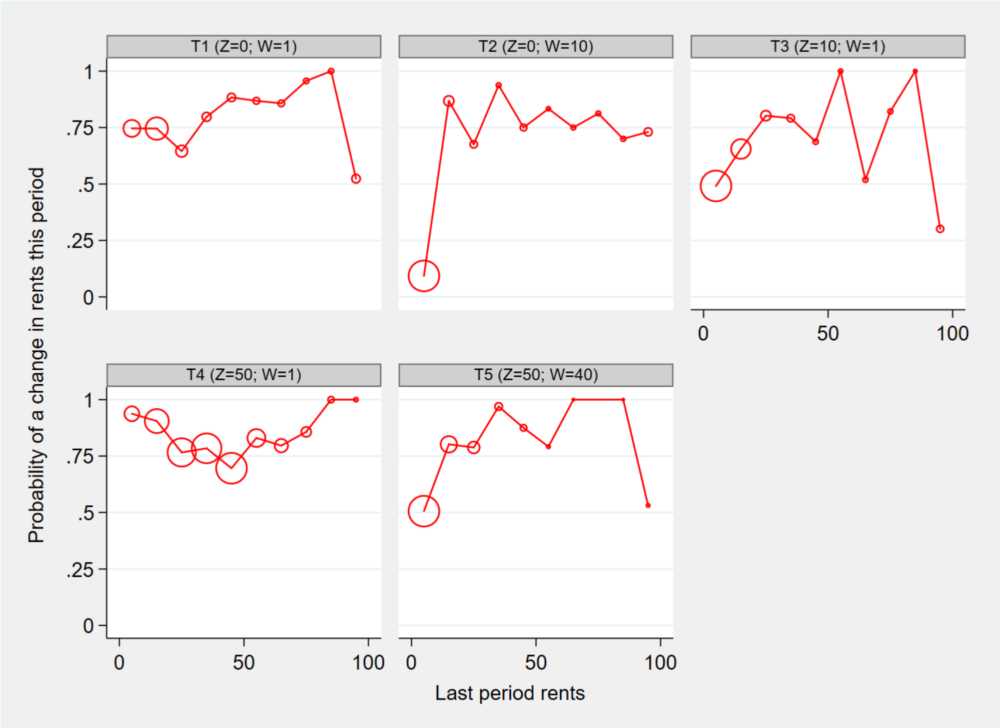


Figure S3: *Stability of rents treatment by treatment, averaged over the last 25 games.*

Table S18 show random effects regressions with standard errors clustered at unique blocks, treatment-by-treatment. The regressions include all 50 games, and have dummies for structural breaks. These dummies take the value one if in game 50-k and zero otherwise, with k={0}, {0,1}, ..., {0, 1, 2, 3, 4}. The dependent variable in these regressions is the first difference in rents. We would expect end-game effects (unravelling of collusive arrangements) to show up in negatively significant coefficients on the dummies for structural breaks.

As is clear from the table all significant endgame effect move in the opposite direction of what should be expected if subjects had managed to establish collusive agreements. Negative coefficients on the structural break dummies appear only at the very last games of treatments T4 and T5, are generally modest in size and are not significantly different from zero at conventional levels.

	T1		T2		T3		T4	T5	
D_50	0.66		1.21		3.44		-1.08	-0.31	
	(1.45)		(2.19)		(2.93)		(2.31)	(2.29)	
D_49-50	1.08	**	0.66	*	1.17	***	-0.15	0.82	***
	(0.47)		(0.40)		(0.21)		(0.15)	(0.31)	
D_48-50	1.01	**	0.49		1.27	***	-0.07	0.91	***
	(0.51)		(0.45)		(0.07)		(0.18)	(0.29)	
D_47-50	0.85		0.64		1.10	***	0.02	1.02	***
	(0.57)		(0.40)		(0.17)		(0.34)	(0.23)	
D_46-50	1.37		0.71	**	1.15	***	0.23	0.96	***
	(0.30)		(0.36)		(0.27)		(0.30)	(0.30)	

Table S18: *Dependent: First difference in rents*

Random effects for subjects, (robust standard errors clustered on unique blocks)

****1%; **5%; *10%.*

H) Sample instructions T1 (Z=0; W=1)

Welcome! You are participating in an experiment financed by xxx.

You will earn money in the experiment. How much you earn depends on the decisions you make, as well as on the decisions made by other subjects.

All interactions are anonymous and are performed through a network of computers. The administrators of the experiment will not be able to observe your decisions during the experiment.

All participants in the experiment are present in this room. They have all been recruited in the same way as you and are reading the same instructions as you are for the first time.

It is important that you do not talk to any of the other participants in the room until the experiment is over.

In the experiment your payoffs are denominated in experimental currency units (ECUs). At the end of the experiment, you will be paid in [local currency] based on your total earnings in ECUs from all the periods of the experiment. The exchange rate from ECU to [the local currency] is:

1 ECU= 0.33 [Local Currency]

The more ECUs you earn, the more cash you will receive.

The experiment

The experiment consists of 50 periods.

Prior to the first period each participant is assigned a label. This label is either "A" or "B".

Labels are assigned randomly. You keep your assigned label through the entire experiment.

In each period, participants are randomly matched into pairs consisting of one A and one B.

In each period, you choose a number on the interval from 0 to 100. You can choose your number with up to three decimals.

You choose your number without knowing the chosen number of your match. Likewise, your match chooses his or her number without knowing your chosen number.

The payoffs in each period is explained in the next section.

Payoffs

In each period, one participant in the match gets a positive payoff equal to her/his chosen number + 1, while the other participant in the match gets a payoff of zero.

A gets the positive payoff if her/his chosen number is **smaller** than B's chosen number.

B gets the positive payoff if her/his chosen number is **smaller** than A's chosen number.

If A's chosen number **equals** B's chosen number, then the outcome is determined by a fair lottery implemented by the computer. In this lottery the payoff of the winner is her/his chosen number +1, while the payoff of the loser is 0.

Examples

1. A's chosen number is 36 and B's chosen number is 52. Thus A's payoff is 37 while B's payoff is 0.

2. A's chosen number is 2 and B's chosen number is 1. Thus A's payoff is 0 while B's payoff is 2.

3. A's chosen number is 12 and B's chosen number is 12. Thus a winner is determined by a fair lottery. The payoff of the winner is 13, while the payoff of the loser is 0.

Feedback

After each period there is a feedback screen. This screen provides information about your current payoff, your accumulated payoff, and the choices of you and your match. You will also see the same information for all previous periods on the feedback screen.

Earnings

After the last period ends, your payoffs in ECU are converted to [the local currency] at the stated exchange rate. Your earnings in [the local currency] will be paid in cash as you exit the lab.

Are there any questions?

I) Sample instructions T3 (Z=10, W=1)

Welcome! You are participating in an experiment financed by xxxx.

You will earn money in the experiment. How much you earn depends on the decisions you make, as well as on the decisions made by other subjects.

All interactions are anonymous and are performed through a network of computers. The administrators of the experiment will not be able to observe your decisions during the experiment.

All participants in the experiment are present in this room. They have all been recruited in the same way as you and are reading the same instructions as you are for the first time.

It is important that you do not talk to any of the other participants in the room until the experiment is over.

In the experiment your payoffs are denominated in experimental currency units (ECUs). At the end of the experiment, you will be paid in [the local currency] based on your total earnings in ECUs from all the periods of the experiment. The exchange rate from ECU to [the local currency] is:

1 ECU = 0.8 [the local currency]

The more ECUs you earn, the more cash you will receive.

The experiment

The experiment consists of 50 periods.

Prior to the first period each participant is assigned a label. This label is either "A" or "B".

Labels are assigned randomly. You keep your assigned label through the entire experiment.

In each period, participants are randomly matched into pairs consisting of one A and one B.

In each period, you choose a number on the interval from 0 to 100. You can choose your number with up to three decimals.

You choose your number without knowing the chosen number of your match. Likewise, your match chooses his or her number without knowing your chosen number.

In each period, after you and your match have chosen your numbers, the computer program draws a random number with three decimals.

The random number is drawn from the interval -10 to $+10$. Any random number in this interval has the **same probability** of being drawn.

The use of the random number and the participants' payoff in each period are explained in the next section.

Payoffs

In each period, one participant in the match gets a positive payoff equal to her/his chosen number + 1, while the other participant in the match gets zero payoff.

A gets the positive payoff if her/his chosen number is **smaller** than B's chosen number + the random number.

B gets the positive payoff if her/his chosen number + the random number is **smaller** than A's chosen number.

Thus, a positive random number is an advantage to A, while a negative random number is an advantage to B.

If A's chosen number **equals** B's chosen number + the random number, then the outcome is determined by a fair lottery implemented by the computer. In this lottery the payoff of the winner is her/his chosen number +1, while the payoff of the loser is 0.

Examples

1. A's chosen number is 60 and B's chosen number is 55 while the random number is + 6.223. Then A' s chosen number is 60 while B's chosen number + the random number is 61.223. Thus A's payoff is 61 while B's payoff is 0.

2. A's chosen number is 60 and B's chosen number is 55 while the random number is ÷ 6.303. Then A's chosen number is 60 while B's chosen number + the random number is 48.697. Thus A's payoff is 0 while B's payoff is 56.

3. A's chosen number is 2 and B's chosen number is 1 while the random number is ÷ 2.357. Then A's chosen number is 2 while B's chosen number + the random number is ÷1.357. Thus A's payoff is 0 while B's payoff is 2.

4. A's chosen number is 2 and B's chosen number is 1 while the random number is + 3.455. Then A's chosen number is 2 while B's chosen number + the random number is +4.455. Thus A's payoff is 3 while B's payoff is 0.

5. A's chosen number is 25 and B's chosen number is 20 while the random number is +5.000. Then A's chosen number is 25 and B's chosen number + the random number is 25.000. Thus a winner is determined by a fair lottery. The payoff to A if A wins is 26, while the payoff to B if B wins is 21.

Feedback

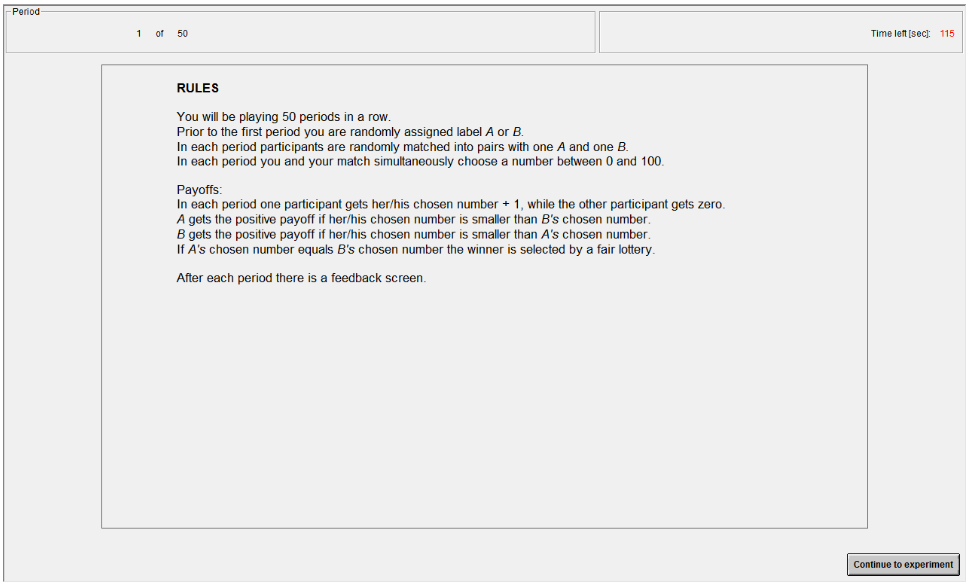
After each period there is a feedback screen. This screen provides information about your current payoff, your accumulated payoff, the choices of you and your match, and the random number drawn by the computer. You will also see the same information for all previous periods on the feedback screen.

Earnings

After the last period is completed, your payoffs in ECU are converted to [the local currency] at the stated exchange rate. Your earnings in [the local currency] will be paid in cash as you exit the lab.

Are there any questions?

J) Sample screens T1 (Z=0;W=1)



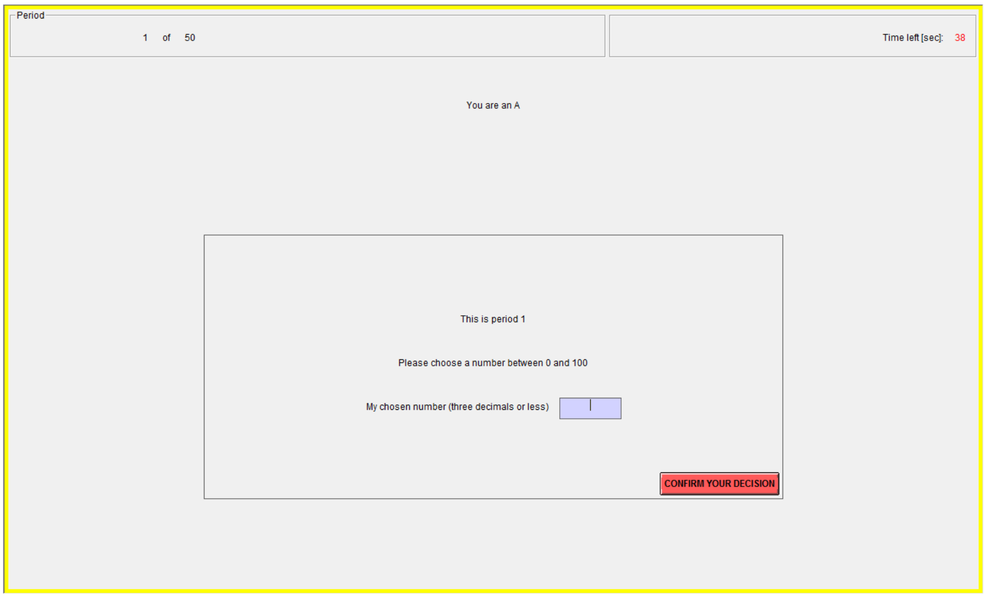
RULES

You will be playing 50 periods in a row.
Prior to the first period you are randomly assigned label *A* or *B*.
In each period participants are randomly matched into pairs with one *A* and one *B*.
In each period you and your match simultaneously choose a number between 0 and 100.

Payoffs:
In each period one participant gets her/his chosen number + 1, while the other participant gets zero.
A gets the positive payoff if her/his chosen number is smaller than *B's* chosen number.
B gets the positive payoff if her/his chosen number is smaller than *A's* chosen number.
If *A's* chosen number equals *B's* chosen number the winner is selected by a fair lottery.

After each period there is a feedback screen.

Figure S4: *Information screen T1 (only on in period 1).*



Period 1 of 50

Time left [sec]: 38

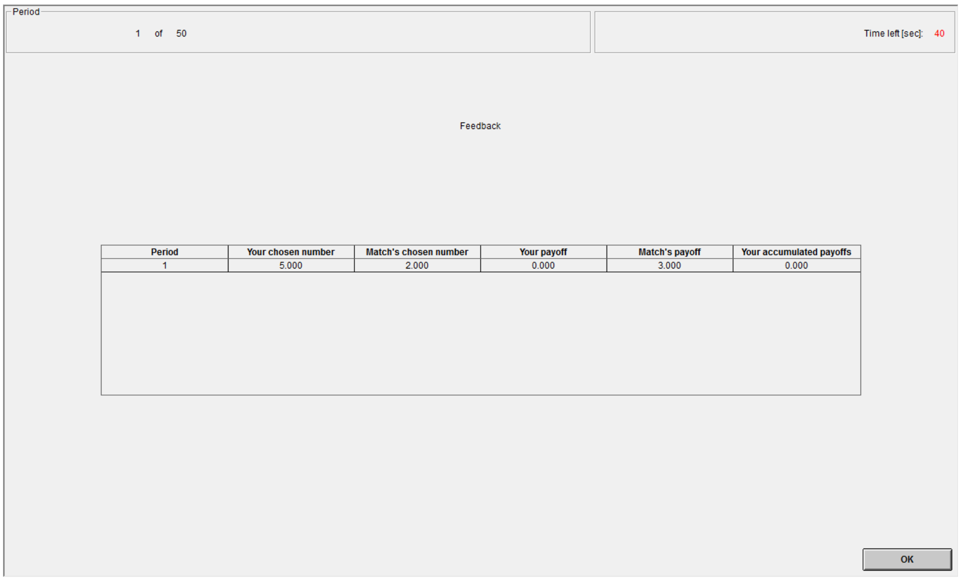
You are an A

This is period 1

Please choose a number between 0 and 100

My chosen number (three decimals or less)

Figure S5: *Decision screen T1.*



Period 1 of 50

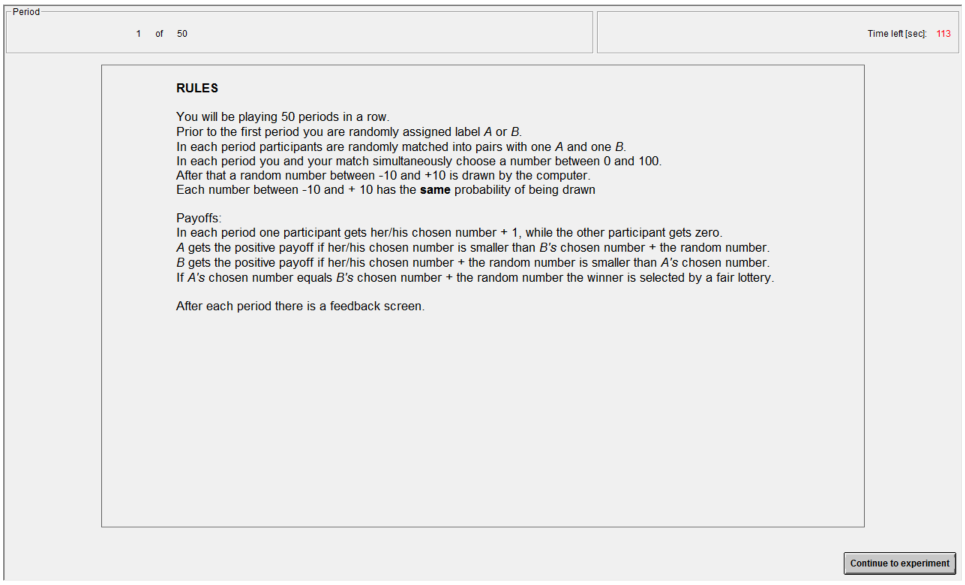
Time left [sec]: 40

Feedback

Period	Your chosen number	Match's chosen number	Your payoff	Match's payoff	Your accumulated payoffs
1	5.000	2.000	0.000	3.000	0.000

Figure S6: *Feedback screen T1.*

K) Sample screens T3 (Z=10;W=1)



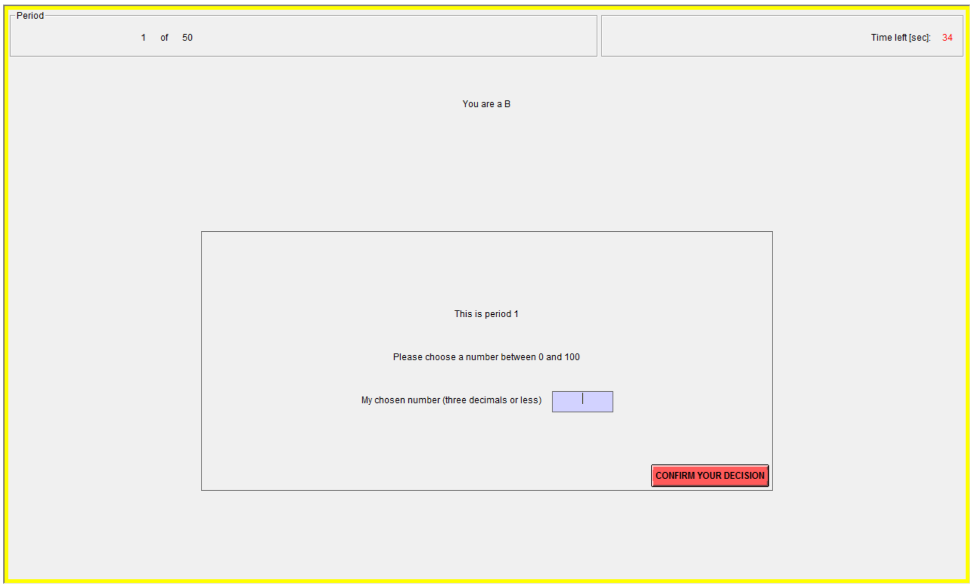
RULES

You will be playing 50 periods in a row.
Prior to the first period you are randomly assigned label *A* or *B*.
In each period participants are randomly matched into pairs with one *A* and one *B*.
In each period you and your match simultaneously choose a number between 0 and 100.
After that a random number between -10 and +10 is drawn by the computer.
Each number between -10 and + 10 has the **same** probability of being drawn

Payoffs:
In each period one participant gets her/his chosen number + 1, while the other participant gets zero.
A gets the positive payoff if her/his chosen number is smaller than *B's* chosen number + the random number.
B gets the positive payoff if her/his chosen number + the random number is smaller than *A's* chosen number.
If *A's* chosen number equals *B's* chosen number + the random number the winner is selected by a fair lottery.

After each period there is a feedback screen.

Figure S7: *Information screen T3 (only in period 1).*



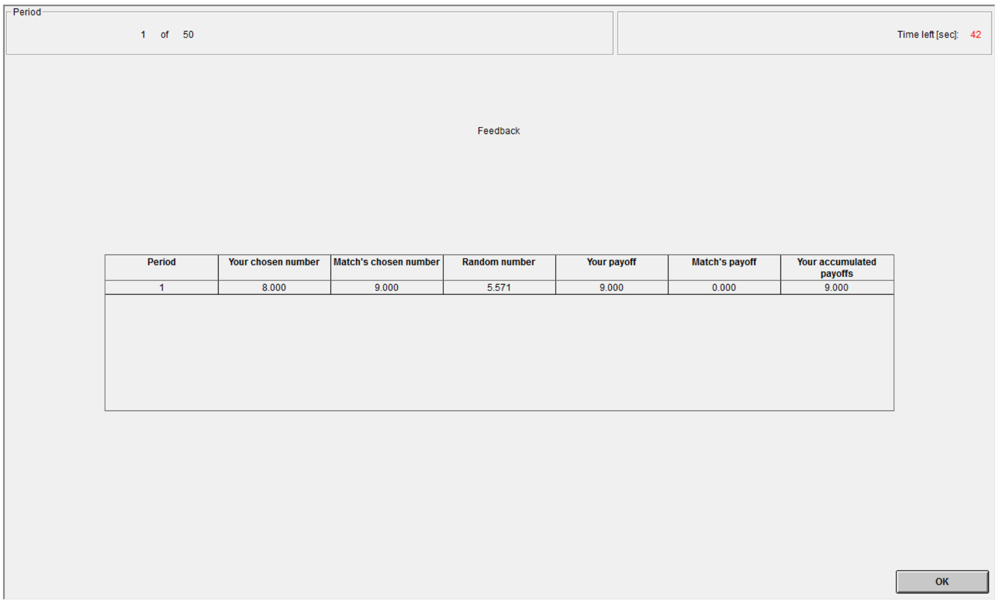
You are a B

This is period 1

Please choose a number between 0 and 100

My chosen number (three decimals or less)

Figure S8: *Decision screen T3.*



Period	Your chosen number	Match's chosen number	Random number	Your payoff	Match's payoff	Your accumulated payoffs
1	8.000	9.000	5.571	9.000	0.000	9.000

Figure S9: *Feedback screen T3.*

References

Cason, T., and C. Noussair (2007): A Market with Frictions in the Matching Process: An Experimental Study. *International Economic Review* 48(2): 665–691.

Davis, D., O. Korenok and R. Reilly (2010): Cooperation without coordination: signaling, types and tacit collusion in laboratory oligopolies. *Experimental Economics* 13(1): 45–65.

Friedman, D., S. Huck, R. Oprea and S. Weidenholzer (2015): From imitation to collusion: Long-run learning in a low-information environment. *Journal of Economic Theory* 155:185–205.

Kreps, D., P. Milgrom and D. R. Wilson (1982): Reputation and imperfect information. *Journal of Economic Theory* 27:253-279.

Kruse, J. B., S. Rassenti, S., S. Reynolds and V. L. Smith (1994): Bertrand-Edgeworth competition in experimental markets. *Econometrica* 62(2): 343–371.

Noussair, C. N., C. R. Plott, and R. G. Riezman (1997):The principles of exchange rate determination in an international finance experiment. *Journal of Political Economy* 105(4): 822–861.

Noussair, C. N., C. R. Plott, and R. G. Riezman (1995): An experimental investigation of the patterns of international trade. *American Economic Review* 85(3): 462–491.

Potters, J. and S. Suetens (2009): Cooperation in experimental games of strategic complements and substitutes. *Review of Economic Studies* 76(3):1125–1147.