Essays in Applied Microeconomics

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Andreas Grunewald
aus Bonn

Bonn 2013

Dekan:	Prof. Dr. Klaus Sandmann
Erstreferent:	Prof. Dr. Paul Heidhues
Zweitreferent:	Prof. Dr. Armin Falk

Tag der mündlichen Prüfung: 22.08.2013

Acknowledgements

First and foremost, I want to thank my supervisors Paul Heidhues and Armin Falk. Their guidance, comments and support have shaped and improved this dissertation in all aspects. Their lessons have helped me to develop an ideal of how research should be conducted and will impact all research that lies ahead of me. They have also deeply impressed me with their creativity motivation and friendliness.

To write this dissertation would have neither been possible nor would it have been fun if it had not been for my co-authors Steffen Altmann, Armin Falk, Emanuel Hansen, David Huffman and Gert Pönitzsch. I have learned a lot from them about all parts of economics in the always enlightening discussions. These discussion have been a major source of inspiration pleasure and motivation for conducting research.

I also want to thank all BGSE members and officials for their organization of the BGSE and the finical support they provided. They have made the research environment at the BGSE the best I can imagine.

I am particularly grateful to my co-students in the BGSE who have made me enjoy to come to the BGSE every day in all these years. Even more important they have also enriched my live massively during the ”non-research” times with their advice friendship and good laughs.

Finally, I want to thank my family for the never ending support and love they give to me every single day.

Contents

List of Figures

List of Tables

1. Introduction

During the last years and especially during the course of the financial crises, economics has been criticized for being agnostic about real world problems. There seems to be a consensus in media that the theoretical strand of economics as a whole lacks the connection to empirical applications and can thus not provide any policy recommendation. In particular, the models of the economic profession are criticized for relying on assumptions about human characteristics that do not replicate observed behavior. The derived implications may thus have no empirical value, or even worse lead to wrong insights. A closer connection of the theoretical solutions to economic problems and their empirical validation has two advantages. Firstly, it leads to a better comprehension of the underlying determinants of human behavior. There may either arise supporting or contradicting evidence with respect to the corresponding assumptions about preferences. This in turn leads to a sound understanding which characteristics play a crucial role in the analyzed situation. Secondly, it helps to filter those models that yield reliable predictions and can thus be used to derive policy advice.

This dissertation tackles three questions theoretically and empirically. First, why do non binding default options influence the behavior of consumers? Second, does contractual incompleteness cause labor market segmentation and unemployment? And third, how should democratic institutions be drafted to foster the quality of political leaders? This approach unifies both a clear cut and concise derivation of hypotheses and an immediate empirical validation of the results. Hence, it ensures an instant relation between the theoretical solution of the considered problems and real world data.

In the first chapter we study how non binding default options influence the behavior of consumers. In any decision environment a default option specifies which alternative is implemented if no active decision is taken by the agent. Although seemingly irrelevant for the behavior of rational consumers, choices are influenced by default options in a wide array of applications (Johnson and Goldstein 2003, Madrian and Shea 2001, Levav et al.

2010). We first provide a theoretical framework to analyze when decision makers should rationally follow default options instead of making active choices. We model decision-making in environments with default options as strategic interaction between a policy maker and decision makers. Our model focuses on two key dimensions that affect behavior of both parties in equilibrium. First, choices of policy makers and decision makers are influenced by the degree of alignment of their interests. Second, the strength of default effects depends on the parties' relative level of information about the decision environment and decision makers' optimal choices.

We test the key predictions of our model in a laboratory experiment exogenously varying these two key dimensions. We find supportive evidence that strategic incentives as well as the level of information affect behavior of policy makers and decision makers. Our findings also provide insights for the evaluation of "libertarian paternalistic" policy interventions as well as consumer protection regulation. Participants in our experiment seem to understand the arising conflict of interest and tailor their decisions accordingly, taking into account the alignment of interests as well as the relative level of information. Hence, defaults are not harmful to those that have a sound understanding which decision alternative is best for them, while improving the decision quality of less informed agents.

The second chapter provides evidence that two important features of labor markets—the existence of involuntary unemployment, and the segmentation of markets according to firms offering "good" and "bad" jobs—may have a common underlying cause. In particular, in the prevalent case that contracts are incomplete, the implicit contracting strategies adopted by firms may simultaneously generate involuntary unemployment, and labor market segmentation. To causally identify the impact of contractual incompleteness on (endogenous) unemployment and market segmentation, we implemented experimental labor markets differing in the absence or presence of verifiable effort and complete contracts. We show that involuntary unemployment is much higher when third-party contract enforcement is absent. Moreover, we show that the necessity to provide implicit performance incentives can lead to a segmentation of the labor market. Firms in both segments earn similar profits, but workers in the secondary sector face much less favorable conditions, and exert less effort, than their counterparts in primary-sector jobs.

The third chapter investigates the effects of power concentrating institutions on the quality of political selection, i.e., the voters' capacities to empower competent politicians. In our model, two candidates compete in an election. They can either propose a risky reform or the riskless status quo. Only able candidates can implement a reform that increases ex ante welfare. Candidates are privately informed about their abilities and motivated by welfare considerations as well as the spoils of office.

In equilibrium, a reform proposal is associated with high ability and thus increases electoral prospects. As a consequence, variations in power concentration involve a tradeoff. On the one hand, higher concentration of power enables the voters' preferred politician to enforce larger parts of his agenda. On the other hand, more power concentration leads to more mimicking by low able candidates, due to increased electoral stakes. We find that full power concentration is desirable if and only if politicians are strongly driven by welfare considerations. If politicians are mainly motivated by the spoils of office, on the contrary, some dispersion of power is optimal.

Additionally, we find that the results of an empirical analysis are in line with this theoretical prediction. In a sample of modern democratic countries we study the interaction effect of power concentration and office motivation of politicians on GDP growth. We find that power-dispersing institutions have positive effects on economic growth only in countries in which voters believe that their representatives are mainly driven by private interests. In countries with a more positive attitude towards the political leaders, in contrast, power concentration yields higher GDP growth.

This thesis has benefited of comments from numerous people including seminar and conference participants as well as journal referees and editors. The first chapter has jointly been developed with Steffen Altmann and Armin Falk and is based on the working paper with the same name. The second chapter builds upon on a joint paper with Steffen Altmann, Armin Falk and David Huffman. An earlier version of this paper was circulated under the title: "Implicit Contracts, Unemployment, and Labor Market Segmentation". The third chapter stems form joint work with Emanuel Hansen and Gert Pönitzsch and is based on the working paper: "Political Selection and the Concentration of Political Power".

2. Information and Incentives as Driving Forces of Default Effects

2.1. Introduction

A substantial body of empirical research has shown that non-binding default options strongly affect consumption and savings decisions. Default effects have been documented, for instance, in enrollment and contribution decisions in retirement saving plans (Madrian and Shea 2001, Choi et al. 2004), choices of insurance contracts (Johnson et al. 1993), car purchases (Levav et al. 2010), or consent to postmortem organ donations (Johnson and Goldstein 2003, Abadie and Gay 2006). Based on these empirical findings, a number of scholars have argued that defaults could be a powerful instrument of "libertarian paternalism" (Sunstein 2012b, Camerer et al. 2003). Since they do not restrict choice sets, defaults could help ill-informed individuals making better decisions, without distorting behavior of better informed agents. In this paper, we argue that both, the strength of default effects as well as their influence on agents' welfare fundamentally depend on the intentions of the default setter, and the agents' information about the decision environment.

We provide a theoretical framework to study the influence of these two factors on the behavior of default setters and decision makers. Our model is based on the premise that default specifications as well as individuals' choices are outcomes of a strategic interaction between a default-setter and decision makers: a default setter ("principal") interacts with a population of decision makers ("agents"). Both the principal and the agents have incomplete private information about which of the available choices is optimal for the agents' welfare. In particular, both parties may be uncertain about the preferences of the agent, the exact consequences of the decision or an underlying state of the world. We assume that agents dispose of different levels of information quality according to the decision environment.

The model shows that the alignment of interest as well as the relative level of information affect default specifications and the strength of default effects. First, the behavior of default setters and decision makers depends on how closely their interests are aligned. In principle, the default option may convey the principal's information about the optimal choice for agents. In any Pareto-efficient equilibrium, defaults are more informative, if the principal is more benevolent towards a given population of agents. Defaults specified by a fully benevolent default setter, for instance, always truthfully reveal her beliefs about the best decision for the agents. In contrast, defaults specified by a fully selfish default setter convey no information ("babbling"). Conversely, the reaction of rational agents to a default option depends on its informative value, and is therefore related to the principal's level of benevolence. Second, the model predicts that default effects differ across subgroups of the population: when defaults are at odds with the decision makers' information, agents are more prone to follow defaults, the lower their private information quality is.

In the second part of the paper, we test the key predictions of our model in a laboratory experiment. As a workhorse, we use a simple binary-choice paradigm in which an agent has to decide whether a set of nine cards, which can be either red or black, contains more red cards or more black cards. The ex-ante likelihood that a given set of cards contains more cards of a given color is 0.5. Before making choices, the principal and the agent receive independent and informative signals on the agent's payoffs from the two available choice options. The principal selects one of the options as default option, and the agent can accept the default or opt out. We exogenously vary whether preferences of the principals are (i) fully aligned, (ii) partially aligned, or (iii) misaligned with those of the agents (FUL, PAR, and MIS treatment, respectively). Within each treatment, we additionally vary the relative level of information of principals and agents. In particular, there are always some agents who are better informed than the principal, and some agents whose information quality is below that of principals.

Our empirical results support the notion that strategic incentives as well as the level of information are crucial determinants of behavior for default setters and decision makers. First, defaults truthfully reveal principals' information in 98% of cases in FUL, but only in 75% of cases in PAR and in 56% in MIS. Benevolent principals, thus, select informative defaults, while defaults specified by fully selfish principals barely convey any information. Second, agents are substantially more likely to accept defaults that are chosen by more benevolent principals, with 90%, 74%, and 58% of agents accepting defaults in FUL, PAR, and MIS, respectively. Third, agents' reaction to defaults strongly depends on the quality of their personal information. In FUL and PAR, agents with low and intermediate levels of information strongly rely on the information transmitted through default options. At the same time, decisions of agents with superior information are barely affected by default options in either of the treatments. Well-informed agents almost always opt out when the default is in conflict to their own information. Finally, our results suggest that defaults of (partially) benevolent default setters can in fact enhance the aggregate quality of individual decisions. While choices of well-informed agents are not distorted by default options, agents with lower levels of information attain higher payoffs by accepting default specifications. However, we also find that the effects on welfare are ambigious, if the default setter is fully selfish. While well informed agents are still not affected by default specifications, those with intermediate information seem to follow defaults to frequently, thereby suffering a welfare loss.

Conventional wisdom holds that default options should be irrelevant for the behavior of rational agents as long as the costs of opting out are trivial (see, e.g., Thaler and Sunstein 2003, Schwartz and Scott 2003). This perspective is based upon the assumption that defaults do not convey any information. In contrast, we suggest that defaults arise from a strategic interaction of a default setter and a population of agents. Rational agents should therefore follow defaults whenever the informational quality of the default is sufficiently high. The idea that defaults are behaviorally relevant since they are perceived as implicit recommendation of the default setter has informally been made in the psychological literature that studies the foundations of default effects (Johnson and Goldstein 2003, McKenzie et al. 2006). McKenzie et al. (2006) as well as Tannenbaum (2011) provide evidence from vignette surveys and questionnaires supporting the notion that consumers stick to defaults in order to follow the default setter's recommendation. However, Brown and Krishna (2004) have argued that customers might be skeptical about defaults set by profit maximizing firms and, as a consequence, rely more heavily on active choices. We provide a formal framework that addresses both points of view. It shows that the informational content of default options, and the extent to which rational consumers account for this information both depend on the degree of alignment of interests of default setters and decision makers.

Several further reasons why defaults could influence behavior have been discussed in the literature (see Sunstein 2012a). These include status-quo effects and loss aversion (Kahneman et al. 1991, Sunstein 2002), pecuniary and non-pecuniary costs of opting out (Schwartz and Scott 2003, Thaler and Sunstein 2003), quasi-hyperbolic discounting and individuals' tendency to postpone active decisions (Madrian and Shea 2001, Carroll et al. 2009), or perceptual limitations of decision makers (Caplin and Martin 2011, Caplin and Martin 2012). All of these papers show that psychological motives can explain default effects in a variety of contexts. However, these motives are often context specific. The channel we introduce in this paper, in contrast, is inherent in every interaction with a specified default independently of the psychological motives at play. A loss averse consumer, for example, should take his experience and his knowledge of the decision environment into account, when choosing whether to stay with a default option or to opt out. Even, if consumers tend to postpone decisions, they should consider the intentions of the default setter in the corresponding choice setting. These aspects have largely been neglected in most of the empirical studies on defaults that have assessed decision-making with default options mainly from an individual-choice perspective. We abstract from the decision context and psychological motives. However, our modeling approach is flexible enough to be subsequently enriched by these factors. Therefore, we provide a framework that allows to study the interaction of these alternative factors with incentives and information in shaping the strength of default effects.

Our results also inform the discussion on the scope of libertarian paternalistic policy interventions (Thaler and Sunstein 2003, Camerer et al. 2003, Glaeser 2005). One of the most important concerns against libertarian paternalism is precisely the fear that consumers do not react optimally towards default setters' incentives and information, and thus accept defaults that are against their best interest. Participants in our experiment seem to understand the differences in the informational quality of defaults pretty well, and tailor their reactions accordingly. In particular, the decisions of well informed participants are not distorted by the presence of default options. At the same time, participants with less information benefit from defaults, if specified by a principal with (partially) aligned interests. Overall, defaults may enhance welfare, if decision makers are aware of the intentions of the default setter and the relative levels of information. However, our empirical results also show some important deviations from this pattern. In particular, agents with intermediate information quality seem to follow defaults too frequently when specified by default setters with misaligned preferences. This suggests that a high level of transparency and a sound understanding of default setters' information and strategic incentives are crucial to ensure that consumers react optimally to default options. Policies that promote such transparency—such as information disclosure requirements—are therefore important for protecting consumers from making mistakes in environments where firms or policy makers use default options.

The remainder of the paper is organized as follows. In the next section we present our model and derive testable implications for the empirical analysis. Section 2.3 outlines the design and procedures of the experiment. In Section 2.4 we present the empirical results, and Section 2.5 concludes.

2.2. The model

In what follows, we provide a framework to study the strategic interaction between a default setter ("principal") and a population of decision makers ("agents"). Building on

Crawford and Sobel (1982), we focus on two dimensions that are inherent in any interaction between a default setter and decision makers. First, we analyze how default-setting and decision making are affected by the strategic incentives of the interacting parties. That is, we aim at providing comparative statics on how alignment of incentives between principals and agents influence default specifications, and agents' tendency to stick to the default. Second, we analyze how asymmetric information between principals and agents influences decision making. In particular, the principal and all agents are heterogeneous in the quality of their information according to which decision alternative maximizes agents' welfare. This heterogeneity enables us to study how the relative level of information between a decision maker and the default setter shapes default effects. All proofs are found in appendix A.1.1.

2.2.1. Model setup

In the model one default-setting principal interacts with a population of agents. These are assumed to be rational decision makers who choose an action z in order to maximize their utility $U_A(z, \theta)$, which is strictly concave in z for all θ. The state of the world, θ, can take value θ_h or θ_l, which determine the preferences of the agents over the choice variable. More precisely, we assume that the partial derivative with respect to z of $U_A(z, \theta_l)$ is smaller than the corresponding derivative of $U_A(z, \theta_h)$ for all z, i.e, higher zs are optimal if the agent puts more probability weight on θ_h. The action z is a discrete choice variable $z \in \{z_1, \ldots, z_m\}$. Both states of the world are equally likely and unobservable for the agent. All players have incomplete information about the true state of the world. This may be due to a complex decision environment, unclear preferences of the agents or uncertainty about an underlying variable that has not been realized, yet. Agents dispose of different levels of information quality. In particular, we assume that the population of agents is distributed according to $f(x)$, with full support over $[\frac{1}{2}, 1]$, where x is the signal strength of agent x. Agents' signals are denoted by $\sigma \in \{\sigma_l, \sigma_h\}$ and have conditional distribution $p(\sigma_l|\theta_l) = p(\sigma_h|\theta_h) = x$.

The principal as well possesses private information about the optimal decision for the agents. Her signal is drawn independently from the ones of the agents and denoted by $\rho \in \{\rho_l, \rho_h\}$ with quality $q = p(\rho_l|\theta_l) = p(\rho_h|\theta_h) \in [\frac{1}{2}, 1]$. Consequently, the principal is always better informed about the state of the world than some agents, and worse than others. While the principal cannot influence the agents' decision directly, she can choose a default option $d \in \{d_1, \ldots, d_n\}$ prior to the agents' decision. The utility of the principal is the weighted sum of the utility of the agent and a term, $b(z)$, that captures a potential conflict of interests between the principal and the agents:

$$U_P(\theta, z) = \mu U_A(\theta, z) + b(z).$$

The parameter μ describes the benevolence of the principal. For small μ, the principal cares only little about the agents' well being and focuses more strongly on her private interests. For large μ, however, preferences of the principal and the agents are strongly aligned. A firm that anticipates a repeated long run interaction with the consumer may for example weigh consumer satisfaction more strongly than a firm that interacts only once with each consumer. Similarly, the alignment of interests between a government and agents may be stronger, than it may be in the case of a customer and a profit maximizing company. To put structure on the preferences of the principal, we assume an upward bias in the sense that $b(z)$ is strictly increasing in z.

The game is divided into four periods. In period 1 nature draws the state of the world, θ, and all private signals. Moreover, all agents and the principal observe their corresponding signal. In period 2 the principal decides on the default option d. In period 3, the default is transmitted to the agents, and they decide individually on which choice z they implement. Dependent on the choice and the true state of the world, payoffs for agents and the principal are realized in period 4. In this setup, a strategy of the principal is a mapping

$$s_P : \rho \to (p(d_1), \ldots, p(d_n)) \in [0, 1]^n$$

which specifies the probability of every default to be chosen for all possible signals ρ. A strategy of the principal, thus, determines the correlation between any default and her signal. For any combination of a private signal σ and default d, agents choose z. A pure strategy[1] of an agent with signal strength x is thus a mapping

$$s_A^x : (\sigma, d) \to z \in \{z_1, \ldots, z_m\}.$$

Define s to be a strategy profile, consisting of a strategy of the principal and the strategy of every agent. As a solution concept we apply Bayesian equilibrium. Hence, all players maximize their utility at any information set occurring with positive probability given their type and the strategies of all opponents.

2.2.2. Equilibrium analysis

In the following analysis we concentrate on equilibria in which all defaults are played with positive probability. This is without loss of generality, since every equilibrium can be

[1]In principle an agent may be indifferent between different actions z and mix over the actions in equilibrium. Nevertheless, the set of agents that are indifferent between at least two messages has mass zero. Hence, extending the analysis to mixed strategies does not allow for any additional insights.

replicated by an output equivalent equilibrium without out of equilibrium defaults. If there is an off equilibrium path default, it is always possible to construct a principal's strategy which mixes over this default and an equilibrium default such that both convey the same information. Consequently, the best response of the agent is to treat them identically. Since the principal is, thus, indifferent between both defaults, the new strategy profile constitutes an output equivalent equilibrium. More precisely, the ex ante probability that a particular utility outcome for the agents and the principal is realized is identical.

Lemma 2.1. *For any Bayesian equilibrium s there exists an output equivalent Bayesian equilibrium in which all defaults are played with positive probability.*

Note that any Bayesian equilibrium without out of equilibrium messages also constitutes a perfect Bayesian equilibrium. Hence, agents choose action z to maximize the expected utility $E(U_A|d,\rho,x)$ at any information set. Once an agent received his signal and the principal has decided upon a default, agent x solves

$$\max_z E_\theta(U_A|d,\sigma,x) = p(\theta_l|d,\sigma,x)U_A(z,\theta_l) + p(\theta_h|d,\sigma,x)U_A(z,\theta_h).$$

Before the agents decide on z, the principal can influence them by choosing a default option to ensure a more favorable outcome to herself. Consider a perfect Bayesian equilibrium with strategy profile s. The principal's expected payoff from playing default k in information set i associated with signal ρ_i, is given by

$$\int_{\frac{1}{2}}^{1} p(\theta_l,\sigma_l|\rho_i,x)U_P(s_A^x(\sigma_l,d_k),\theta_l) + p(\theta_h,\sigma_l|\rho_i,x)U_P(s_A^x(\sigma_l,d_k),\theta_h)$$
$$+ p(\theta_l,\sigma_h|\rho_i,x)U_P(s_A^x(\sigma_h,d_k),\theta_l) + p(\theta_h,\sigma_h|\rho_i,x)U_P(s_A^x(\sigma_h,d_k),\theta_h)dF(x).$$

Whenever the principal receives a high signal, her preferences are aligned with those of the agents. This is true independently of the level of μ, since we assumed an upward bias of the principal. In this case, a higher z maximizes both, her selfish interests $b(z)$ and the utility of the agents. Hence, the principal chooses the default which induces the highest z. As a consequence, all messages that are played with positive probability after a high signal must induce the same z and are qualitatively identical. The remaining defaults reveal that the principal received a low signal. According to this distinction we bisect the set of defaults $\{d_0,\ldots,d_n\}$. First, those defaults that are exclusively played after a low signal $\{d_0,\ldots d_{k-1}\} = D_l$, and second default options $\{d_k\ldots,d_n\} = D_h$ which are played with positive probability after a high signal.

Lemma 2.2. *In every equilibrium the principal will send at most two qualitatively different messages.*

Using the lemma we are able to describe the information transmission by the principal by a single parameter $c = p(d_0|\rho_l) + \cdots + p(d_{k-1}|\rho_l)$. The larger c, the more often the principal truthfully reveals her signal after a low signal, leading to more information transmission. Note that the meaning of every default arises endogenously in equilibrium and is not exogenously given. Clearly, all permutations of messages also support an equilibrium. We name all defaults played with positive probability after a high signal a high default (d_h) and those that are exclusively played after a low signal a low default (d_l).

2.2.3. Principal's level of benevolence

The decision of rational agents to follow defaults specified by a principal is likely to depend on the principal's level of benevolence. A principal with opposing interests will use his information to extract a higher rent for herself. In contrast, a more benevolent principal will specify a default that is more trustworthy for the agents and should thus more likely be followed. In the extreme case of full alignment of preferences, the principal has incentives to truthfully transmit all of his information.

Proposition 2.1. *Fixing everything else, there exists a $\bar{\mu}$ such that for all $\mu > \bar{\mu}$ the Pareto-dominant equilibrium exhibits complete information transmission. Moreover, there exists a $\underline{\mu}$ such that for all $\mu < \underline{\mu}$ defaults exhibit no informational content in equilibrium.*

If preferences are not fully aligned, the principal may choose to conceal some of her information. While the existence of equilibria is still clear for this case there may emerge multiple equilibria exhibiting information transmission. To rank those, we employ the Pareto-efficiency criterion. Agents always prefer an equilibrium corresponding to a larger c to one with less information transmission, since more information yields a better decision and thus a higher expected payoff. The same is true for the principal, if she receives a high signal, because preferences are aligned in this case. If the principal receives a low signal she is indifferent between all mixed equilibria. To see this, suppose the principal chooses a message from the set D_l in a mixed equilibrium. Since this action reveals her signal, the inherent information and the expected payoff associated with a low default are equal in all mixed equilibria. Furthermore, the principal is indifferent between a low and a high default. Hence, the expected payoff after a high default must also be identical across mixed equilibria. Overall, the principal prefers the mixed equilibrium with the highest information transmission rate. If the expected payoff following a high default is larger than the one corresponding to a low default for full information transmission, the principal clearly prefers the full information equilibrium to all mixed equilibria. We conclude that the equilibria featuring the highest information transmission rate are Pareto-efficient.

Lemma 2.3. *Consider two equilibria with corresponding $c_1 < c_2$. Then the equilibrium corresponding to c_2 Pareto-dominates the one belonging to c_1.*

In the following, we focus on the question how a Pareto-efficient equilibrium depends on the model's main parameters. Define $c^{pd}(\mu)$ to be the information transmission rate in any Pareto-dominant equilibrium. As argued above, μ represents the degree to which principal's and agents' preferences are aligned. For increasing μ, any Pareto-efficient equilibrium becomes more favorable for the agents.

Proposition 2.2. *$c^{pd}(\mu)$ is weakly increasing in μ. Agents internalize this effect and exhibit more pronounced reactions to defaults, i.e, $s^x_A(\sigma_i, d_h)$ is weakly increasing and $s^x_A(\sigma_i, d_l)$ is weakly decreasing in μ.*

In any equilibrium with full information transmission this is trivially the case. Consider the case of a Pareto-dominant equilibrium with a mixed strategy of the principal. Then she is indifferent between specifying a high or a low default in response to a low signal:

$$\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s^x_A(\sigma, d_h)) - U_P(\theta, s^x_A(\sigma, d_l))|x, \rho_l]dF(x) = 0.$$

With increasing μ the principal's preferences become more strongly aligned to those of the agents. As a consequence, her utility of a truthful report increases relative to the utility of concealing information, hence, the difference above is decreasing in μ for all x. This effect is offset along equilibrium path by a higher transmission rate c. While the second part of the difference does not change with c - playing a default from the set D_l always conveys the same information and thus induces the same action z-, the first part must be increasing in c. We show in the appendix, that if this were not the case, there would always exist an equilibrium with a higher transmission rate which is a contradiction to the assumption of being in a Pareto-efficient equilibrium. In general, the proposition implies that a higher level of benevolence leads to a higher informational content of the defaults, and thus ceteris paribus to a higher incentive for agents to adapt their choice towards the default. In particular, in the case of a conflicting private signal and default, agents weigh the default more strongly. Hence, default effects are stronger in a given population of agents if interests of the population and the default setter are more aligned.

2.2.4. The quality of information

Differences in agents' knowledge about the decision environment are an intuitive explanation why the stickiness of defaults might differ across subgroups of the population. Intuitively, agents that are less familiar with the choice environment rely more heavily on the information entailed in the default option. In our model, such differences in agents' information are captured by the agents' signal strength, $x = p(\theta_l|\sigma_l) = p(\theta_h|\sigma_h)$.

Proposition 2.3. *In equilibrium, the strategies of any two agents with $x_1 < x_2$ exhibit the following properties:*

$$s_A^{x_1}(\sigma_l, d_i) \geq s_A^{x_2}(\sigma_l, d_i) \qquad s_A^{x_1}(\sigma_h, d_i) \leq s_A^{x_2}(\sigma_h, d_i) \quad \textit{for} \quad i \in \{l, h\}$$

Agents are less susceptible to adapt their decision towards a given default if the quality of their personal information increases. The relationship between agents' knowledge and the degree of default adherence is most clearly seen if the default and the agent's private signal are in conflict. For instance, the strategy $s_A^x(\sigma_l, d_k)$, which describes agents' behavior after a high default and a low personal signal, is decreasing in the information quality x. Consequently, agents with more informative low signals shift "further" away from the high default and rely more strongly on their own information about the optimal decision.

2.3. Design of the experiment

2.3.1. The game

In the experiment, we aim at testing the key comparative statics of the model empirically. Firstly, we endogenously vary the alignment of interests between default setters and decision makers. We consider three different situations in which the preferences of the default setter and the agents are (i) fully aligned, (ii) partially aligned, or (iii) misaligned (in what follows, we refer to the three conditions as the "FUL", "PAR", and "MIS" treatment, respectively). Secondly, for a given level of benevolence, our empirical approach ensures controlled variation in the relative level of information between the default setter and the agents. Some agents are better informed than the principal about which decision maximizes their payoff, while others are less informed.

As a workhorse for implementing these treatment conditions, we use a simple paradigm in which one principal interacts with one agent. In each period, the agent has to decide whether a set of nine cards, which can either be red or black, contains more red cards or more black cards. Each card in each period is drawn independently with probabilities $p(Red) = p(Black) = 0.5$, i.e., the ex-ante likelihood that a given set of cards contains more cards of a given color is 0.5. Before making choices, the principal and the agent receive independent signals about the composition of the current set of cards.

The principal receives the signal via a message on her screen. The message can either indicate that the current set of cards contains "more black cards", or "more red cards". The signal is private information of the principal. Her signal strength, however, is common knowledge and held constant at $q = 0.8$. Whether a principal receives a correct or wrong signal in a given period is determined randomly and independently between principals, periods, and sessions.

Agents receive information about the number of red and black cards in the current set, by a subset of cards, which is privately revealed to the agent. In each period, a coin flip determines whether the first *two* or the first *five* cards in the set are uncovered for a given agent.[2] The signal-generating mechanism for agents ensures that we obtain controlled variation of the information quality for different types of agents. Since each of the revealed cards for the agent is black or red with probability 0.5, we obtain five different levels of signal strengths. The resulting distribution of agents' signal qualities is reported in Table 2.1. Since the principal's signal strength is always 0.8, the agent is informed worse than the principal in about 56% of cases (Columns 1 and 2 in Table 2.1); in about 25% of cases he has about the same signal strength (Column 3 in Table 2.1), and in 19% of cases (Columns 4 and 5), the agent is better informed than the principal. As a within-subject dimension, we can test whether agent types 1–5 systematically differ in their behavior, holding principals' information quality and benevolence constant (Proposition 2).

Agent Type	1	2	3	4	5
Signal Quality	0.5	0.69	0.77	0.94	1.00
Occurrence Probability	0.25	0.313	0.25	0.156	0.031
Example	1 black	3 black	2 black	4 black	5 black
	1 red	2 red	0 red	1 red	0 red

Table 2.1.: *Distribution of agent's signals.*

Participants in the experiment are not shown Table 2.1, but the signal-generating mechanism for principals and agents is common knowledge. In particular, the principal knows the procedure how the agent is informed, and consequently the distribution of signal qualities for the agent. However, she is not informed about the number or colors of the cards that are revealed to the agents. Hence, the principal essentially plays against a population of agents as depicted in Table 2.1, although principals and agents are matched one-to-one in a given period.

After having received her signal, the principal selects the default option for the current period. She can either specify "more red cards" or "more black cards" as the default. In a next step, the agent is informed about the default option. He can then accept the default or opt out and take an active decision. To accept or change the default option in the experiment, the agent has to press the respective button displayed on his screen. In case the agent accepts the default option, the selected default is implemented as final decision for both the principal and the agent. If the agent presses "opt out", a new screen pops

[2] Figure A.1 in appendix A.1.2 depicts an example of an agent's information screen.

up on which the agent can (actively) choose between "more red cards" and "more black cards" as her ultimate decision for this period.[3] After the agent has taken his decision, he is asked to state his perceived certainty that his choice was correct on an 8-points Likert scale. In a final stage of the game, principal and agent are provided with a feedback screen, on which the entire set of cards for the current period is revealed. Furthermore, players are informed about the agent's final decision and the resulting payoffs for both players.

2.3.2. Treatments and payoffs

Agents earn points, if their decision in a given period is correct (i.e., the chosen color matches the color that occurs more frequently in their current set of cards). In each of the three treatments, payoffs for agents were calculated as follows:

$$\pi_A = \begin{cases} 50 \text{ points} & \text{if decision correct} \\ 0 & \text{if decision wrong} \end{cases}$$

Payoff functions for principals differed across treatment conditions. In treatment FUL, principals' payoffs are perfectly aligned with those of the agents. Thus, a principal receives 50 points if the agent's decision in a given period is correct, and 0 otherwise:

$$\pi_P^{FUL} = \begin{cases} 50 \text{ points} & \text{if A's decision correct} \\ 0 & \text{if A's decision wrong} \end{cases}$$

In contrast, principals in the treatment with misaligned preferences (MIS), receive 50 points if and only if the agent's decision in a given period is "more red cards". They receive 0 points if the agent's decision is "more black cards":

$$\pi_P^{MIS} = \begin{cases} 50 \text{ points} & \text{if A's decision "red"} \\ 0 & \text{if A's decision "black"} \end{cases}$$

This payoff function resembles a principal intending to direct the decision of the agent towards one particular alternative, without taking the welfare consequences for the agent into account, i.e, a fully selfish principal. To induce partial benevolence, each matching group in our third treatment, PAR, consisted half of benevolent and half of selfish principals with payoff functions π_P^{FUL} and π_P^{MIS}, respectively. Agents are not informed about which type of principal they are matched with in a given period. From agents' (ex ante) perspective, this treatment is thus equivalent to interacting with a partially benevolent principal.

[3] We neither impose a time limit for principals setting the default nor for agents making their final decision. This procedure ensures that the cost of opting out of the default is minimal, while at the same time avoiding mistakes due to time pressure or accidental clicks.

2.3.3. Parameters and procedures

Overall, we conducted 12 sessions of the experiment; four sessions for each treatment. In each of the sessions, we had 12 principals and 12 agents that interacted for 50 periods. Subjects within a session were divided into two matching groups with 12 participants each. Principals and agents within a given matching group were randomly rematched between periods, yielding 8 independent observations per treatment for the non-parametric tests reported below. Points earned throughout the experiment were converted at an exchange rate of 100 points = 1 Euro. Overall, sessions lasted about 120 minutes, and subjects earned on average 24.32 euros (about 32 USD at the time of the experiment), including a showup fee of 4 euros.

All sessions were carried out in the BonnEconLab, the laboratory for economic experiments at the University of Bonn. The experiment was computerized using the software z-Tree (Fischbacher 2007), and subjects were recruited with the online recruitment system by Greiner (2003). A total of 288 subjects (96 in each treatment) took part in the experiment. Subjects were mainly undergraduate university students from all majors, and participated in only one of the treatment conditions. To ensure common knowledge of the rules and structure of the experiment, a summary of the instructions for the respective treatment was read out aloud at the beginning of each session. Participants then received detailed written information about the experiment.[4] The experiment started only after all participants had answered several control questions correctly.

2.3.4. Hypotheses

Applying our model to the setup and parameters of the experiment yields the following predictions for differences in behavior between treatments and agent types.

Hypothesis 2.1. *Principals' propensity to truthfully reveal their signal through defaults is highest in FUL, intermediate in PAR, and lowest in MIS.*

In particular, principals in FUL should always truthfully reveal their signal. In contrast, defaults specified by fully selfish principals (MIS) should convey no information (see Proposition 2.1). Since 50% of principals in PAR are benevolent and 50 % are selfish, the truthfulness of defaults in this treatment should lie in between MIS and FUL.

Hypothesis 2.2. *Agents' aggregate propensity to accept defaults should be highest in FUL, intermediate in PAR, and lowest in MIS.*

[4] A translation of the verbal summary and the instructions is available upon request.

This aggregate hypothesis is derived from a more specific sub-hypothesis. Proposition 2.2 implies that the strength of default effects is weakly increasing in the benevolence of the principal. Applying Proposition 2.2 to the parameters of the experiment it predicts that agents' propensity to accept a default that is in conflict with their private information should be strictly higher in FUL than in PAR, and strictly higher in PAR than in MIS for "low-information types" (type 1, 2, and 3 in Table 2.1).[5] In contrast, the model predicts no treatment difference in default adherence after conflicting signals for "high-information types" (type 4 and 5 in Table 2.1). Similarly the application of Proposition 2.3 to each of the treatments yields the following hypothesis for behavioral differences in the within-subjects dimension (behavior of low-information vs. high-information types within a given treatment):

Hypothesis 2.3. *In case of conflicting signals, low-information agents should be strictly more likely to accept defaults than high-information agents in FUL and PAR. There should be no difference in acceptance rates between low-information and high-information types in MIS.*

Finally, the above predictions for differences in behavior yield the following hypothesis for differences in the overall quality of agents' decisions as measured by the resulting monetary payoffs.

Hypothesis 2.4. *For low-information types, the quality of decisions in FUL should be strictly higher than in PAR which in turn should be strictly higher than in MIS. The quality of decisions for high-information types should not be affected by principals' level of benevolence (FUL=PAR=MIS).*

2.4. Results

In this section, we present the results of the experiment. We first summarize the behavior of principals and analyze whether the informational content of defaults differs across treatments (Hypothesis 1). We then focus on the agents, and study how agents react to defaults in FUL, PAR, and MIS (Hypothesis 2). In a next step, we analyze differences in agents' behavior, depending on their relative level of information (Hypothesis 3). Finally, we compare the quality of decisions for the different types of agents across treatments (Hypothesis 4).

[5] Since some predictions for individual types of agents who have lower signal quality than the principal do not differ, we jointly denote types 1, 2, 3 from Table 2.1 as "low-information types" for ease of exposition. Types 4 and 5 are denoted as "high-information types", accordingly.

2.4.1. How do principals set defaults?

Figure 2.1 summarizes principals' behavior in the different treatments. The figure depicts the average frequency of defaults that are specified according to the private signal of principals about the state of the world, ρ. In line with Hypothesis 1, we observe strong treatment differences in the likelihood that defaults truthfully reveal principals' private signal. Fully benevolent principals (FUL) almost always reveal their signal (in 98% of cases). In contrast, principals reveal their private signal only in 75% of cases in PAR. Selfish principals set the default according to their private signal in only 56% of cases. Hence, defaults in the MIS treatment convey hardly any information about the principals' knowledge of the state of the world. The overall difference in the information quality of

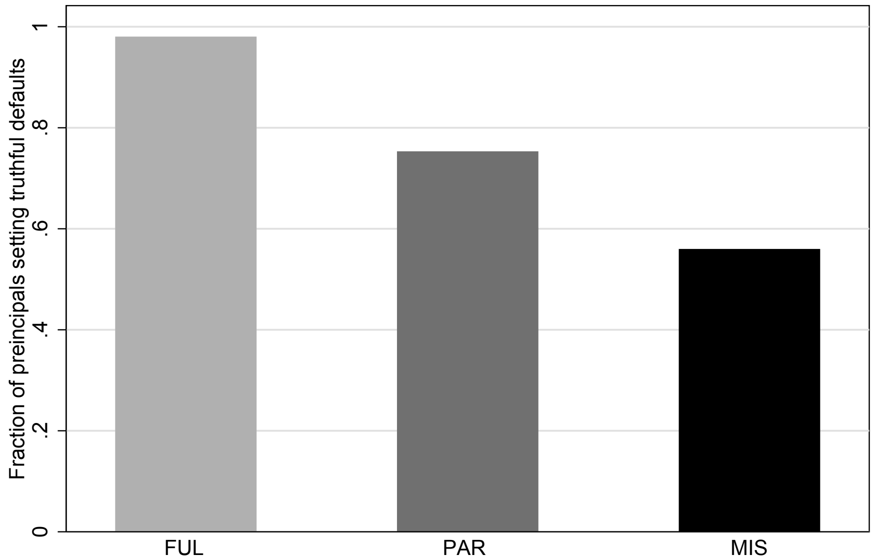


Figure 2.1.: *Frequency of principals setting default according to their signal. Average values per period in FUL, PAR, and MIS treatment.*

defaults is highly significant for all pairwise treatment comparisons (Fisher-exact tests, $p < .001$ for FUL vs. PAR, FUL vs. MIS, and PAR vs. MIS).[6] Figure A.2 in the appendix A.1.2 shows that principals' default-setting strategy is relatively stable over time in all treatments. The strong difference between treatments is already observed in the first periods, and we find no significant time trend in either of the treatments.

Overall, principals' behavior in the experiment matches the predictions accurately. A small, but noteworthy exception is observed in the MIS treatment. The model predicts that defaults set by fully selfish principals convey no information in equilibrium. That is,

[6]Unless otherwise noted, all non-parametric tests are based on matching-group averages. Reported p-values are always two-sided.

the likelihood of a principals' signal to be red or black, should be exactly 50% independently of the observed default. Principals' behavior in the experiment comes close to this prediction. However, the actual frequency of informative defaults in MIS is 56%. There could be two potential explanations for this deviation from the model's point predictions. First, the higher informativeness of defaults could be explained by a failure of principals to completely shade their private information. Second, some principals in MIS could be not fully selfish in the sense that they deliberately provide agents with information about their private signal. This could, for instance, be due to preferences for honesty (Charness and Dufwenberg 2006, Erat and Gneezy 2012, Fischbacher and Heusi 2008), aversion towards payoff inequalities (Fehr and Schmidt 1999), or other forms of social preferences. Put differently, the true preferences of some principals might not coincide with the monetary incentives induced in the MIS treatment.

To shed light on which of the two potential explanations is likely to drive the deviation, we study behavior of individual principals in the MIS treatment. As a measure for the truthfulness of reports we calculate the average frequency of a realized black default after a principal received a black signal.[7] We then estimate an OLS model with the principals' individual frequency as dependent variable and different potential determinants. As proxies for the "cognitive mistakes" explanation, we include principals' scores in the Cognitive Reflection Test, CRT (Frederick 2005), and principals' final math grade in high school.[8] As a proxy for the "social preference" explanation, we include the "Honesty-Humility-Scale", HHS—a subscale of the HEXACO personality questionnaire designed to measure an individual's inclination to avoid manipulation of others for personal gain (Ashton and Lee 2009).[9]

Table A.1, in appendix A.1.2, shows that principals who score higher on the Honesty-Humility-Scale have a weakly significant higher likelihood to truthfully reveal a black signal in the MIS treatment (p=.064). In contrast, the proxies for cognitive reflection and math abilities are not significantly related to principals' behavior. All in all, the results provide tentative evidence that the "too informative" defaults observed in the MIS are due to some principals' preferences for honest behavior, rather than due to strategic errors by the principals.

[7]We concentrate on the case of black signals since principals almost unanimously select red defaults after a red signal (in 98.6% of cases). That is, principals mostly choose to shade their private information by always selecting "red" as the default (the overall frequency of red defaults in MAL is 93.4%).

[8]The CRT is a three-item questionnaire that measures an individual's inclination to suppress a spontaneous, but wrong answer. High school grades range from 1-6 with 1 being the best grade.

[9]The HHS ranges from -60 to +40.Higher scores indicate a higher inclination to avoid manipulations.

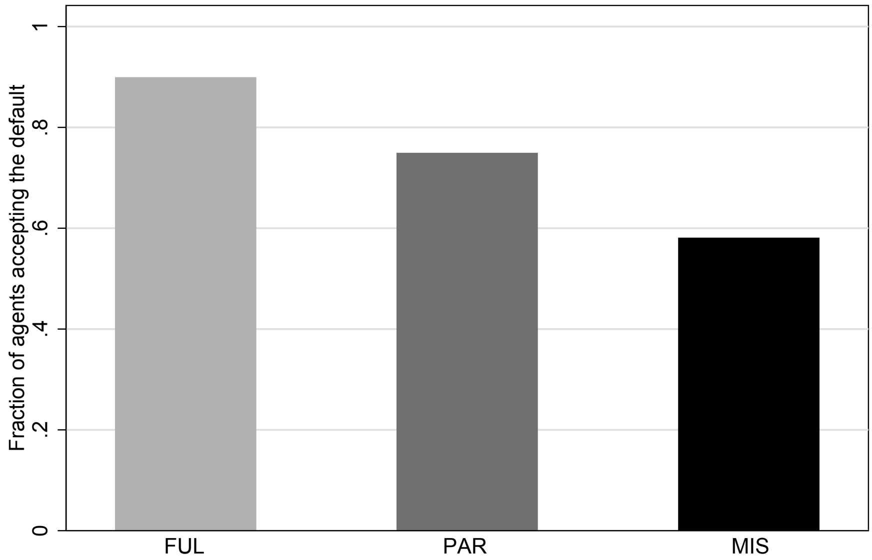


Figure 2.2.: *Frequency of default adherence. Average values in FUL, PAR, and MIS.*

2.4.2. How do agents react to defaults?

In the next step, we turn to the analysis of agents' reactions to defaults in the different treatments. Figure 2.2 depicts the aggregate frequencies of default adherence for agents in FUL, PAR, and MIS. Agents accept the default specified by principals in 90% of cases in FUL, 74% of cases in PAR, and 58% of cases in MIS. All pairwise treatment differences in agents' behavior are statistically significant (Fisher-exact tests, $p < .001$ for FUL vs. PAR, FUL vs. MIS, and PAR vs. MIS). Figure A.3 in appendix A.1.2 depicts the frequency of default acceptance in the different treatments over time. Again, we observe a strong difference between treatments already in the first periods and relatively stable behavior over time. The only treatment exhibiting a significant decrease in default acceptance over time is the MIS treatment.[10] Similar to principals' behavior, agents match the predictions of the theoretical model well. The aggregate treatment comparisons for agents' default acceptance support the comparative static predictions of our model (Hypothesis 2).

We conclude that agents account for differences in the principals' benevolence and the resulting differences in the informativeness of defaults on the aggregate level. This also holds for each type of agent. The top panel of Figure 2.3 depicts the average frequency of default acceptance for the five different types of agents in the different treatments.[11] Sub-

[10] A linear time trend is significant at the 5% level in a probit estimation where the dependent variable is 1 if an agent accepts the default, and 0 otherwise (p=.030, accounting for potential clustering in standard errors on the matching-group level).

[11] Agents are ordered according to their signal quality. Type "A69" in Figure 2.3 are agents who received an a signal with 69 % correlation in a given period (Type 2 in Table 2.1).

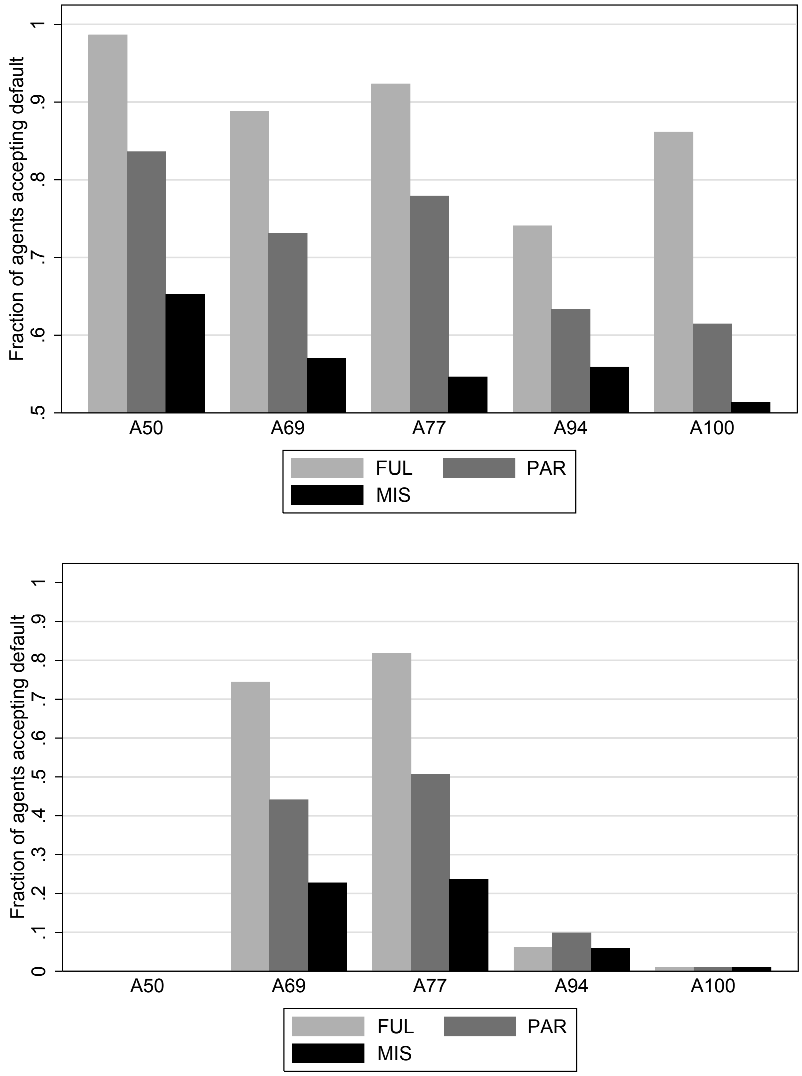


Figure 2.3.: *Fraction of agents accepting the default. Upper panel: average aggregate values for different agent types in FUL, PAR, and MIS. Bottom panel: cases where default is in conflict with agents' private signal; average values for different agent types in FUL, PAR, and MIS.*

stantiating our previous results, we find that the likelihood of accepting defaults increases in the benevolence of the principal for each individual type of agent.

Next we analyze agents' reactions to defaults that are at odds with their private signal. The bottom panel of Figure 2.3 shows the behavior of the different types of agents, if they face a default that is in conflict with their private signal (e.g., the agent observes three black and two red cards, but a red default).[12] The data depicted in the graph allow us to directly test the sub-hypothesis implied by Proposition 2.2. In particular, we predicted that low-information types are strictly more likely to accept defaults in FUL than in PAR

[12]Note that type A50 is excluded from the analysis. Since this types has no informative signal (1 red card and 1 black card), there can be no conflicts between own signal and the default.

after conflicting signals, and strictly more likely in PAR than in MIS. In contrast, we expect no treatment effects for agents who have superior information than the principal.

Indeed, we find that the default adherence rate for low-information types if the default is in conflict with their private signal is increasing in the benevolence of the principal (Fisher-exact test for low-information types. FUL vs. PAR: $p < .001$, FUL vs. MIS: $p < .001$, PAR vs MIS: $p = .010$). On the contrary, the reactions of high-information types do not differ significantly across treatments (Fisher-exact test for high-information types. FUL vs. PAR: $p = 1.000$, FUL vs. MIS: $p = .619$, PAR vs MIS: $p = .619$).

The bottom panel of Figure 2.3 also allows to test Hypothesis 3, which predicts that high-information types are less likely to accept conflicting defaults than agents with lower quality information when facing a (partially) benevolent default setter. This within treatment effect is born out by the data (Wilcoxon signed-rank tests for high-information types vs. low-information types, $p = .012$ for FUL and PAR). However, the difference remains significant when analyzing the MIS treatment, while we hypothesized that there are no type-specific differences in behavior as a response to selfish defaults (Wilcoxon signed-rank test, $p = .012$). The top panel of Figure 2.3 shows that the diverging reactions of agents to conflicting defaults induce overall differences in type specific default acceptance rates. Agents who are better informed than the principal (i.e., Type A94 and A100) are less likely to accept defaults than agents with inferior information quality if the principal is (partially) benevolent. A Wilcoxon signed-rank test confirms that high-information types behave significantly different from low-information types in both the FUL and PAR treatment ($p = .012$, for FUL and PAR). When interacting with selfish principals, the effect is less pronounced and turns out to be insignificant (Wilcoxon signed-rank test, $p = .161$). This is in line with the prediction of our theoretical model: Since defaults convey no information in a babbling equilibrium, all types of agents should merely rely on their private signal.

Overall, the empirical results support Hypothesis 2 and Hypothesis 3 from our model. However, we again find some deviations from the model's predictions. In particular, agents with low and intermediate levels of information seem to trust defaults in the MIS treatment "too much". This becomes most evident when looking at behavior of type A69 and A77 in the case of conflicting signals (bottom panel of Figure 2.3). Our model predicts that all types of agents completely ignore defaults set by a fully selfish principal. For conflicting signals, we thus expect default acceptance rates close to zero for A69 and A77, whereas the acceptance rates in the experiment are about 20% for both types of agents.[13]

[13] A similar effect is observed in the PAR treatment where we observe acceptance rates of 40-50% instead of 20-30% as predicted by our model.

We have already seen that some principals in the MIS treatment do not behave in a fully selfish way. That is, defaults in the MIS treatment are—on the aggregate level—not fully uninformative from the agents' perspective. This raises the question whether the default acceptance rate by low-information agents that we observe is caused by "too much trust" in selfish defaults (i.e., a mistake by low-information agents) or rather a best response to the informational content of defaults, that agents experience during the experiment. If agents' choices are a best response to the behavior of principals, the experienced profits of agents should be as least as high as in a hypothetical situation without defaults in which agents always follow their private signal. The next section discusses, if the decision quality of agents is positively influenced by the presence defaults.

2.4.3. Do defaults improve decisions?

In a final step of our empirical analysis, we turn to the question whether defaults in our experiment improve the overall quality of agents' decisions, and how this depends on the agents' level of information and the principals' level of benevolence. Figure 2.4 depicts the percentage change in the agents' decision quality compared to a hypothetical situation in which agents always follow their private signal. This behavior resembles a setting without default, since agents can only adapt their decision according to their private signal. To evaluate this situation, we take the actually realized signals and sets of cards from the experiment. Positive values indicate that agents profited from defaults, while negative values mean that they would have attained higher profits without default options.

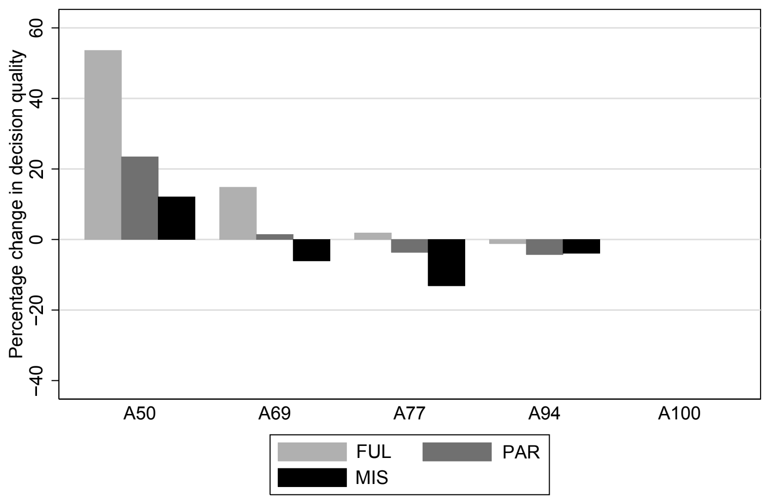


Figure 2.4.: *Percentage change of agents' decision quality compared to a hypothetical situation, in which agents always follow their signal, for different agent types and treatments.*

For high-information types (A94, A100), figure 2.4 indicates that profits are hardly affected by the presence of defaults independently of the alignment of preferences. This reflects the observation that these types react appropriately to their high signal quality and only rely on their private signal. Thus, there are little or no distortions in the decisions of well-informed agents.

For agents with lower information quality, the extent to which default options are welfare enhancing depends on the agent's level of information and, more importantly, on the benevolence of the default setter. Compared to the situation without default specifications, we observe an unambiguously positive effect of default options in the FUL treatment where principal's and agents' information are fully aligned. Agents with lower information quality make use of the informativeness of defaults, which are set by a better informed default setter, and thereby attain higher profits.

For the PAR and MIS treatment, our findings on the beneficial or detrimental effects of default options for agents with low information quality are somewhat mixed. First, we observe unambiguously positive effects on the decision quality of agents with no information (Type A50). For Type A69 who has intermediate level quality, defaults set by partially benevolent or selfish default setters have almost no profit consequences. Their decision quality is, on average, not distorted relative to the benchmark without default. Finally, profits of Type A77 suggest that agents with informational levels similar to the one of the principal do not discount the informational content of the default strong enough to account for the misalignment of preferences[14]. Overall, this finding suggests that a sound understanding of the relative quality of private and default setters' information and strategic incentives are crucial for agents to reap the benefits of defaults without bearing detrimental consequences if defaults are specified by principals with misaligned interest.

2.5. Conclusions

In the first part of this paper, we have provided a theoretical model focusing on two key channels for behavioral consequences of non-binding default options—the strategic incentives of default setting institutions, and the default setter's and decision makers' relative level of information about the decision environment. While both effects have been discussed informally in the literature, our model provides a unifying framework for analyzing how individual behavior and the overall strength of default effects are shaped by both factors. We derive two testable predictions. First, default options should have

[14]This lack of discounting has also been documented by Cain et al. (2005), who find that agents trust the information given by selfish advisers too strongly.

a stronger impact on behavior, if interests of the default setter and decision makers are more closely aligned. Second, whenever decision makers experience conflicts between their own information about optimal choices and a default option, those with less information are more prone to accept the default.

In the second part of the paper, we test these predictions in a laboratory experiment that allows us to exogenously vary the parameters of interest. Overall our empirical results match the theoretical predictions well. We find that the informational content of defaults and, thus, the strength of default effects are increasing in the alignment of interests. In the second dimension, we show that agents with lower information quality are indeed more susceptible to accept default options, if these are in conflict with an agent's individual information.

On a more general level our analysis suggests that a more integrative perspective on defaults is important. It ought to incorporate both, psychological factors that may strengthen default effects like a status quo bias or present biased preference (Samuelson and Zeckhauser 1988, Carroll et al. 2009) but also the incentives created by the strategic interaction that is inherent in the default setting process. Our model is flexible enough to be extended such that the interaction of these effects can be studied. Further interesting extensions include an integration of the possibility that agents can acquire additional information (Caplin and Martin 2012), the possibility of heterogeneous preferences of agents for a given state of the world and to allow for imperfect information about the opponents interests.

Our findings also provide insights for the evaluation of "libertarian paternalistic" policy interventions. In particular, they indicate that agents do take the default setter's strategic incentives and information into account, and condition their acceptance of default options on both factors. This suggests that defaults might indeed only influence the part of the population that has inferior information about making "good" decisions. At the same time, decisions of individuals that have a sound understanding of what is the best option for themselves are not distorted by the presence of default options. However, our empirical results also suggest that a high level of comprehension of default setters' information and strategic incentives are crucial to ensure that consumers react appropriately to default options.

3. Contractual Incompleteness, Unemployment, and Labor Market Segmentation

3.1. Introduction

This paper provides evidence that two important features of labor markets—the existence of involuntary unemployment, and the segmentation of markets into firms offering "good" and "bad" jobs to apparently similar workers—may have a common underlying cause. In particular, both phenomena may jointly arise as a consequence of contractual incompleteness in market environments in which work effort is not third-party verifiable. We also provide evidence supporting a specific mechanism for how contractual incompleteness can cause these two phenomena, in which the implicit incentive strategies adopted by firms to address the contract enforcement problem play key role.

Intuitively, when effort is not verifiable, firms may adopt an implicit incentive strategy for eliciting high work effort that involves paying relatively high wages, and conditionally renewing workers' contracts based on their performance. With diminishing marginal product of labor, however, high wage payments can make it profitable for firms to hire fewer workers than technologically feasible, since the gains from higher overall production might be more than offset by higher wage costs. The result of such job rationing is endogenous involuntary unemployment. If a critical mass of firms rations jobs, however, a secondary employment sector could also emerge, where firms profitably fill all vacancies and pay relatively low wages. Such firms are able to pay lower rents, and elicit relatively lower but non-minimal effort, because of the unemployment pressure in the market. This might give rise to a segmented labor market in which the strategies of offering "good" high-rent jobs and "bad" low-rent jobs are—in equilibrium—equally profitable for firms. We show that the qualitative features of this intuition can be captured in a simple formal model, which builds on the classic efficiency wage frameworks by Shapiro and Stiglitz (1984) and Akerlof and Yellen (1990).

In order to provide empirical evidence on the causal impact of contractual incompleteness on unemployment and market segmentation, we study behavior of firms and workers in competitive experimental labor markets. All firms in our markets share the same production technology which exhibits decreasing returns to scale from labor but ensures that full employment is technologically efficient. In our main treatment, the Incomplete Contracts treatment (*IC treatment*), work effort is observable to firms but not verifiable to third parties. Firms may, however, use implicit incentives to elicit non-minimal work effort in this treatment. In a control treatment, the Complete Contracts treatment (*C treatment*), effort is verifiable and contracts are explicitly enforced. If a worker accepts a contract in this treatment, he thus has to provide the contractually stipulated effort level.

Our first main empirical result is that contractual incompleteness causes a strong increase in the level of unemployment. Our data also reveal important differences in how labor markets function in the presence or absence of explicit contract enforcement. In line with the hypothesized mechanisms, we find that firms in the IC treatment use implicit incentives involving conditional contract renewal and paying strictly positive worker rents. At the same time that they pay high wages, however, some firms choose to offer fewer vacancies than possible. Endogenous unemployment in the IC treatment arises as a byproduct of this job rationing decision. Given that employed workers earn substantial rents, unemployment in the IC treatment is involuntary. In the C treatment where effort is explicitly enforced, labor market outcomes differ substantially along all these dimensions. Firms pay much lower wages and reap the major share of production surplus. Employment relations are shorter than in the IC treatment, and the overwhelming majority of firms does not ration jobs. As a result, endogenous unemployment in this treatment is very low and mostly voluntary, being caused by workers who do not accept existing contract offers. We also find support for the underlying mechanism that is hypothesized to drive these treatment differences, namely an impact of incompleteness on the profitability of different contractual instruments: paying positive rents and using contingent contract renewal increases firm profits in the IC treatment, while being counterproductive or irrelevant for firms in the C treatment.

Our second main empirical finding is that contractual incompleteness leads to a stable coexistence of different job types. After an initial phase in which we observe a trend towards job rationing in the IC treatment, a plateau is reached such that unemployment stabilizes at a high level, and a relatively constant fraction of firms continues to operate without rationing job offers. Whereas in the initial phase job rationing is the more profitable strategy for firms, in this later phase firms earn similar profits regardless of whether or not they ration jobs. Workers, however, earn substantially lower rents and exert lower effort in firms that do not ration jobs. In the long run, the situation in the IC treatment thus resembles a segmented labor market in which some workers are employed in "primary-sector" jobs characterized by high worker rents, relatively stable employment relationships, and job rationing, while other workers work under less favorable conditions in "secondary-sector" jobs (see, e.g., Doeringer and Piore 1971, or Saint-Paul 1996). Given the equal profitability of these alternative firm strategies, the segmentation has the character of a stable market equilibrium. By contrast, market segmentation is not observed in the C treatment where firms' strategy of not rationing jobs and paying low worker rents pervades the market. This indicates that the emergence of market segmentation is directly linked to contractual incompleteness. We also find support for a key mechanism hypothesized to sustain segmentation, which is an impact of unemployment pressure on worker behavior: workers in the IC treatment are significantly less likely to shirk when reduced market activity indicates lower job finding chances.

The first important contribution of our paper lies in empirically identifying a direct causal link between contractual incompleteness and involuntary unemployment. While efficiency-wage theory has long hypothesized that this link may exist (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1989, Akerlof and Yellen 1990, MacLeod and Malcomson 1998), establishing this key claim of the theory is difficult if not impossible using field data. The empirical literature has made important contributions on other aspects of the efficiency-wage hypothesis, particularly the relationship between rents and worker performance (for a survey see Katz 1986). However, evidence on key variables like worker effort has necessarily been indirect, because effort is inherently difficult to measure in settings where efficiency wages would be relevant. As a solution, previous field studies have related indirect proxies for effort, such as worker discipline problems or survey measures of workplace performance, to wage premiums (e.g., Cappelli and Chauvin 1991, Campbell III and Kamlani 1997), or to measures of dismissal barriers and firing threat (e.g., Ichino and Riphahn 2005). An experimental approach is complementary to these studies because of the possibility to exogenously vary the degree of contractual incompleteness, and to measure the impact on involuntary unemployment.[1] In our setup we can also induce or accurately measure key variables such as worker effort and ability, or firms' production technology. Thus, we are able to precisely assess whether contractual incompleteness influences worker rents or decisions on job rationing and contract acceptance, as hypothesized in the theory.

The second main contribution of our paper concerns understanding the foundations

[1]For a general discussion of the role of lab experiments in studying labor market institutions see, e.g., Charness and Kuhn (2011).

of dual labor markets. The theoretical literature on dual labor markets has traditionally argued that market segmentation can result from contract enforcement problems and efficiency wages, if monitoring technologies differ exogenously across segments (Bulow and Summers 1986). Some theoretical approaches have made segmentation endogenous, arising due to non-linearities in monitoring technology, differences in setup and adjustment costs, or on-the-job search (Albrecht and Vroman 1992, Saint-Paul 1996, as well as Board and Meyer-ter-Vehn 2011). We provide further theoretical insights on how market segmentation can emerge endogenously, with diminishing marginal product in the production function, despite homogeneous technology across firms. More importantly, we provide the first empirical evidence on the endogenous emergence of market segmentation due to contractual incompleteness. Some earlier experimental papers have observed firms offering different types of jobs which are more and less attractive from workers' perspective (Brown et al. 2004, Bartling et al. 2012). Importantly, however, in these studies jobs that are good and bad for workers also exhibit strong and systematic differences in firm profits. In contrast, we find that firms who use the good-job and bad-job strategy are equally profitable, consistent with market segmentation being supportable as a stable equilibrium outcome.

A number of other papers has used experimental techniques to study the consequences of contractual incompleteness in labor market settings (e.g., Fehr et al. 1993, Brown et al. 2004, Linardi and Camerer 2010). Our paper differs from this literature in that it incorporates an analysis of endogenously arising unemployment. By contrast, unemployment in earlier papers was exogenously given, ruled out by the design of the experiment, or determined by exogenous stochastic shocks. Previous studies have shown that contractual incompleteness can lead to an adoption of implicit incentive strategies that involve rent payments and contingent contract renewal (e.g., Brown et al. 2004).[2] We add a missing dimension to this literature, showing how contractual incompleteness can also affect aggregate-level market outcomes, and how these in turn reinforce workers' incentives. Our findings thus provide a missing empirical link, showing that contractual incompleteness, the use of implicit incentives, unemployment, and market segmentation can all be intimately related.

In the remainder of the paper, we first present the setup and procedures of the experiment, before discussing theoretical hypotheses in section 3.3. In section 3.4, we present our empirical results, and section 3.5 concludes.

[2] For recent surveys on lab and field experiments that study contractual and non-contractual solutions to alleviate moral hazard in the labor market see Charness and Kuhn (2011) and Bandiera et al. (2011). See Brown and Zehnder (2007) and Brown and Serra-Garcia (2012) for applications in credit markets.

3.2. Design and Procedures of the Experiment

To study the impact of contractual incompleteness on unemployment and market segmentation, we implemented experimental labor markets where we exogenously varied the verifiability of work effort. As our workhorse, we used a variant of the gift-exchange game (Fehr et al. 1993). In the market, firms and workers interacted for 18 periods. Each period consisted of a market phase where firms offered employment contracts and hired workers, and a work phase where work effort was determined. In our main treatment, the IC treatment, effort was not verifiable and workers thus could depart from the contractually agreed upon effort level. By contrast, the effort level stipulated in the employment contract was explicitly enforced in our control treatment, the C treatment. Keeping everything else identical (production technology, supply and demand of labor, etc.) while varying the verifiability of work effort allows us to causally identify the effects of contractual incompleteness on labor market outcomes.

3.2.1. The Market Phase

Firms were the contract makers in the market phase. To offer a contract, firms stipulated a non-contingent upfront wage payment, w, and a desired level of effort, $\hat{e}$. Firms could make two types of contract offers: public offers that were available to all workers and could also be observed by all other firms, or private offers that were only available to one specific worker. Public offers allowed firms to reach the entire market if they wanted to fill a vacancy regardless of a particular trading partner. Private offers made it possible for firms to target specific workers. This is a necessary feature if firms want to apply a strategy that involves systematic rehiring of workers based on their previous performance. If an employer wanted to (re)hire a specific worker via a private contract offer, she had to specify the ID of the worker in the contract offer. In this case, only the selected worker was informed about the contract offer, and only this worker could accept the offer.

In a given market period, each firm could hire up to two workers. As long as none of her contract offers had been accepted, a firm could make as many private and public offers as she wanted. A worker could accept any public contract offer available in the market, and any private offer he had received. Workers were not informed about the number of private offers in the market as a whole, but they could infer labor market conditions and the tightness of the market from the number of public offers observed in a given period. Once a worker accepted a contract offer, he was not allowed to accept further offers in this period. Additionally, all other outstanding offers of the respective firm were removed from the list of available contracts. The firm could then decide to open a second vacancy and hire another worker, by entering new contract offers. This feature of opening first and second vacancies sequentially was implemented to prevent "accidental hiring", such that a firm who wanted to employ only one worker but entered multiple contract offers had a second offer accepted before being able to withdraw her remaining contract offers. Note that while firms and workers could endogenously build up long-term employment relationships by repeatedly offering and agreeing on (private) contract offers, it was not possible for market participants to directly announce or sign a multi-period employment contract.

The market phase ended when all firms had filled both vacancies, or when all firms had indicated that they did not want to offer further vacancies.[3] At the end of the market period, workers received a summary of their own contract terms, and information on whether and under which conditions their firm had employed a second worker. This information was mainly provided to ensure common knowledge within a firm on whether the firm operated as a one-worker or two-worker firm. While complete information regarding co-workers' wages may not be fully realistic, some degree of transparency is likely present in many work settings: even with a firm policy encouraging wage secrecy, workers may have a reasonably accurate idea about co-workers' earnings.[4]

3.2.2. The Work Phase

After the end of the market phase, workers who had accepted a contract offer entered the work phase in which actual work effort, e, was determined. Since effort was contractible in the C treatment, effort levels corresponding to the contractually stipulated ones were exogenously implemented by the experimenter ($e = \hat{e}$). By contrast, work effort was observable by the firm, but not verifiable to third parties in the IC treatment. Therefore, a worker could exert equal, less, or more effort than stipulated in his employment contract. Workers' effort choices, together with firms' wage payments, determined material payoffs of firms and workers. Before the next period started, a firm and its worker(s) were informed about work efforts and the resulting payoffs for the firm and the workers employed by this firm.

[3] We also had a maximum trading time of 200 seconds for each market phase. This constraint was, however, only binding in few occasions (mostly in the C treatment). The impact of the time limit on unemployment and other market outcomes reported below is thus limited and confined to the control treatment.

[4] Empirically, co-worker wages have no significant impact on workers' effort choices in our setup. Efforts in the IC treatment strongly depend on a worker's own contract terms, but they are not significantly related to either the wage or the desired effort level of the co-worker (results can be found in Table A.2 of the appendix). The finding that wage inequalities per se might not affect behavior is in line with recent evidence on social comparison processes in similar setups (Charness and Kuhn 2007, Gächter et al. 2012).

3.2.3. Parameters and Procedures

Participants' roles were randomly assigned at the beginning of the experiment and kept constant throughout all market periods. In every market, we had 17 workers and 7 firms. Since firms could employ at most two workers, this implies that three workers were "exogenously" unemployed in each period.

A worker's payoff in a given period, π_W, was given by

$$\pi_W = \begin{cases} w - c(e) & \text{if worker accepted a contract } [w, \hat{e}] \\ 0 & \text{if unemployed} \end{cases}$$

A worker who remained unemployed in a given period received a payoff of 0 points. An employed worker received the wage w specified in his contract and had to bear the cost of the work effort he provided, $c(e)$. The set of feasible efforts and wages was given by $e \in \{1, 2, ..., 10\}$ and $w \in \{0, 1, 2, ..., 100\}$. As illustrated in Table 3.1, we induced a convex (monetary) effort-cost schedule in the experiment.

Effort level e	1	2	3	4	5	6	7	8	9	10
Cost of effort $c(e)$	0	1	2	4	6	8	10	12	15	18

Table 3.1.: *Schedule of effort costs.*

A firm's profit depended on the number of workers hired, the wage(s) paid, and the effort exerted by the worker(s). Firms' production technology was characterized by decreasing returns to scale. Decreasing returns are often argued to arise with increases in firm size, for instance due to higher bureaucratic or coordination costs in larger organizations. We conceptualized this as a reduction in workers' productivity if a firm hired two workers. Specifically, each unit of effort by a worker increased output (and the firm's payoff) by 10 points if only one worker was employed by the firm. If two workers were employed, each unit of effort increased the firm's payoff by 7 points. This corresponds, for instance, to a work environment where workers in larger firms need to spend some of their time doing administrative tasks that are not directly productive. The payoff of a firm, π_F, can thus be summarized as follows:

$$\pi_F = \begin{cases} 10e_1 - w_1 & \text{if one worker employed} \\ 7(e_1 + e_2) - w_1 - w_2 & \text{if two workers employed} \\ 0 & \text{else} \end{cases}$$

e_1 (e_2) denotes the effort provided by the first (second) worker, and w_1 (w_2) is the wage paid to the first (second) worker employed by the firm. Note that this specification of the production technology implies that efficiency is maximized when two workers are employed and maximum effort is exerted. Payoff functions π_F and π_W, workers' cost schedule $c(e)$, and the number of firms and workers in the market were common knowledge.

The experiment was carried out in the BonnEconLab at the University of Bonn. A total of 240 subjects, mainly university students from all majors, took part in the experiments. Every subject participated only in one session, and we conducted five independent market sessions for each treatment. At the beginning of a session, participants received detailed information about the rules of the experiment.[5] The experiment started only after all participants had answered several control questions correctly. In addition, subjects played one trial period of the market phase to ensure that they understood how to use the computer program. Sessions lasted about 110 minutes and subjects earned on average 25.49 euros (about 35 USD at the time of the experiment), including a showup fee of 8 euros. The experiments were computerized using the software "z-Tree" (Fischbacher 2007); subjects were recruited with the online recruitment system by Greiner (2003).

3.3. Behavioral Predictions

The treatments described in the previous section allow us to identify the causal impact of contractual incompleteness on unemployment and labor market segmentation in our setting. Furthermore, using the data from the experiment we can investigate whether both phenomena arise in a way that is consistent with a specific theoretical equilibrium. A simple model, which is derived in the theoretical appendix, informs our hypotheses. In the model, we show how unemployment and segmentation can be part of a market equilibrium when effort is non-verifiable, and how these aggregate-level outcomes arise along with a very specific set of strategic behaviors by workers and firms. These individual-level mechanisms become additional qualitative predictions, which should be satisfied empirically if unemployment and segmentation are to be explainable by the type of equilibrium formalized in the model.

Our model builds on two important strands of efficiency-wage theory. One is the "shirking version" of Shapiro and Stiglitz (1984), in which materially selfish agents are motivated to work by the prospect of earning future rents, and a threat of being fired in case of shirking. The other is the "gift-exchange version" of Akerlof and Yellen (1990), in which workers are fair-minded, in the sense of experiencing a psychological benefit or cost of fulfilling a contract, depending on the generosity of the rents offered in their current contract. Our model incorporates both motivations: a fraction of agents is assumed to have fairness

[5] A translation of the instructions can be found in section A.2.3 of the appendix. To rule out that differences in participants' experiences from their employment relationships outside the lab could bias our results, instructions were framed in a neutral goods-market language.

concerns, while the rest is materially selfish. This assumption is in line with abundant previous evidence from the lab and field, showing that a mix of selfish and fair types is typically present in a given population (see, e.g., Fehr et al. 1993, Bewley 1999, Fehr and Gächter 2000, Cohn et al. 2012). To match the experimental setting, our model features a finite horizon.[6]

3.3.1. Non-verifiable effort

The presence of both fair and selfish types has important implications for the type of market equilibria that can emerge in finite-horizon settings where effort is non-verifiable, as is the case in our IC treatment. Intuitively, the presence of some fair types may lead firms to pay non-minimal wages in the final period, anticipating that fair agents voluntarily provide non-minimal effort in response to fair wages. This generates a rent from being employed in the final period. The prospect of earning this final-period rent, in turn, opens up possibilities for equilibria in which firms use implicit incentives to motivate agents in the pre-final period(s). There can thus exist equilibria in the IC treatment that involve firms paying rents and conditioning contract renewal on workers' previous performance; in pre-final periods of such equilibria, selfish and fair types pool and fulfill their contracts in order to qualify for reemployment and earning future rents.

Our model shows how—in an environment where effort is non-verifiable—there exist equilibria in which involuntary unemployment can arise as a byproduct of the implicit-incentive strategies adopted by firms. Intuitively, starting from a situation where unemployment is low, the rents needed to deter shirking may be rather high because shirkers who are fired can relatively easily find another job. If a firm adopts a strategy of paying high wages and worker rents, however, decreasing returns to scale can make it profitable to "stay small". Increasing the number of workers at a given wage yields a proportional increase in wage costs, but a lower increase in output. If wages are high, the absolute increase in the wage bill can be larger than the gains in terms of extra output, so that it can be profitable to hire fewer workers than technologically feasible. Endogenous unemployment arises as a byproduct of such job rationing, and since firms pay positive rents, this unemployment is involuntary.[7] The model identifies several key components that should

[6]For a more abstract model that is tied less directly to the specifics of the experimental design, see the discussion paper version of our paper, in which we show in an infinite-horizon setting how unemployment and segmentation can be part of a stationary labor market equilibrium under contractual incompleteness.

[7]Anticipating an impact of higher unemployment pressure on worker effort potentially provides an additional motive for firms to ration jobs. This is likely a second-order concern, however, since an individual firm has only a modest impact on the aggregate unemployment rate in our experiment. In the limit, as the number of market participants increases this motive vanishes entirely.

be observed in the IC treatment if such an equilibrium arises. First, the profit-maximizing strategy of firms should involve paying strictly positive worker rents, and conditional contract renewal such that workers who shirk have a higher likelihood of being dismissed than workers who provide the contractually stipulated effort level. Second, from a worker perspective, the long-run costs of shirking in terms of forgone future rents should more than offset the short-run gains due to lower effort costs. Third, unemployment should emerge due to firms deciding to offer less vacancies than possible, rather than workers rejecting available job offers.

The model also shows how an equilibrium with non-verifiable effort can involve market segmentation, emerging jointly with involuntary unemployment. Intuitively, as some firms adopt the strategy of eliciting effort with high rents ("good jobs") and rationing jobs, the resulting unemployment reduces the job-finding chances for unemployed workers, and therefore the attractiveness of shirking for those employed. If the unemployment pressure in the market is strong enough, a fraction of firms might be able to operate equally profitably by hiring two workers, and offering jobs involving lower worker rents, and somewhat lower effort levels ("bad jobs"). Firms with this low-wage, low-effort strategy can have a larger optimal size since doubling a low wage leads to a small absolute cost increase, and this can be smaller than the gain in extra output even with decreasing returns to scale.[8] Importantly, without sufficient unemployment pressure, the low rents of secondary-sector firms might not be able to prevent shirking, and offering bad jobs would thus not be profitable for firms. In the appendix, we derive sufficient conditions for segmentation to be part of a stable market equilibrium, such that the strategies of offering good jobs and bad jobs are equally profitable for firms, and neither type of firm has an incentive to deviate. If a segmentation equilibrium arises in our experiment, the model thus predicts that one-worker firms and two-worker should coexist in the market, and that firms in both segments are equally profitable. Furthermore, the firms who ration jobs should be the ones offering high wages and worker rents, and workers in such firms should exert high effort levels. In firms who hire two workers, wages and offered worker rents are predicted to be lower, as are worker effort levels.

3.3.2. Explicit contract enforcement

The mechanisms we have described as potential sources of unemployment and market segmentation are directly linked to the inability to verify work effort; they thus do not

[8]The tendency for larger firms to implement a lower effort level than smaller firms follows directly from profit maximization and DRS: As we show in appendix A.2.1, the optimal effort is always lower for a larger firm, since the impact of effort on output is smaller at a flatter portion of the production function.

apply in our control treatment where effort is contractible. In the appendix we discuss equilibrium characteristics for an environment with explicit contract enforcement, as is the case in the C treatment. Holding constant other assumptions besides the verifiability of work effort, the equilibrium is characterized by all firms hiring two workers and thus zero endogenous unemployment. Firms elicit maximum effort from workers, while paying wages just slightly above worker effort costs. Because effort is explicitly enforced, there is also no need to engage in repeated contract renewal with specific workers. Since all firms use a homogenous strategy of offering "bad jobs", there is no market segmentation.

Intuitively, firms can elicit high effort without paying high rents or using conditional contract renewal because contractibility of effort eliminates a source of worker bargaining power: not only are firms on the short side of the market, but workers have no option to shirk once they are employed. The presence of fair types has relatively minor implications for behavior in an environment with explicit contract enforcement, as workers cannot shirk on the job.[9] Firms strictly prefer hiring two workers, due to the efficiency of maximum employment for our production technology. We thus hypothesize that firms in the C treatment maximize profits by paying very low rents, and filling both vacancies. Furthermore, contingent rehiring of specific workers should be irrelevant for firm profits, and firms are thus predicted to more heavily rely on public offers in the C treatment. When a firm wants to fill a vacancy without concern of hiring a particular worker, public offers have the advantage that they reach the entire market.

Our discussion has illustrated how involuntary unemployment and market segmentation can be part of a market equilibrium when effort is not verifiable. It is well known, however, that there are typically multiple possible equilibria in repeated games (see, e.g., Fudenberg and Maskin 1986), and this is true as well in our setup. For example, there is a range of different equilibria for our IC setting, involving different equilibrium separation rates between firms and non-shirking workers. Our model illustrates how equilibrium characteristics depend on the equilibrium separation rate on which the players endogenously "coordinate". For example, according to our model, segmentation equilibria in environments with excess supply of labor always involve positive equilibrium separation

[9]Fairness could play a role in the decision of whether or not to reject a contract. Theories of fairness, and ample empirical evidence, however, indicate that a key motive underlying fairness is a desire to punish unfair actions by an opponent. In a setting like the C treatment, where receivers of offers face competition, rejection of contract offers has limited effectiveness as a punishment strategy and even offers involving very low worker rents are thus likely to be accepted (see, e.g., Fehr and Schmidt 1999).

rates, whereas market segmentation is not possible with a separation rate of zero.[10] While equilibrium multiplicity limits the predictive power of repeated game models, in our view they are nevertheless very useful for heuristic purposes. The specification of "candidate" equilibria gives rise to additional testable hypotheses for behavior, which must be satisfied if the data are to be explained by a particular type of equilibrium. For instance, equal profitability of both firm types and positive separation rates between firms and non-shirking workers are both necessary features, and thus testable qualitative implications of an equilibrium involving market segmentation. The presence of multiple equilibria also makes it even more important to study actual behavior in a narrowly controlled environment, as it is ultimately an empirical question which type of equilibrium emerges.

To summarize, the equilibria described above for the IC and C settings suggest the following qualitative predictions : (1) endogenous involuntary unemployment in the IC treatment, in contrast to maximum employment in the C treatment; (2) firms in the IC treatment using strictly positive rent payments, and conditional contract renewal with private offers, as a strategy to provide implicit incentives, in contrast to firms in the C treatment using minimal rents and more strongly relying on public contract offers; (3) segmentation of the market in the IC treatment, such that firms earn similar profits but differ qualitatively with respect to worker rents, effort, and job rationing, in contrast to no segmentation in the C treatment with all jobs involving minimal rents; (4) the secondary-sector strategy in the IC treatment being supported by unfavorable market conditions, which make workers willing to provide non-minimal effort for low rents; (5) a positive separation rate for firms and non-shirking workers in the IC treatment. The discussion of our empirical results in the following section will be structured according to these main predictions.

3.4. Results

In this section we first analyze the impact of contractual incompleteness on the level of unemployment, before examining in more detail whether the channels through which unemployment emerges are in line with the mechanisms featured above (section 3.4.1). A similar structure applies to section 3.4.2, where we study the influence of contractual incompleteness on labor market segmentation.

[10] As discussed in the theoretical appendix, when there are zero job separations in a no-shirking equilibrium, there are no labor-market flows and the value of unemployment is zero. In this case the value of unemployment does not depend on the number of one-worker firms, eliminating a key mechanism for sustaining segmentation.

3.4.1. Contract Enforcement and Unemployment

Figure 3.1 depicts the average unemployment rates for the C treatment (black) and the IC treatment (grey), as well as the fraction of workers who are exogenously unemployed due to excess supply of labor (dashed grey line). We observe strong differences in unemployment between treatments ($p < 0.01$).[11] In the C treatment where contracts are explicitly enforced, unemployment remains close to the minimum possible level: in most periods, the unemployment level lies only 1–5 percentage points above the baseline level that is due to excess supply of labor. The unemployment pattern looks markedly different, by contrast, when effort is not verifiable. In particular, we observe a sharp increase in unemployment over the first seven market periods, before unemployment stabilizes and remains high for the rest of the game.[12] Overall, contractual incompleteness thus causes a strong increase in the level of unemployment.

Result 1: *We observe strong differences in unemployment between treatment conditions. Under explicit contract enforcement (C treatment), unemployment levels are close to the minimal possible level. When effort is not verifiable (IC treatment), unemployment rises strongly before stabilizing at a relatively high level.*

Rows (1) and (2) of Table 3.2 shed further light on the driving forces behind the differences in unemployment. As could be inferred from the low levels of endogenous unemployment in the C treatment, we find that firms in this treatment almost always try to fill two vacancies (in 96.7% of cases) and, with an acceptance rate of 97.8%, workers essentially always accept. In the IC treatment, by contrast, firms open up only 67.5% of the possible vacancies. At the same time, available contract offers are accepted in 99.8%

[11]This test is based on a panel estimation with session-level random effects. Unless otherwise noted, the same procedure is also applied for all tests on other aggregate-level outcomes in the remainder of this section. Test statistics for individual-level data (e.g., wages in Table 3.2 below) are based on estimations with individual-level random effects and standard errors accounting for potential clustering at the individual level. Results are robust to applying non-parametric tests based on session-level averages for analyzing the main treatment differences, or to using alternative estimation strategies, such as allowing for multi-level random effects at the session and individual level, or clustering of standard errors at the session level. Reported p-values are always two-sided.

[12]Regressing per-period changes in unemployment on a constant, we estimate an average change of $\Delta_t = 3.6$ percentage points for the first seven period in the IC treatment ($p = 0.044$). If we do the analog for the remaining periods, the average change is much smaller and insignificant ($\Delta_t = 0.6$; $p = 0.517$). Not surprisingly, the same holds in an estimation framework that uses linear time trends and accounts for a potential "endgame effect" in unemployment in the final period of the IC treatment. We find a significantly positive trend for the first seven periods ($p < 0.01$), but not for the remaining periods or the final period alone ($p = 0.154$ and $p = 0.228$, respectively).

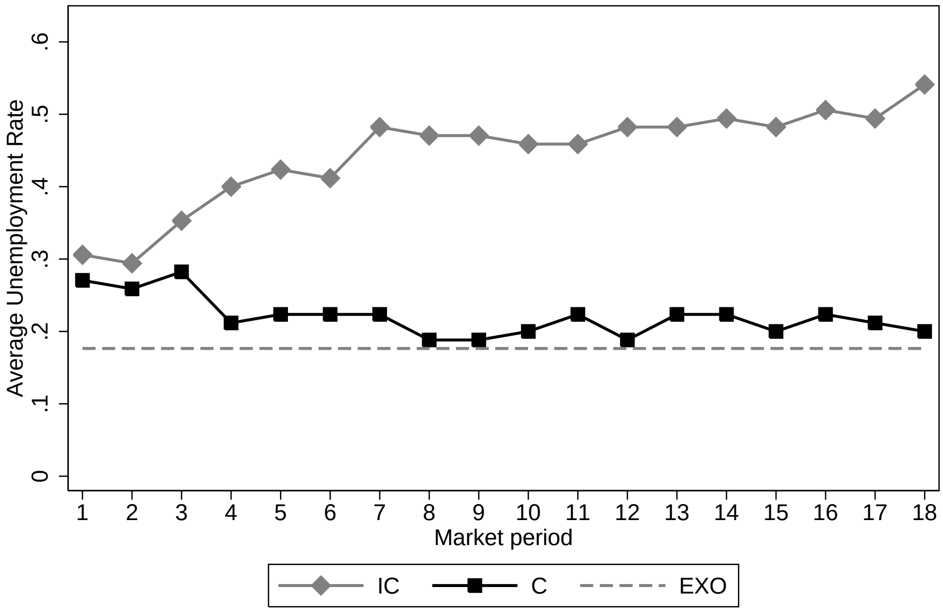


Figure 3.1.: *Unemployment over time. Average unemployment rate in the IC treatment (grey) and the C treatment (black). Fraction of exogenously unemployed workers (dashed grey).*

of cases.[13] Firms' decision to offer fewer vacancies than possible—i.e., to ration jobs—is thus the main source of endogenous unemployment in the IC treatment.

Rows (3)–(9) of Table 3.2 summarize further differences in the characteristics of employment relationships, in terms of the contract terms being offered, worker behavior, and the duration of employment relationships. In the C treatment, firms pay worker rents of only 5.7 points, hire workers via public offers in 84% of cases, and elicit effort close to the maximum possible level (9.65). In contrast, wages and offered worker rents are substantially higher in the IC treatment, reaching levels of 34.6 and 21.1 points, respectively. In addition, 71.7% of concluded contracts in this treatment are initiated via private offers, and long-term employment relationships are frequently observed. For instance, 60% of firms rehire the same worker for at least 9 consecutive market periods. Effort levels are lower than in the C treatment, but at an average level of 7, they lie far above the minimum. Although effort is not verifiable, workers in the IC treatment on average deviate from the contractually stipulated level by only 1.2 points. These findings provide first indications that firms in the IC treatment successfully use contract renewal and worker rents to establish implicit performance incentives. The co-existence of high

[13]In the few cases in which we observe available contracts offers being rejected, these exhibit very low or even negative worker rents. On average, the rejected contract offers stipulate worker rents of only 0.9 points (1.9 points in the C treatment, and -6.6 points in the IC treatment). Rejection of such offers with extremely low worker rents might be an indication of workers' fairness concerns (see the discussion in the theoretical appendix).

		C treatment	IC treatment	
(1)	Fraction of possible vacancies offered	0.967	0.675	$p < 0.01$
(2)	Fraction of offered vacancies accepted	0.979	0.998	$p < 0.01$
(3)	Wages	22.80	34.60	$p < 0.01$
(4)	Rents offered by firms ($w - c(\hat{e})$)	5.74	21.12	$p < 0.01$
(5)	Realized worker rents ($w - c(e)$)	5.74	23.59	$p < 0.01$
(6)	Fraction of private contracts	0.160	0.717	$p < 0.01$
(7)	Fraction of firms with employment relationships $\geq$ 9 market periods	0.029	0.600	$p < 0.01$
(8)	Effort	9.65	6.97	$p < 0.01$
(9)	Desired effort	9.65	8.20	$p < 0.01$

Table 3.2.: *Market characteristics. Mean values across treatments. The p-value for row (7) is derived from a linear-probability model where we regress a dummy variable equal to 1 if a given firm has at least one employment relationship for $\geq$ 9 consecutive market periods during the experiment on a treatment dummy. Reported p-values for all other tests are derived using the procedures described in Footnote 11.*

worker rents and job rationing through firms also demonstrates that unemployment in the IC treatment is involuntary. At the same time, the observation of low rents and a lack of repeated contract renewal in the C treatment indicate that firms manage to exploit their high bargaining power when contracts are explicitly enforced.

In Table 3.3 we investigate two additional hypotheses about differences in the functioning of employment relationships, which cannot be directly inferred from the level differences in Table 3.2. Columns (1) and (2) report linear-probability estimates on the determinants of firms' contract renewal decisions. This allows us to examine whether the greater prevalence of long-term relationships in the IC treatment indeed reflects a policy of *conditional* contract renewal, with rehiring and separation decisions depending on a worker's previous performance. Comparing rehiring rates across treatments, the positive coefficient for the IC treatment in Column (1) confirms that contract renewal is overall more prevalent when effort is not verifiable. The negative coefficient on the indicator for shirking, however, demonstrates that firms in the IC treatment do indeed strongly condition contract renewal decisions on whether a worker fulfilled the contract. The average likelihood that a contract is renewed drops from 70.9% in the case of contract fulfillment to only 26.8% if a worker deviates from the contractually stipulated effort level. The estimates also imply that separation rates in the case of contract fulfillment are lower, but strictly positive in the IC treatment. This is important since—according to our model—equilibria involving endogenous unemployment and market segmentation exhibit non-zero separation rates in equilibrium. Column (2) adds controls for other important characteristics of employment relations, including the rents offered in the current period's contract,

	Dependent variable:			
	1 if worker is re-hired in $t+1$		Future rents	
	(1)	(2)	(3)	(4)
IC treatment	0.639***	0.232***	153.899***	211.060***
	(0.061)	(0.072)	(14.896)	(21.012)
Shirking (t)	-0.427***	-0.205***	-74.028***	-70.156***
	(0.057)	(0.058)	(14.081)	(10.902)
Offered rents (t)		0.003		0.768*
		(0.002)		(0.425)
IC × Offered rents (t)		0.002		0.483
		(0.003)		(0.635)
Contract renewed (t)		0.416***		4.717
		(0.083)		(4.769)
IC × Contract renewed (t)		0.045		50.617***
		(0.098)		(11.767)
Constant	0.065***	-0.007	28.480***	56.495***
	(0.020)	(0.022)	(1.949)	(5.974)
Market period	no	yes	no	yes
N	1,935	1,935	2,042	2,042

Table 3.3.: *Rehiring and long-run incentives. Random-effects models; the reported standard errors (in parentheses) account for potential clustering at the individual level. Columns (1)–(2): the dependent variable equals 1 if a firm renews the contract of a worker through a private contract in period $t+1$. Columns (3)–(4): the dependent variable "future rents" is the sum of a worker's earnings from period $t+1$ until period 18. "Shirking" is an indicator equal to 1 if the worker deviates from the contractually stipulated effort level in the current period (i.e., $e_t < \hat{e}_t$). "Contract renewed" is an indicator equal to 1 if a firm has interacted with the same worker in period $t-1$, and rehired the worker through a private contract offer in period t. "IC × Offered rents" and "IC × Contract renewed" are interaction terms of the respective measure with the treatment dummy. Columns (2) and (4) additionally control for a linear time trend, as well as an interaction term of the time trend with the treatment dummy. *** / ** / * indicate significance on the 1-percent / 5-percent / 10-percent level.*

and an indicator for past contract renewal, which equals 1 if a firm has already renewed a worker's employment contract in the current period (via a private contract offer). In this specification shirking continues to be a crucial factor in firms' contract renewal decisions in the IC treatment.[14] This indicates that firms in the IC treatment systematically engage in conditional contract renewal rather than, e.g., just having a stronger taste for repeatedly interacting with a given worker. Further estimates (available upon request)

[14] By design, shirking is not possible in the C treatment, and thus we do not include an interaction term between shirking and the treatment dummy. Interestingly, the coefficient of "Contract renewed" in the C treatment is significant and positive (Column 2). This only reflects a small number of firms, however, who engage intensely in contract renewal in this treatment; as illustrated in Column (1), the likelihood that a firm in the C treatment rehires her worker is very low. Our estimates from Table 3.4, discussed below, further indicate that this strategy does not increase the respective firms' profits.

show that the results reported in Table 3.3 are also robust to using alternative measures of worker slacking, such as the degree of deviation from the contractually stipulated effort level.

Columns (3) and (4) of Table 3.3 show that the observed differences in rent payments and contract renewal strategies have important consequences for workers' long-term earnings prospects. Confirming the observations from Table 3.2, the treatment dummy in Column (3) illustrates that workers in the IC treatment generally earn higher future rents than those in the C treatment. More importantly, the coefficient for shirking indicates that—within the IC treatment—the long-run benefits of contract fulfillment in a given period considerably outweigh the short-run gains from shirking.[15] Controlling for other aspects of contract terms yields similar results (see Column (4) of Table 3.3). This underlines that workers in the IC treatment face strong implicit performance incentives. The relatively high effort levels in this treatment are thus understandable from a long-term incentive perspective.

The data from the IC treatment also illustrate the relevance of both (implicit) material incentives and fairness concerns. This can best be seen in the final market period, in which the potential for future interactions and the prospect of qualifying for future rents vanish. Nevertheless, we find that 46.1% of workers in the IC treatment provide above-minimal efforts in this periods, on average choosing an effort level of 6.0. This underlines the relevance of voluntary gift-exchange in our setting, paralleling what has commonly been observed in similar labor-market settings in the lab and field (see, e.g., Fehr et al. 1993, Brown et al. 2004, Cohn et al. 2012). At the same time, our data on effort provision also illustrate the importance of future rents and contingent contract renewal for motivating workers in the IC treatment: those workers who only exert minimal effort in the final period on average provide efforts of 7.7 in period 17 in which the potential for future interaction is still intact.

Result 2: *Firms in the IC treatment pay higher workers rents, engage more heavily in repeated contract renewal, and offer fewer vacancies than firms in the C treatment. The combination of job rationing and high worker rents implies that endogenous unemployment in the IC treatment is involuntary. The strategies of firms in the IC treatment establish implicit performance incentives that imply a cost of shirking.*

So far, our findings on treatment differences in worker rents, the prevalence of conditional contract renewal, and other key characteristics of employment relationships are all

[15]For instance, a worker who shirks in period 10 earns on average 96.5 points less during periods 11-18 compared to a worker who provides the contractually stipulated effort in the same period. These losses in future rents are higher than the maximally possible short-run gains of shirking due to reduced effort costs.

	Dependent variable: firm profit from a given contract			
	(1)	(2)	(3)	(4)
Offered rents	-0.990***		-0.991***	-0.869***
	(0.121)		(0.116)	(0.112)
(Offered rents)^2	-0.002		-0.002	-0.004
	(0.003)		(0.003)	(0.002)
IC × Offered rents	2.229***		1.894***	1.529***
	(0.203)		(0.186)	(0.193)
IC × (Offered rents)^2	-0.014***		-0.012***	-0.006
	(0.005)		(0.004)	(0.004)
Contract renewed		-1.596	1.432	0.865
		(3.909)	(1.437)	(1.297)
IC × Contract renewed		17.948***	11.458***	9.802***
		(4.309)	(2.386)	(2.268)
IC treatment	-42.961***	-28.971***	-43.105***	-41.695***
	(1.991)	(1.441)	(1.703)	(2.701)
Constant	52.160***	46.405***	52.087***	49.600***
	(0.520)	(0.755)	(0.552)	(1.659)
Market period	no	no	no	yes
Final period	no	no	no	yes
N	2,042	2,042	2,042	2,042

Table 3.4.: *Profitability of contractual instruments. Random-effects models; the reported standard errors (in parentheses) account for potential clustering at the firm level. The dependent variable is the level of a firm's earnings from a given employment contract. See Table 3.4 for definitions of the remaining variables. Column (4) additionally controls for a linear time trend, an indicator equal to* 1 *in period 18 to capture a possible endgame effect, as well as interaction terms of the respective variables with the IC treatment.* *** / ** / * *indicate significance on the 1-percent / 5-percent / 10-percent level.*

consistent with our theoretical hypotheses. However, our model also suggests a specific mechanism for *why* these differences emerge, namely that the absence of explicit contract enforcement makes specific contractual instruments profitable for firms. Table 3.4 provides a test of this key underlying mechanism. As explanatory variables, we focus on the types of contractual instruments studied in Table 3.2, relating to worker rent levels and contingent contract renewal. Column (1) shows that higher rent payments decrease the profitability of a contract for firms in the C treatment. In the IC treatment, by contrast, we see a qualitatively opposite relationship. Firm profits are increasing in the level of rents that a firm offers to her worker; this holds up to a point, after which further increases in wage costs dominate the profit increases due to higher work efforts. Column (2) reports estimates on the profitability of repeated contract renewal in both treatments, using the "renewed contract" indicator from Table 3.3. We find that firm profits in the IC treatment are strictly higher in employment relationships that involve repeated rehiring. In contrast, firm profits in the C treatment are not systematically affected by whether firms and workers interact repeatedly. In further specifications, we allow for a simultaneous influence of both contractual instruments, and add controls for potential differences in dynamics across treatments (Columns (3) and (4) of Table 3.4). This does not affect our main findings. In sum, we find important qualitative differences in how contractual instruments affect firm profits, and we see that the level differences in the variables shown in Table 3.2 are in line with these profit incentives. We discuss the profitability of differences in job rationing in the next section, where we study the impact of contractual incompleteness on market segmentation.

Result 3: *Contractual incompleteness leads to a qualitative change in the profitability of contractual instruments. Paying higher rents decreases profits, and repeated contract renewal is irrelevant for firms when contracts are explicitly enforceable. Paying higher rents and using repeated rehiring has a positive impact on firm profits when work effort is not verifiable.*

3.4.2. Contractual Incompleteness and Labor Market Segmentation

In this section we turn to investigating the impact of contractual incompleteness on labor market segmentation, and we analyze whether individual behavior is consistent with our theoretical hypotheses for how segmentation can be an equilibrium phenomenon. We have already seen in Figure 3.1 that, after a strong initial increase the level of unemployment in the IC treatment stabilizes during the later phase of the experiment. Given that endogenous unemployment is almost exclusively driven by job rationing, this also implies that the proportion of firms who ration jobs is increasing initially, but then reaches a plateau. Indeed, a relatively stable fraction of about 25% of firms continues to employ two workers during the later phase of the IC treatment. This provides a first indication that there might be market segmentation under contractual incompleteness, with two different firm types coexisting in the long run.

For market segmentation to be an equilibrium phenomenon according to our model, however, it is necessary that the firm strategies characterizing the different market segments are equally profitable. This can occur if the unemployment pressure arising due to job rationing by some firms is sufficient to allow the non-rationing strategy to be viable for other firms. This is exactly what we observe in our data. Job rationing is strictly profitable for firms in the early periods of the IC treatment, in which unemployment pressure is relatively low, and in which we observe strong dynamics towards adopting the one-worker strategy. On average, one-worker firms earn roughly 50% more than two-worker firms during the first seven market periods (average per-period profits are 37.2 and 24.7 points, respectively). The difference in firm profits is statistically significant ($p < 0.01$).[16]

[16]In line with the idea that firms ration jobs in order to pay high worker rents and elicit high work effort,

In contrast, in the remaining periods where unemployment has reached a plateau and the fraction of two-worker firms stabilizes, profits between one-worker firms and two-worker firms do not differ significantly anymore. The average per-period profits for periods 8–18 are 36.7 and 35.3 points in one-worker and two-worker firms, respectively ($p = 0.361$). The finding that, in the long run, profits of one-worker firms and two-worker firms are similar is consistent with the emergence of a segmentation equilibrium.

		C treatment			IC treatment		
		1-worker		2-worker	1-worker		2-worker
(1)	Firm Profits	77.12	$p < .01$	95.46	36.74	$p = .361$	35.32
(2)	Wages	22.88	$p = .115$	21.40	42.33	$p < .01$	27.70
(3)	Rents offered by firms	4.88	$p = .313$	3.74	26.90	$p < .01$	16.00
(4)	Realized worker rents	4.88	$p = .313$	3.74	29.21	$p < .01$	17.55
(5)	Effort	10.00	$p = .161$	9.88	7.91	$p = .718$	6.48
(6)	Fraction of private contracts	0.192	$p = .082$	0.170	0.900	$p = .012$	0.775
(7)	Employment duration	1.00	*NA*	1.87	6.76	$p < .01$	4.48
(8)	Fraction of firms	0.068		0.932	0.738		0.262

Table 3.5.: *Market segmentation in the IC and C treatment. Mean values of market characteristics during the late phase of the experiment (periods 8–18). The reported p-values are derived from random-effects estimations in which the respective dependent variable is regressed on a dummy equal to 1 if a contract comes from a one-worker firm (standard errors account for clustering on the individual level).*

While both strategies yield similar profits for firms during the later market phase, workers in two-worker firms face much less favorable contract terms than their counterparts in one-worker firms (see the rightmost columns of Table 3.5). On average, firms who employ two workers offer about 40% lower worker rents than one-worker firms. At the same time, workers in one-worker firms work somewhat harder: average effort is approximately 20% higher, although this difference turns out to be statistically insignificant. Overall, however, the substantially higher wages in one-worker firms result in much higher realized earnings for workers in those firms.[17] Thus, firms using the one-worker strategy and two-worker strategy in the IC treatment offer "good" jobs and "bad" jobs, respectively, consistent with the endogenous emergence of a segmented labor market. Interestingly,

we observe that 67.4% of firms who switch to a one-worker strategy in the IC treatment subsequently increase their wage payments, and that this increases the profits of the respective firms.

[17] Notably, we find a negative correlation between wages and firm size in the IC treatment. It is important to bear in mind, however, that our design rules out some of the most important factors that are typically discussed as reasons for positive firm size-wage differentials, e.g., larger firms hiring higher-quality workers in terms of observed and unobserved skills, having higher degrees of unionization, or facing stronger monitoring difficulties (e.g., Brown and Medoff 1989, Abowd et al. 1999).

the two segments in the IC treatment also seem to differ in the general stability of employment relationships: two-worker firms are somewhat more likely to hire their workers through public contract offers, and the overall employment duration in two-worker firms is also shorter (see line (7) and (8) of Table 3.5).

In contrast to the situation in the IC treatment, we do not see any indication of market segmentation under explicit contract enforcement (see the leftmost part of Table 3.5). In the few occasions in which firms do hire only one worker in the C treatment, wages and worker rents are only slightly above the values for two-worker firms (the differences are 1.5 and 1.2 points, respectively). Similarly, efforts are about 1% higher, and the fraction of private contracts is 2.2 percentage points higher in one-worker firms. While the latter effect turns out to be weakly significant, the differences between segments are generally much smaller than in the IC treatment. Most importantly, the data on firm profits demonstrate that employing a one-worker strategy in the C treatment is clearly suboptimal from a firm's perspective. Firms who hire only one worker in the C treatment earn almost 20 points less than the ones employing two workers. This underlines why the latter strategy dominates the market, with 93% of firms using a two-worker strategy when contracts are explicitly enforced.

Result 4: *The absence of explicit contract enforcement leads to a segmentation of the labor market. In the long run, two types of firms coexist in the market when effort is not verifiable. These earn similar profits, but differ qualitatively with respect to wage payments, worker rents and effort provision.*

As a final step we investigate more directly a key mechanism underlying our theoretical explanation for segmentation, which is a feedback from unemployment pressure to behavior of workers. The emergence of the secondary sector is possible in equilibrium, because the unemployment pressure arising as a byproduct of job rationing makes workers less likely to shirk and willing to put in higher effort for a given wage. If this is true, we should observe workers being more likely to shirk on the job when they receive information that signals low unemployment pressure and high chances to acquire a job. While workers in the experiment did not have precise information on the level of unemployment or the job acquisition rate in a given period, they could infer the tightness of the labor market from activity in the contracting stage. The most salient indicator of less favorable market conditions from a worker's perspective is the number of public contract offers in a given period. Since unemployed workers disproportionally have to rely on public contract offers for finding a new job, a low number of public offers indicates high costs of unemployment and, consequently, higher unemployment pressure for those employed.

Columns (1)–(3) of Table 3.6 demonstrate that a decrease in the number of public job

	Dependent variable:					
	1 if $e < \hat{e}$			Effort		
	(1)	(2)	(3)	(4)	(5)	(6)
# public offers	0.022***	0.020***	0.020**	-0.069**	-0.058**	-0.070**
	(0.006)	(0.006)	(0.008)	(0.028)	(0.025)	(0.033)
Wage	-0.019***	-0.019***	-0.019***	0.112***	0.110***	0.110***
	(0.001)	(0.002)	(0.002)	(0.007)	(0.008)	(0.008)
Desired effort	0.111***	0.111***	0.111***	0.223***	0.223***	0.222***
	(0.009)	(0.009)	(0.009)	(0.042)	(0.042)	(0.043)
Contract renewed		-0.044	-0.045		0.248	0.274
		(0.048)	(0.048)		(0.201)	(0.203)
Market period			-0.000			-0.008
			(0.005)			(0.023)
Final period	0.379***	0.382***	0.385***	-3.147***	-3.162***	-3.085***
	(0.073)	(0.073)	(0.087)	(0.490)	(0.486)	(0.540)
Constant	0.120*	0.130**	0.135	1.403***	1.349***	1.446***
	(0.065)	(0.064)	(0.089)	(0.229)	(0.214)	(0.383)
N	849	849	849	848	849	849

Table 3.6.: *Shirking and effort provision in the IC treatment as a function of market conditions. Random-effects models; the reported standard errors (in parentheses) account for potential clustering at the firm level. Column (1) – (3): the dependent variable is a dummy equal to 1 if the worker deviates from the contractually stipulated effort level. Column (4) – (6): the dependent variable is the level of effort provided in a given period. "# public offers" is the number of public contract offers available in the market in a given period. "Final period" is a dummy variable equal to 1 in the final period of the game.* *** / ** / * *indicate significance on the 1-percent / 5-percent / 10-percent level.*

offers in the market is associated with a significant reduction in workers' propensity to shirk. This holds after controlling for the contract terms faced by a worker, an "endgame dummy" to account for the sharp increase in shirking in the final period of the experiment, and a general time trend. Paralleling the observations on shirking, Columns (4)–(6) of Table 3.6 show that information signaling lower job finding chances is associated with a general increase in workers' performance in terms of effort level. This illustrates how changing market conditions feed back into workers' behavior under contractual incompleteness.

Our data also allow us to rule out alternative explanations for why the secondary-sector strategy becomes profitable. In particular, we find no evidence that workers who are employed in two-worker firms during the late phase of the experiment are inherently more willing to provide higher efforts for a given wage. In other words, there is no support for a sorting explanation in which two-worker firms become more profitable over time, on the basis of eventually finding especially diligent workers. Instead, our data indicate that a within-worker change in the willingness to provide effort under tighter labor market conditions accounts for the increased profitability of two-worker firms.[18]

Result 5: *Workers' behavior in the IC treatment changes in accordance with tightening market conditions. This contributes to the increasing profitability of two-worker firms and the long-run segmentation of the labor market.*

3.5. Conclusions

In this paper, we have provided empirical evidence and theoretical insights that involuntary unemployment and labor market segmentation may both arise as a consequence of contractual incompleteness. Our experiments show a causal impact of contractual incompleteness on the emergence of involuntary unemployment. While this link has long been discussed in theory, direct causal evidence on its relevance has remained scarce (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1989, Akerlof and Yellen 1990). Our findings also provide new insights on the efficiency-wage foundation of dual labor markets (Bulow and Summers 1986, Albrecht and Vroman 1992, Saint-Paul 1996). More precisely, we show theoretically and empirically that contractual incompleteness can lead to a segmented labor market, arising endogenously along with unemployment, although firms face identical technological constraints.

Our analysis suggests several avenues for future research. First, it would be interesting to study how the availability of additional contractual instruments affects unemployment and market segmentation. This holds for other explicit enforcement strategies (e.g., piece rates), as well as for other contracting schemes that require no verifiability of workers' performance, such as voluntary bonus payments or possibilities to sign multi-period employment contracts (e.g., MacLeod and Malcomson 1998, Huck et al. 2011). Second, our setup could be enriched by labor market institutions such as unemployment insurance or employment protection, which might affect workers' on-the-job behavior and employee turnover. This would be especially interesting since our model predicts that the emergence and characteristics of market segmentation depend on the job separation rate in equilibrium. Finally, our theoretical and empirical framework can be used to analyze how hiring decisions, rent payments, and other labor market outcomes interact with the available production technology. In our setup, decreasing returns to scale imply that one-worker

[18] This result comes from restricting the analysis to workers who are mainly employed in two-worker firms during the later phase of the IC treatment. We find that the willingness to provide effort for a given wage for these workers increases between the early and late phase. Thus, we observe a within-worker change, which is consistent with adaption to market conditions, but not with an explanation where two-worker firms eventually found a particularly hard-working group of workers.

firms and two-worker firms face an implicit reservation value of using the respective other hiring strategy. In many labor markets, there are additional factors that directly influence a firm's hiring incentives. For instance, firms might face market entry costs, have different outside opportunities available, or they might differ in the degree to which they can substitute capital and labor inputs. Studying how these factors affect market performance under contractual incompleteness is an important extension to our analysis.

4. Political Selection and the Concentration of Political Power

4.1. Introduction

Unfortunately, most politicians are neither benevolent nor omniscient. It is thus the role of political institutions to enforce the voter's interest within the political process. From the founding of modern democracies in the 18th century to recent constitutional drafts in Egypt and Lybia, political thinkers have been engaged in finding the best institutions for centuries. A central question in constitutional design has been whether political power should be concentrated on a group of political agents, typically the one winning the general election, or dispersed between different groups.[1] Strikingly, there are pronounced cross-country differences along this dimension, with classical extreme cases being the United Kingdom with concentrated power on the one hand, and Belgium and Switzerland with dispersed power on the other hand.

The Federalist Papers highlight two channels through which constitutions affect social welfare: the selection of competent politicians and the reduction of moral hazard of politicians in office.[2] The economic literature on the first channel, moral hazard, consistently finds that power-dispersing institutions increase welfare as they help to discipline egoistic incumbents. In contrast, economists have yet little to say about the second channel, political selection (see Besley 2005). Selecting competent politicians, however, is non-trivial. Since voters base their ballot on their perceptions of candidates' competence (Stokes et al., 1958; King, 2002; Pancer et al., 1999), politicians exert considerable effort in appearing

[1] For a discussion of this crucial issue and its relation to various specific institutions, see Lijphart (2012), Lijphart (1999) and Tsebelis (2002).

[2] "The aim of every political Constitution, is or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of society; and in the next place, to take the most effectual precautions for keeping them virtuous whilst they continue to hold their public trust" (Madison, 1788b)

competent and virtuous during electoral campaigns. This affects voters' capacity to identify and empower able candidates. A comprehensive appraisal of political institutions thus has to account for whether institutions hinder or enforce the selection of competent candidates for office.

The aim of this paper is to study the effects of power-concentrating institutions on campaigns and political selection. We consider a pre-election setup in which candidates are privately informed about their quality and are partly office motivated. Voters infer candidates' qualities from their campaigns. In this setup, we identify a trade-off that arises for changes in the level of power concentration. On the one hand, higher concentration of power implies a better allocation of political influence to competent candidates, as long as political campaigns provide at least some information to voters. We refer to this positive effect on welfare as the *empowerment effect*. On the other hand, more concentration of power increases the desire of office motivated candidates to win the election. Thus, mimicking of good candidates becomes more profitable, resulting in increasingly distorted policy choice. Campaigns convey less information on the competence of individual candidates and voters are less able to select high quality politicians. We call this negative welfare effect the *behavioral effect*.

We formalize our argument by a simple model in which two candidates compete in a public election by making binding policy proposals. In particular, candidates can either commit to a risky reform or to the (riskless) status quo. Candidates differ in their abilities, which are unobservable to the electorate. Only highly able candidates can increase expected welfare by adopting the risky policy, while less able candidates should stick to the status quo. Voters observe policy proposals and draw inferences about the candidates' abilities. In equilibrium, a reform proposal is associated with high ability and reforming candidates win the election more often than those proposing the status quo. Politicians do not only care about welfare but are also office motivated. This creates incentives for low-ability candidates to mimic the policy choice of their more able counterparts at the cost of adopting inefficient policies.

Variations in the level of power concentration involve a trade-off between the empowerment effect and the behavioral effect. The size of these effects depends on the relative weights of office motivation and policy motivation in the candidates' utility function. We find that the optimal level of power dispersion is higher, the more politicians are driven by office rents. If and only if politicians care predominantly about implementing efficient policies, it is optimal to concentrate power completely in the hands of the election winner. Conversely, if office rents are a strong component of candidates' motivation, some dispersion of political power enhances voter welfare.

The basic intuition behind this result is the following. Candidates' office motivation induces mimicking and distorts their policy choice. These distortions are fueled by the concentration of power because higher concentration increases electoral incentives. For particularly high office motivation, it is then optimal to reduce the resulting inefficiencies by dispersing political power among candidates. Hence, power dispersing institutions are beneficial if politicians are mainly office motivated.

We generalize the model in different aspects, finding that the qualitative results do not change. First, we introduce a continuous policy space such that candidates may choose the magnitude of reform they propose rather than limiting their choice set to a reform and the status quo. Second, we relax the assumption of binding policy commitments by introducing the possibility to withdraw a proposed reform after the election, under some circumstances. Third, we allow for heterogeneous policy preferences in the electorate. In this setting, we additionally show that increasing power dispersion reduces inequality in the society. If the social planner is inequality averse, the optimal level of power dispersion is consequently higher.

Data from international surveys like the *International Social Survey Panel* indicate considerable cross-country differences in voters' assessments of politicians' office motivation. Assuming that they mirror actual differences in politicians' motivation, our theoretical analysis gives rise to a testable hypothesis: Countries in which politicians are predominantly office motivated benefit from power dispersion. In contrast, countries with policy motivated politicians reduce welfare if they disperse political power.

In a cross-country design, we investigate the interaction effect of power dispersion and politicians' motivation on social welfare. For this purpose we combine measures of political institutions[3] with data on the perceived motivation of politicians. As a measure for the performance of the political system, we use growth in per capita GDP. Due to the availability of this data our analysis is restricted to eighteen established democracies. For this set of countries, the data provide support for our hypothesis. For countries with highly office motivated politicians, we find a positive relationship between power dispersion and growth, while we observe a negative relationship for countries with mainly policy motivated politicians.

Our analysis intends to identify the economic effects of power-dispersing institutions, which limit the office-holders discretion. Many economists have addressed this question abstracting from the problem of political selection. For a homogeneous set of politicians, power-dispersing institutions increase voter welfare. For example, Lizzeri and Persico (2001) demonstrate that office-motivated politicians campaign for more public good pro-

[3] We use Lijphart's index of the executive-parties dimension which considers five categories of political institutions to order political systems according to the implied dispersion or concentration of power (Lijphart, 1999).

vision and less pork-barrel spending under proportional representation than under the winner-takes-all regime. For the term in office, Persson and Tabellini (2003) show that voters are more able to discipline an incumbent if power is separated between multiple political agents.

These papers abstract from any heterogeneity in candidate quality and thus from the role of political selection.[4] The importance of incorporating the selection aspect into the analysis of political institutions is emphasized by Besley (2005). The process of selecting politicians that are qualified for office has two aspects. First, to choose among competing candidates the one who holds most promise to design and implement efficient policies. This part of political selection is based on electoral campaigns, which are typically studied in pre-election models (see, e.g., Downs, 1957; Lindbeck and Weibull, 1987). Second, to keep in office only politicians who adequately performed during the term. This aspect of political selection is based on the behavior of an incumbent, which is analyzed in post-election models in the spirit of Barro (1973) and Ferejohn (1986). While pre-election models are based upon the idea that politicians can commit to policies through campaigns, post election models assume that commitment is not possible. However, models of both types identify a common pattern. If an unobservable trait of politicians, e.g., competence, is important to voters, politicians will do their best to signal this trait. This affects politicians' behavior both before the election (Callander and Wilkie, 2007; Callander, 2008; Kartik and McAfee, 2007) and after it (Majumdar and Mukand, 2004) and should have an impact on the voters' capacity to select and retain the politician they want.

The role of institutions for political selection only recently came to attention and has so far only been studies in post-election settings. A first model addressing this question is Maskin and Tirole (2004). It investigates conditions under which the voter prefers political decisions to be taken by accountable politicians instead of non-accountable judges. Maskin and Tirole (2004) argue that holding public officials accountable in reelections is not optimal for all kind of policy decisions. While they do not compare alternative democratic institutions, this approach is taken by Smart and Sturm (2006) in a closely related paper. They study variations in the level of accountability through the introduction of term limits. Depending on the amount of noncongruent politicians, a limit of two terms as applied in many modern democracies is shown to be optimal.

Closest to our paper is the analysis by Besley and Smart (2007), who study the effects of several fiscal restraints on political selection in a post-election setting. Similarly to Maskin and Tirole (2004) and Smart and Sturm (2006), they show a trade-off between

[4]The assumption that candidates differ in a quality dimension, sometimes referred to as "valence issue", is applied in a large number of papers, including Adams (1999), Ansolabehere and Snyder (2000), Groseclose (2001), Aragones and Palfrey (2002), Sahuguet and Persico (2006) and Krasa and Polborn (2011).

disciplining incumbents and improving political selection. Whenever an institution allows to discipline bad incumbents, i.e., to make them adopt welfare-enhancing policies, this prevents effective political selection because voters are unable to distinguish a disciplined but bad politician from a good one. Our pre-election model produces a different trade-off. If voters have to infer the ability of candidates from their campaigns, dispersing power leads to both better policy choice and better selection, but comes at the cost of giving some political power to low-ability candidates. Besley and Smart (2007) consider four fiscal restraints, such as limits on the government size and transparency, which limit the office-holders' discretion. Our focus, in contrast, is on power-dispersing institutions, such as proportional representation, federalism, or public referenda. Interestingly, Besley and Smart (2007) find that three of the four restraints only increase voter welfare if there are sufficiently many good politicians. This contrasts our result according to which power dispersion is optimal if and only if the candidates are strongly driven by egoistic motives.

Finally, we also relate to a growing empirical literature on democratic systems and their effects on fiscal policy. The analyzes often focus on specific political institutions (see, e.g., Feld and Voigt, 2003; Persson and Tabellini, 2004; Enikolopov and Zhuravskaya, 2007; Blume et al., 2009; Voigt, 2011). In contrast, we apply a classification of political systems based on the implied dispersion of political power, thus encompassing a broad range of institutions. Lijphart (1999, 2012) as well as Armingeon (2002) study the influence of the dispersion of power on various political and economic outcomes. While Lijphart (1999) finds no effect of power dispersion on economic variables, Armingeon (2002) finds a negative effect of power dispersion on unemployment and inflation. Complimenting these findings, we show that the effect of power dispersion on growth in GDP positively interacts with the strength of politicians' office concerns.

4.2. The model

Our model studies the effects of institutions on candidates' campaigns and the empowerment of competent politicians. Candidates differ in quality, more precisely in the ability to implement welfare-enhancing policies. They are privately informed about their abilities and commit to a policy prior to the election. Voters observe candidates' campaigns and vote based on the expected welfare each candidate provides. We depict political institutions in reduced form, by means of how much political power is concentrated in the political system. With higher concentration of power, the candidate receiving a majority of votes is more capable to enforce his agenda.

The game consists of three stages. At the first stage, nature independently draws both candidates' abilities a_1 and a_2, which are privately revealed to the candidates. At the second stage, both candidates simultaneously make binding policy proposals, x_1 and x_2. At the third stage, voters observe the proposals, update their beliefs about the candidates' abilities and cast their votes. Based on the election result, the power allocation rule divides political power between both candidates. Finally, each candidate's power determines the enacted policy.

While the basic model serves to clarify the main arguments, we discuss a number of modifications in Section 4.7. In particular, we allow for a continuous policy space, a form of limited commitment and heterogeneity in the voters' policy preferences. Importantly, these modifications do not alter the main results of the basic model.

4.2.1. Voters

There is a continuum of fully rational and risk neutral voters of mass one who have preferences both over policy and candidates. Voters have identical preferences regarding the considered policy field. In this policy field, either the status quo can be maintained ($x_i = 0$) or a reform can be implemented ($x_i = 1$). All voters receive a positive payoff from the reform if and only if it is successful. More precisely, we assume that a successful reform yields a return of 1 to each voter while a failing reform leads to a return of zero. Whenever a reform is adopted, all voters bear a cost of c. Maintaining the status quo gives a certain payoff of zero.

Voters might also care about other policy fields and about the candidates' ideologies or personal characteristics. We account for these preferences in the tradition of the probabilistic voting model by assuming that voters have heterogeneous candidate preferences (Lindbeck and Weibull, 1987).[5] If candidate 1 is in power, voter k receives an additional utility of μ_k, while we normalize the additional utility if candidate 2 enters office to zero. Let μ_k be distributed according to some symmetric cdf Ω with expected value zero. Furthermore, let Ω have full support on the interval $[-1, 1]$ to guarantee heterogeneity in the resulting voting preferences. Altogether, if candidate i is in power and sets policy x_i, voter k receives a utility of

$$V_k(x_i, i) = \begin{cases} (i)\mu_k + 1 - c & \text{reform succeeds} \\ (i)\mu_k - c \quad \text{if} & \text{reform fails} \\ (i)\mu_k & \text{status quo is maintained,} \end{cases}$$

where (i) denotes the indicator function which is one if $i = 1$ and zero otherwise. Voter k prefers candidate 1 if and only if he expects $V_k(x_1, 1)$ to be larger than $V_k(x_2, 2)$. Voters vote sincerely, i.e., each voter casts his vote for his preferred candidate. Hence, candidate

[5] Note that our results are independent of whether these candidate preferences are subject to an additional aggregate shock as in Lindbeck and Weibull (1987).

i's vote share depends positively on the voters' belief about the payoff he provides, and negatively on the belief about the payoff provided by his opponent.

4.2.2. Candidates

There are two candidates running for office. Each candidate i can either commit to a reform ($x_i = 1$) or to the status quo ($x_i = 0$). More able candidates design better reforms, i.e., reforms that are more likely to succeed. We measure candidate i's ability by the implied probability of a successful reform, $a_i \in [0, 1]$. Hence, policy x_i set by candidate i provides an expected payoff of $x_i(a_i - c)$ to each voter. We refer to it as the candidate's welfare contribution.

Prior to the election, nature independently draws both candidates' abilities from the cumulative distribution Φ. Let the corresponding density function ϕ have full support on $[0, 1]$ and be continuously differentiable. After observing his ability, each candidate i commits to policy $x_i \in \{0, 1\}$. Thus, the strategy X_i of politician i is a mapping from abilities to policy commitments.

Candidates care about gaining political power (office motivation) as well as about the expected welfare contribution of the policy they have designed (policy motivation). The utility function of politician i is given by

$$U_i(a_i, x_i) = f(v_i, \rho)\left[\theta + x_i(a_i - c)\right], \tag{4.1}$$

where $\theta > 0$ denotes the relative weight of office motivation. Candidate i's power $f(v_i, \rho) \in (0, 1)$ equals the probability that he can enforce his policy proposal x_i. It depends on his vote share v_i and the parameter of power concentration ρ, representing the set of political institutions. To simplify notation, this utility function is formulated at an ex interim stage, i.e., taking the expected payoff after the election but before the reform outcome has been realized. We also omit the dependence of v_i on both candidates' strategies and actions.

Note that candidate i only cares about how expected welfare is affected by his policy choice, not about voter welfare in general. This way to formulate policy preferences of politicians has been introduced by Maskin and Tirole (2004), who phrase it *legacy motive*. It captures the idea that politicians have a desire to leave a positive legacy to the public.

4.2.3. Institutions of power allocation

We model political institutions by a power allocation rule f that translates election results into an allocation of political power, i.e., each politician's probability to implement his policy proposal. Technically, candidate i's power $f(v_i, \rho)$ depends on his vote share v_i and on the level of power concentration ρ implied by the set of political institutions.

Definition 4.1. *The continuously differentiable function* $f : [0,1] \times \mathbb{R}_+ \to [0,1]$ *is a power allocation function if it satisfies:*

1. *monotonicity in* v_i*:* $\frac{\partial f(v_i,\rho)}{\partial v_i} > 0 \;\forall\; \rho$
2. *symmetry in* v_i*:* $f(v_i, \rho) = 1 - f(1 - v_i, \rho)$
3. *piece-wise monotonicity in* ρ*:* $\frac{\partial f_i(v_i,\rho)}{\partial \rho} > 0 \;\forall\; v_i \in (1/2, 1)$.

The first property rules out that candidates receive a larger amount of political power if they gain less votes in the election. The second establishes anonymity, i.e., the constitution does not treat candidates differently. The third allows us to interpret any rise in ρ as an unambiguous increase in the concentration of power. The higher is ρ, the larger is the amount of power assigned to the election winner, i.e., the candidate that gains more than half of the votes.

This modeling approach allows to study a large variety of institutional differences. Figure 4.1 illustrates how political institutions can be represented by power allocation functions. Each of the four panels depicts two examples of allocation functions. Throughout, the solid line represents an institutional setup that concentrates power more strongly than the one corresponding to the dashed line.

Panel I depicts two stylized allocation rules frequently used to compare electoral systems in the theoretical literature (see, amongst others, Lizzeri and Persico 2005). The solid line represents institutions that fully concentrate power in the hands of the election winner. The step function is the standard way to model plurality voting, also known as winner-takes-all system. The dashed line represents proportional representation, which implies a lower concentration of political power and is often modeled by the identity function $f(v, \rho) = v$.

A less simplistic representation of these two systems is shown in Panel II. Here, the amount of power depends on the margin of victory, e.g., because delegates might occasionally vote against the party lines. Plurality voting tends to generate clear-cut majorities, as the winning party typically receives a share of parliamentary seats beyond its vote share. In contrast, the allocation of seats corresponds closely to vote shares under proportional representation. Thus, the dashed curve for the proportional system is flatter than the one for plurality voting.

In Panel III, the dashed line represents a political system with a supermajority requirement for certain policy decisions (as employed in Germany and the US). This requirement generates additional steps in the power allocation function, since some policies can only

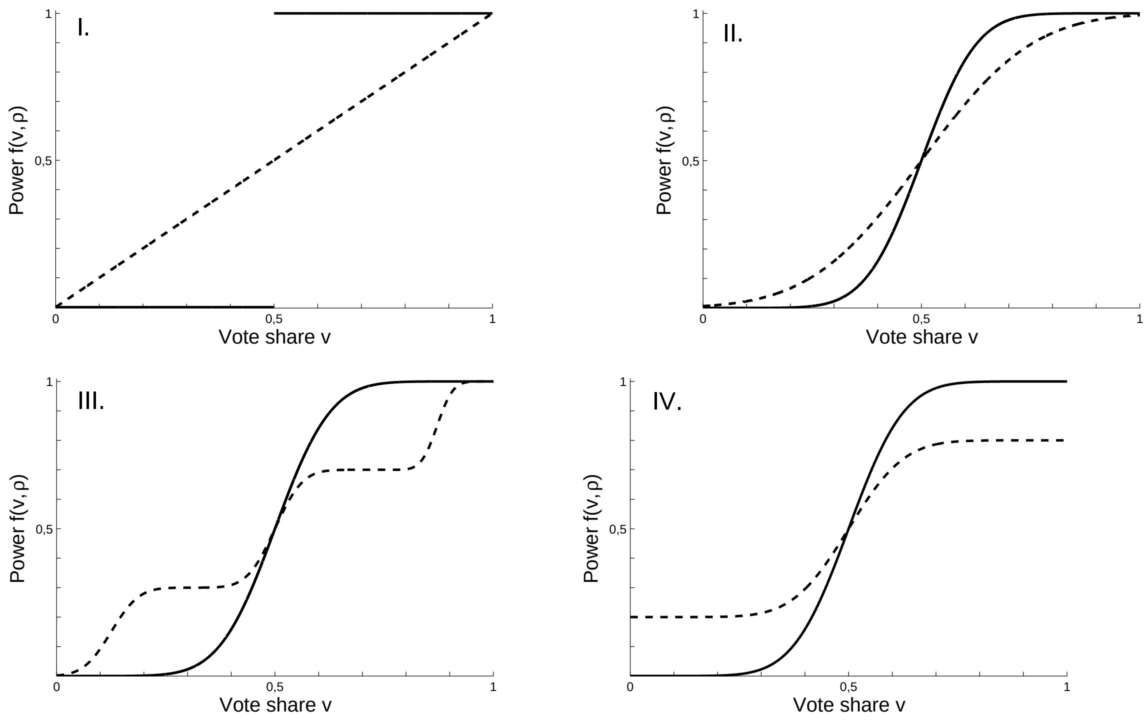


Figure 4.1.: Political institutions and the corresponding power allocation rules.

be enforced after a landslide victory. In contrast, the solid line corresponds to a system as applied in the UK where any decision can be taken by a simple majority.

Finally, the dashed line in Panel IV depicts the use of direct democratic measures as employed for example in Switzerland. Even with a landslide victory in the election, the winning party cannot always implement their agenda. The opposition party can block policies via a referendum or even enforce their own proposals. Thus, only a limited part of political power is at stake in the parliamentary election (similar arguments can be made with respect to federalism, bicameralism or a constitutional court).[6]

4.2.4. Equilibrium concept and normative criterion

To solve this game, we study Perfect Bayesian equilibria. Thus, an equilibrium of this game consists of a strategy profile and a belief system such that (1) both candidates play mutually best responses at the announcement stage, anticipating the winning probabilities for each vector (x_1, x_2) that are implied by the voters' beliefs, and (2) the voters' belief system σ is derived from the candidates' strategies X_1, X_2 according to Bayes' rule everywhere on the equilibrium path.

[6]Note that our definition of the power allocation function does not allow to compare all institutional settings. For example, the dashed lines in Panels III and IV of Figure 4.1 cannot be ordered unambiguously in terms of power concentration, as they intersect more than once.

In the following, we compare the effects of changes in power concentration, i.e., in parameter ρ. As normative criterion, we use a utilitarian welfare function in ex ante perspective, i.e., expected welfare before candidates' abilities are drawn:

$$W(\rho, \theta) = \int_0^1 \int_0^1 \phi(a_1)\phi(a_2) \sum_{i=1}^{2} f(v_i, \rho) X_i(a_i)(a_i - c) \; da_2 da_1.$$

Welfare is, hence, given by the weighted sum of the politicians' welfare contribution, integrated over all possible combinations of candidates' ability. The weights correspond to the candidates' power, $f(v_i, \rho)$. Note that welfare is calculated using equilibrium strategies, which are functions of the parameters ρ (power concentration) and θ (candidates' office motivation).[7]

4.3. Benchmark case: perfect information

If individual abilities are observable to the electorate, voters condition their ballot on candidates' abilities and reform proposals. In particular, the fraction of citizens voting for a candidate is increasing in his welfare contribution.

Candidates' choices are driven by two motives. They seek to obtain power and to contribute to voter welfare. Under perfect information, these motives are fully aligned: Each candidate maximizes his electoral prospects by proposing the policy with highest welfare contribution. Hence, a reform is only proposed by high-ability candidates with $a_i \geq c$. In contrast, a candidate with ability $a_i < c$ gains more power by proposing the status quo instead of a reform with a negative welfare contribution. Thus, candidates' policy choices are undistorted: A politician proposes to implement a reform if and only if the reform enhances welfare. As a consequence, candidates with higher ability, i.e., those who propose to reform, receive higher vote shares in the election.

This result has a direct welfare implication. While variations in the power concentration parameter ρ do not distort candidates' behavior, higher concentration of power allocates more power to candidates' with higher welfare contribution. Hence, welfare strictly increases with the level of power concentration ρ. The following Proposition summarizes these results (all proofs are found in Appendix A.3.1).

Proposition 4.1. *Under perfect information, candidates propose a reform if and only if $a_i \geq c$. Welfare is maximized if political power is completely concentrated.*

[7]We neglect the candidates' utilities when calculating welfare.

4.4. Imperfect information

In the following, we derive the equilibrium properties under the assumption that candidates are privately informed about their abilities. Under imperfect information, voters form beliefs about politicians' welfare contribution on the basis of their policy proposals. The vote share of candidate i thus depends on the proposals only, and not on candidates' abilities. However, the higher a candidate's ability a_i, the higher is his welfare contribution from a reform. A candidate's incentive to propose a reform is thus monotonically increasing in his ability. Hence, the optimal behavior of candidates exhibits the cutoff property, i.e., all candidates with ability greater or equal to some cutoff value α_i propose a reform, while candidates with lower ability propose the status quo.

Lemma 4.1. *Given any belief system σ and any strategy of the opponent X_j, the optimal strategy of candidate i can be characterized by a unique cutoff $\alpha_i \in [0, 1]$ such that*

$$X_i(a_i) = \begin{cases} 1, & \text{if } a_i \geq \alpha_i \\ 0, & \text{if } a_i < \alpha_i. \end{cases}$$

In the following, we denote strategies only by their corresponding cutoffs. Candidates compare the utility of proposing a reform and of proposing the status quo, given the strategy of the opponent and voters' belief system σ. They propose a reform if the following utility difference is positive:

$$R_i(a_i, \alpha_i, \alpha_j, \rho) = E[U_i(a_i, x_i = 1)|\alpha_i, \alpha_j, \sigma] - E[U_i(a_i, x_i = 0)|\alpha_i, \alpha_j, \sigma].$$

We refer to this utility difference as the reform incentive function R_i.

Proposition 4.2. *Two classes of Perfect Bayesian equilibria exist. In class one, equilibria exhibit symmetric cutoffs $\alpha_i = \alpha_j = \alpha$ smaller than c. In class two, at least one of the cutoffs is equal to 1. Class one is non-empty and consists of divine equilibria only, while all equilibria in class two are not divine.*

To refine the set of equilibria, we apply the divinity criterion proposed by Banks and Sobel (1987), which restricts the feasible set of out-of-equilibrium beliefs. In any divine equilibrium, a deviation from equilibrium actions must be attributed to the type that profits most of it. For all equilibria in class two, at least one of the candidates always proposes the status quo even if he has the highest ability possible. This can only be optimal if voters are convinced that candidates proposing a reform have low ability. In this case a reform proposal is associated with a substantially lower vote share than the status quo proposal. According to the divinity criterion, however, the reform proposal necessarily needs to be attributed to the most able type of candidate, whose incentive to propose a reform is highest among all candidates. Hence, such an equilibrium cannot be divine.

The first class of equilibria, in contrast, is robust to the divinity criterion. If $\alpha \in [0, c)$, all actions are played in equilibrium and there are no out-of-equilibrium beliefs. If the cutoffs are equal to zero, even the least able candidates propose a reform. If any candidate deviates to the status quo, his associated welfare contribution is zero independently of his ability. Since the welfare contribution of such a deviation is known, out-of-equilibrium beliefs about abilities are irrelevant for the voting decision. Hence, restricting these will not eliminate this equilibrium. In the following, we only consider equilibria of this class.

To see that all equilibria in this class are characterized by $\alpha \leq c$, note that a reform proposal is always associated with a positive welfare contribution. In other words, the expected vote share increases if a reform is proposed. If voters associated a negative welfare contribution with a reform proposal, candidates with ability below c would never choose to propose a reform. Otherwise, they would suffer from a negative welfare contribution as well as from a loss in expected office utility. Clearly, this is a contradiction: If only candidates with ability above c were to choose a reform, the associated welfare contribution could not be negative. Hence, a reform is associated with a higher vote share than the status quo. It follows that candidates with ability above c always choose to reform. They gain not only from their welfare contribution but also from an increase in expected office rewards. Thus, the equilibrium cutoffs must be below c.

Definition 4.2. *A perfect Bayesian Nash equilibrium with $\alpha > 0$ is an informative equilibrium.*

Next, we derive the equilibrium condition for informative equilibria.[8] In any informative equilibrium, candidates with high ability choose to propose a reform while candidates with low ability choose the status quo. Policy proposals hence convey an informative signal about candidates' abilities. The cutoff type is indifferent between proposing a reform or refraining to do so. Thus, the equilibrium cutoff α is implicitly defined by $R_i(\alpha, \alpha, \alpha, \rho) = 0$. The resulting equilibrium condition is

$$R(\alpha, \rho) = \underbrace{\theta \left(f(v^r, \rho) - \frac{1}{2} \right)}_{\text{Change in office utility}} + \underbrace{\left[\frac{1}{2} + \Phi(\alpha) \left(f(v^r, \rho) - \frac{1}{2} \right) \right] (\alpha - c)}_{\text{Change in welfare contribution}} = 0, \tag{4.2}$$

where v^r represents the vote share from proposing a reform, when facing an opponent who proposes the status quo.

Both aspects of the politicians' preferences can easily be distinguished in Equation (4.2). The change in office utility stems from the expected increase in office rewards due to a

[8]We discuss non-informative equilibria in the extensions.

reform proposal. This term is always positive, since a reform proposal is associated with a higher vote share. However, the politician also cares about the welfare contribution that is induced by his proposal. The first part of the change in welfare contribution stands for the probability that a proposed reform is implemented. The second part is the welfare contribution of a reform implemented by the cutoff type. Note that this term represents a loss, since the ability of the cutoff type α is below the cost c. The cutoff type is hence willing to make a negative welfare contribution to increase his chances to enter office. Next, we establish the uniqueness of divine equilibria.

Assumption 4.1. *The ability distribution $\phi(a)$ is bounded from above with $\phi(a) < \frac{1+\Phi(a)}{c-a}$ for all $a < c$.*

Assumption 4.1 is a regularity assumption on the ability distribution, which is fulfilled for example for the uniform distribution. It ensures that the reform incentive is monotonically increasing in the cutoff. Hence, both incentive functions cannot intersect more than once and there is only one symmetric equilibrium. For the remainder of the paper, we take Assumption 4.1 as given.[9]

Proposition 4.3. *There is a unique divine equilibrium. Moreover, it exists $\tilde{\theta}(\rho) \in \mathbb{R}^+ \cup \{\infty\}$, such that this equilibrium is informative if and only if $\theta < \tilde{\theta}(\rho)$.*

For any given level of power concentration, the existence of informative equilibria depends on the level of office motivation. If office motivation is not too large ($\theta < \tilde{\theta}(\rho)$), cutoffs are larger than zero, and the unique divine equilibrium is informative. Otherwise, all candidates propose a reform, and the equilibrium is non-informative. If the average ability is smaller than the costs, the unique equilibrium is informative for all $\theta < \infty$. In this case, a non-informative would imply a negative welfare contribution from a reform so that candidates with low ability would prefer to propose the status quo.

4.5. The effects of power-concentrating institutions

Empirically, democratic countries differ strongly in their political institutions and the levels of power concentration implied by them. As we have argued in Subsection 4.2.3, our framework allows to represent these differences by means of an appropriate power allocation function f. In this section, we study the effects of these variations, captured by changes in the parameter of power concentration ρ.

[9] If Assumption 4.1 is not given, multiple equilibria may arise. The following analysis is not changed, if we restrict it to the welfare optimal equilibrium.

4.5.1. Effects on candidates' behavior

The power allocation function f determines the electoral incentives of political candidates. Under perfect information, variations in power concentration leave the behavior of politicians unaffected. As low-ability candidates are not able to mimic their more able counterparts, policy choice is always efficient (see Proposition 4.1).

With asymmetric information and office-motivated candidates, in contrast, policy choice is distorted as some low-ability candidates propose welfare-reducing reforms. It turns out that political institutions affect the magnitude of these policy distortions.

Proposition 4.4. *In any informative equilibrium, increasing power concentration ρ leads to the proposal of more inefficient reforms: $\frac{d\alpha}{d\rho} < 0$.*

Consider some level of power concentration ρ_0. By construction, the cutoff type with ability $a_i = \alpha_0 < c$ is indifferent between proposing the reform and the status quo. We find that after an increase in power concentration, the cutoff type strictly prefers to propose the reform. In particular, his utility of proposing the status quo decreases while his utility of proposing the reform increases.

If the cutoff type proposes the status quo, his welfare contribution is equal to zero. Thus, his utility is only determined by the office rents he receives according to his expected amount of power. With increasing power concentration, office rents are reduced because he receives less power when running against a reforming opponent.

If the cutoff type proposes a reform, his utility is composed of two parts. He receives office rents but also incurs a utility loss due to a negative welfare contribution. As he is indifferent between proposing a reform and the status quo, the sum of the two is positive. With increasing power concentration, both office rents and negative welfare contribution increase by the same factor. Hence, the sum of both, representing his utility from a reform proposal, also increases by this factor.

Consequently, with higher levels of power concentration, status quo proposals lead to lower utility while reform proposals become more profitable. The equilibrium cutoff is thus decreasing in the level of power concentration.

4.5.2. Welfare effects

In the following, we study the effects of power-concentrating institutions on the performance of the political system. The objective function is the utilitarian welfare function formulated in ex ante perspective, i.e., before the candidates' abilities are drawn.

With privately informed candidates, the relation between power concentration and welfare is not as clear-cut as under perfect information. On the one hand, there is still a positive *empowerment effect* of power concentration. Whenever both policies are proposed, the majority of votes goes to the reforming candidate, who provides a higher expected welfare than the candidate proposing the status quo (see Section 4.4). Consequently, any increase in power concentration ρ assigns more power to the appropriate candidates.

On the other hand, the previous section demonstrated a negative *behavioral effect* of power concentration. By reinforcing the electoral stakes, a stronger concentration of power leads to the proposal of more inefficient reforms. This reduces the information revealed during the campaigns and limits the voters' capacities to allocate power to high-ability candidates. We show below that the relative size of the two effects, and thus the overall sign, depend on the average payoff from a proposed reform. This in turn is determined by the level of power concentration itself. Under weak conditions, however, there is a unique welfare optimum.

Assumption 4.2. *The ability distribution has a non-decreasing reversed hazard rate* $\Phi(a)/\phi(a)$*.*

In the following, we assume that this regularity condition is satisfied.

Lemma 4.2. *The welfare function* W *is strictly quasi-concave in* ρ*.*

Lemma 4.2 implies that the welfare function has at most one maximum in ρ. In other words, the behavioral effect and the empowerment effect are equalized at most once, and the optimal level of power concentration is well-defined.

To prove Lemma 4.2, we analyze how the empowerment effect and the behavioral effect are influenced by changes in power concentration. The positive empowerment effect results because, with increasing ρ, a reforming candidate receives more power if he runs against an opponent proposing the status quo. Thus, its size depends positively on the average payoff from a proposed reform and the probability that a candidate advocates the status quo, $\Phi(\alpha)$. The empowerment effect is consequently highest at $\rho = 0$, i.e., completely dispersed power. With increasing ρ, the average reform payoff becomes smaller as more inefficient reforms are conducted. At the same time, the probability $\Phi(\alpha)$ is reduced because less types propose the status quo. Thus, the empowerment effect is strictly decreasing in ρ, as illustrated in Figure 4.2.

The negative behavioral effect results because increasing power concentration leads to a reduction in the cutoff α. The size of this effect depends on, first, how harmful this decline in α is to welfare, and second, how sensitively α reacts to changes in power concentration. The effect of changes in power concentration on the size of both factors is directly opposed. On the one hand, a marginal decrease in α reduces the average reform payoff by $\phi(\alpha)(\alpha - c)$. With completely dispersed power ($\rho = 0$), the cutoff is equal to c, and a marginal change in α does not affect welfare. With increasing power concentration, α departs further from c and the additionally proposed reforms are increasingly harmful for welfare. Hence, the

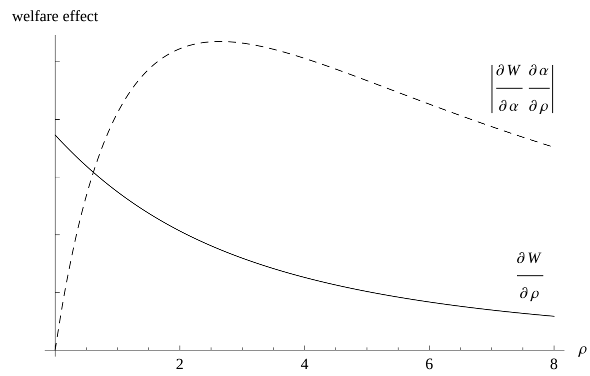


Figure 4.2.: *The welfare effects of a change in ρ for a logistic power allocation function with mean $\mu = 0.5$ and scale parameter $\beta = 1/\rho$, μ_k distributed according to $\mathcal{N}(0, 0.5)$, a uniform ability distribution, $\theta = 1$, and $c = 0.6$. The solid line represents the (positive) empowerment effect, the dashed line represents the (negative) behavioral effect. The optimal level of ρ is attained at the intersection of both lines.*

size of the first factor increases in ρ. On the other hand, the sensitivity of α depends on the additional vote share a candidate gains by proposing a reform. For higher levels of power concentration, the average payoff from a proposed reform is smaller and so is the additional vote share. Thus, the cutoff α reacts less sensitively to further increases in ρ and the size of the second factor decreases in ρ. As a consequence, the behavioral effect is not monotonic in ρ (see Figure 4.2).

In Appendix A.3.1, we show that the sign of the overall welfare effect depends only on the relative sizes of the probability to face a status quo proposing candidate, $\Phi(\alpha)$, and the marginal welfare contribution, $\phi(\alpha)(c - \alpha)$. The standard assumption of a monotone hazard rate guarantees that, at every local extremum, the empowerment effect is strictly increasing in the cutoff relatively to the behavioral effect. Since α depends negatively on ρ, the relative size of the empowerment effect is strictly decreasing in ρ. Thus, there cannot be a local minimum and at most one local maximum, which corresponds to the definition of quasi-concavity.

Proposition 4.5. *If and only if office motivation is below some threshold level $\bar{\theta}$, welfare is maximized by full concentration of power. If instead $\theta > \bar{\theta}$, it is optimal to disperse power to some degree $\rho^*(\theta) \in (0, \infty)$, and the optimal concentration of power is strictly decreasing in the candidates' office motivation: $\frac{d\rho^*}{d\theta} < 0$.*

Proposition 4.5 establishes a relation between candidates' motivation and the optimal level of power concentration. For two reasons, changes in office motivation affect the relative size of empowerment effect and behavioral effect. First, increasing office motivation θ implies that candidates respond more strongly to electoral incentives. Thus, the negative behavioral effect is reinforced. Second, increasing θ results in larger policy distortions, i.e., a lower cutoff α. As argued above, the reduction in α implies that the negative behavioral effect increases relatively to the positive empowerment effect at the optimal level of concentration. In total, an increase in office motivation makes power concentration less beneficial for voters.

Regarding the optimal level of concentration, we have to distinguish two cases. First, consider the case of mainly policy-oriented candidates, in which mimicking is not prevalent. For $\theta < \bar{\theta}$, the behavioral effect is sufficiently small to be dominated by the positive empowerment effect for all levels of ρ. Consequently, welfare is maximized by full concentration of power. For higher levels of office motivation, $\theta > \bar{\theta}$, both effects outbalance each other at some interior level of power concentration $\rho^* \in (0, \infty)$, representing the optimal political institutions. With any further rise in θ, the relative increase of the behavioral effect implies that both effects are equalized at a strictly lower level of ρ. Hence, the more office-motivated politicians are, the more dispersion of power is optimal.

Intuitively, office motivation represents a source of inefficiency in our model because it induces policy distortions, which are aggravated by high levels of power concentration. The more severe the implied inefficiency is, the more beneficial it is to attenuate these distortions through power-dispersing political institutions.

4.6. Empirical analysis

In this section, we analyze whether data for established democracies is in line with our model predictions. Our model establishes a relationship between democratic institutions and the efficiency of implemented policies. Crucially, the effect of institutions depends on the motivation of politicians. Proposition 4.5 states that power concentration is always conductive to the implementation of efficient policies at low levels of office motivation, $\theta < \bar{\theta}$. At higher levels of office motivation, however, the optimal level of power concentration is interior and power should not be maximally concentrated. Moreover, the optimal degree of power concentration declines for further increases in office motivation. Since the welfare function is quasi-concave, the implications of our model can be summarized in the following Hypothesis.

Hypothesis 4.1. *The effect of power concentration on welfare depends on the level of politicians' office motivation. Power concentration has positive effects on welfare if politicians are mainly policy motivated. In contrast, if politicians are mainly office motivated, the welfare effect of power concentration is significantly smaller or even negative.*

We discuss an operationalization of the relevant variables and an empirical test of this Hypothesis in the following.

4.6.1. Operationalization

The test of our model's predictions requires three basic measures. As the dependent variable, we need a measure of efficient policies. Key independent variables are the degree of power dispersion within the political system and the extent of politicians' office motivation.

Any efficient policy, as described by our model, benefits the public as a whole. As a measure for *efficient policies*, we use growth in real GDP per capita (World Bank). It provides a concise and objective measure of developments that bear the potential of welfare improvements for the general public. Growth has been used as outcome variable by a number of other empirical studies on political institutions as Feld and Voigt (2003) and Enikolopov and Zhuravskaya (2007). Other frequently used outcome measures relate to fiscal policy (see Voigt, 2011), which is not addressed in our model.

Several measures of democratic institutions have been discussed in the literature. The *allocation of power* that is implied by political institutions is well described by Lijphart's index of the executive-parties dimension (Lijphart, 1999). This well-established measure quantifies how easily a single party can take complete control of the government. High values of the index correspond to high dispersion of power within the political system. It focuses on economically developed countries with a long democratic tradition and hence covers 36 countries. The measure is based on the period 1945-1996. New Zealand underwent major constitutional changes after 1996 and is thus excluded from the analysis. Its inclusion, however, does not change the qualitative results.

While *office motivation* cannot be measured objectively, indication for it may come from surveys of voters.[10] The International Social Survey Panel (ISSP) includes questions on voters' opinion about politicians.[11] The item relevant to our study was included in its 2004 survey, which was performed in most democratic states: 'Most politicians are in politics only for what they can get out of it personally.' Agreement with this statement is coded on a five point scale. We use the mean points of all survey participants in a country as our measure for the importance of office motivation in our model. That means, we assume that differences in this item reflect differences in politicians' motives.

For an easy interpretation of regression results, we normalize the indices for both office motivation and power dispersion to range between zero and one. High values indicate

[10] Alternatively, one could use measures that are based on experts' assessments like the Corruption Perception Index from Transparency International and the Worldwide Governance Indicators from Kaufmann et al. (2009). However, these indices focus on rent extraction and not on private motivations of politicians in general.

[11] Other surveys, as the World Values Survey, the Global Barometer Survey, the Eurobarometer, or the European Value Survey query trust or confidence in institutions, such as the political parties and the national parliament. Such questions only indirectly relate to politicians' motivation.

pronounced office motivation of political leaders or a strong dispersion of political power, respectively.

4.6.2. Design

Our analysis focuses on countries with a similar degree of democratization. We require that all countries be established democracies as identified by the 2002 Polity IV Constitutional Democracy index (Marshall and Jaggers, 2010). All countries have to feature an index of 95 or higher, which excludes Venezuela from the sample. The remaining 18 countries in the sample are similar with respect to their economic characteristics. In particular, all countries are highly economically developed as classified by the World Bank. They furthermore feature a Human Development Index (HDI) of at least 0.9 in 2004, which places them in the top quintile of all countries.[12]

The time-invariant regressors require a cross-country analysis. To address problems of reverse causality, we use 2004 as the base year for the regression. Our explained variable captures growth after 2004, while our explanatory variables are measured at or before this year. We control for variables that may be correlated with both our explanatory variables and our explained variable. Most notably, past economic performance affects growth (see, e.g., Barro, 1991; Sala-i-Martin, 1994) and may also alter voters' perception of politicians. We hence control for GDP per capita of 2004. Also other variables have been shown to robustly affect growth, such as capital accumulation, school enrollment rates, life expectancy, or openness of the economy (see, e.g., Sala-i-Martin, 1997). To capture such influences and to keep the number of explanatory variables low, we add past growth in real GDP per capita (from 1991 to 2004) to the regression. All variables used in our regression are summarized in Appendix A.3.3.

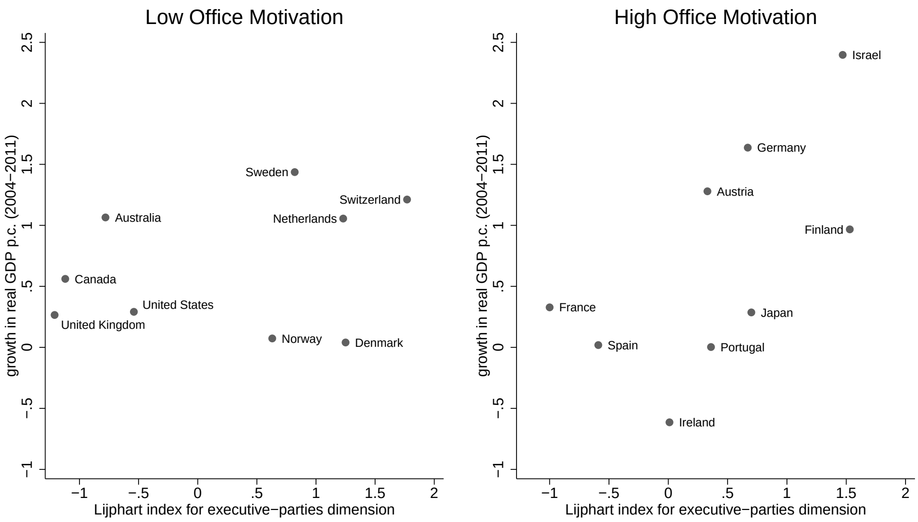
4.6.3. Results

As a first step in our analysis, we split the country set at the median value of politicians' office motivation. Figure 4.3 plots growth against the dispersion of power for the two sets of countries. The left panel depicts the relationship for countries in which politicians' office motivation is below its median value, while the right panel depicts the relationship for countries in which politicians' office motivation is above its median value. The figure suggests that power dispersion has only small effects if politicians are mainly policy motivated, whereas power dispersion is beneficial for growth if politicians are mainly office

[12]The similarity in socioeconomic development was formulated as a major prerequisite for cross-country analyses in Armingeon (2002).

motivated. For both groups of countries, the unconditional correlations between power dispersion and growth support this observation.[13]

Figure 4.3.: Relationship between power dispersion and growth



A statistical test of the effect of power dispersion on economic growth is performed in the OLS regressions presented in Table 4.1. Test statistics are based on White heteroscedasticity-consistent standard errors. Column (a) displays the regression results if the interaction term between power dispersion and politicians' office motivation is omitted. The unconditional effect of power dispersion on economic outcomes is insignificant.

This picture changes if the interplay between power dispersion and politicians' motivation is taken into account. Column (b) reports the corresponding regression results. Note the positive and significant coefficient of the interaction term between power dispersion and office motivation. It implies that power dispersion is more positively related to growth, the more office motivated politicians are. The inclusion of the interaction term in the regression also strongly increases the explanatory power of the econometric model. The adjusted R^2 increases from 0.19 to 0.49. In the analysis of political institutions, neglecting the interplay between power dispersion and politicians' office motivation would thus ignore a relevant aspect.

As it turns out, power dispersion goes along with either increased or decreased growth prospects depending on the level of politicians' office motivation. The conditional effect of power dispersion at the lowest and the highest level of office motivation are reported in

[13]For countries with high levels of politicians' office motivation, there is a positive and weakly significant relationship between growth and power dispersion (Pearson's correlation coefficient, $\rho = 0.618$, $p = 0.0759$), while there is no significant relationship between the two variables for countries with low levels of politicians' office motivation (Pearson's correlation coefficient, $\rho = 0.291$, $p = 0.447$).

Table 4.1.: OLS regression results

	Growth in real GDP per capita (2004-2011)	
	(a)	(b)
Power dispersion	0.852	−3.565*
	(0.490)	(1.637)
Office motivation	−0.125	−6.382**
	(1.090)	(2.666)
Power dispersion · office motivation		**8.948****
		(3.520)
Real GDP per capita in 2004 (in $ 1000)	−0.0247	−0.0436*
	(0.0251)	(0.0214)
Growth in GDP per capita (1991-2004)	−0.267***	−0.352***
	(0.0789)	(0.0586)
Constant	1.530	5.410***
	(1.045)	(1.583)
adjusted R^2	0.19	0.49
F	4.38	14.17
N	18	18

Standard errors are provided in brackets. ***, **, * indicate significance at the 1-, 5-, and 10-percent level, respectively.

Table 4.2. At the lowest level of office motivation, power dispersion is negatively related to growth whereas this relationship turns out to be positive at the highest level of office motivation. Our analysis thus leads to the following result.

Result 1. *The higher is office motivation the more positively is the link between power dispersion and growth. Furthermore, power dispersion is positively related to growth if and only if politicians' office motivation is high. If politicians' are mainly policy motivated, power dispersion comes along with reduced growth.*

This result is in support of the hypothesis derived from our model. We do not only observe that the interaction effect is as predicted, but also that the effects of power dispersion changes its sign as suggested by the theory. In total, we conclude that the data is in line with the model presented in this paper.

Table 4.2.: Effect of power dispersion

	low office motivation	high office motivation
Coefficient	−3.565*	5.382***
Standard error	1.637	1.955

The table depicts the coefficient of power dispersion for the lowest level of office motivation ($\theta = 0$) and for the highest level of high office motivation ($\theta = 1$). ***, **, * indicate significance at the 1-, 5-, and 10-percent level, respectively.

4.6.4. Discussion of empirical results

We conduct several robustness checks for our empirical test. In the following, we discuss the use of different indicators for our main variables, the period of the financial crisis in our data, and an alternative explanation for our result.

We check whether the positive and significant interaction term between power dispersion and politicians' office motivation is robust to the use of different measures for our key variables. Instead of politicians' motivation from the ISSP, we also use confidence in political parties as contained in the third wave of the World Values Survey (WVS), which was concluded in 1998. Using this measure and adjusting the GDP and growth variables to the survey date, the interaction effect remains positive and significant (p=0.009, F=1141.31, N=10). Unfortunately, the set of countries covered both by the third wave of the WVS as well as by Lijphart is small. Other surveys on politicians' office motivation have been conducted only very recently and are thus not applicable within our research design.

The measure for power dispersion by Lijphart (1999) is available in a more current version from Armingeon et al. (2011). The use of this indicator yields a highly significant interaction term (p=0.009, F=9.95, N=17). Armingeon et al. (2011) also provide a modified index which focuses on institutional factors only. It is based on the variables "electoral disproportionality" and "number of parties" and is invariant to behavioral factors such as "absence of minimal winning coalitions" included in the original index. Using this measure instead, the results remain significant (p=0.061, F=15.99, N=17). We also use three different measures which depict important aspects of power dispersion and find similar patterns. For the index for checks and balances (Keefer and Stasavage, 2003) and a plurality electoral system dummy (Beck et al., 2001), the interaction effect shows the expected sign ($p \leq 0.087, F \geq 8.91, N = 18$). For the nine-categorial type of electoral system (IDEA, 2004), however, the coefficient is insignificant ($p = 0.159, F = 8.53, N = 18$).

One might fear that our result is influenced by the Financial crisis which affected output beginning in 2008. To ensure that the Financial crisis does not drive patterns in the data, we may restrict explained GDP growth to the years 2004-2007. For this shorter period the interaction term between power dispersion and office motivation remains weakly significant ($p=0.092$, $F=4.20$, $N=18$). Using the World Values Survey for our measure of politicians' office motivation we can expand explained GDP growth to the years 1998-2007. This data provides a similar picture ($p=0.062$, $F=127.54$, $N=10$). An alternative approach to deal with the financial crisis is to exclude countries that were particularly affected. The result is robust to the exclusion of any one country from the analysis ($p \leq 0.081$) and to the exclusion of any subset of the countries Ireland, Spain, and Portugal ($p \leq 0.075$), which were affected disproportionately.

Finally, the empirical observation in this section could have a different explanation. This may be the case if our measure for politicians' office motivation captures politicians' preferences not only for office but also for rent extraction. Our measure is in fact likely to represent a mix of both motives. However, it is not evident that institutions interact with rent seeking as predicted by our model. In particular, models that allow the prevalence of rent seeking to vary among politicians (Smart and Sturm, 2006; Besley and Smart, 2007) deliver ambiguous results.

4.7. Modifications and extensions

Our model is flexible enough to incorporate several modifications and extensions. In a first step, we extend the analysis to non-informative equilibria. In a second step, we generalize the setting in three dimensions. First, we introduce a continuous policy space such that candidates may choose the magnitude of reform they propose rather than limiting their choice set to a complete reform and the status quo. Second, we show that limited commitment, i. e., the possibility to withdraw a proposal after the election with a certain probability, does not change the results. Third, we allow for heterogeneous policy preferences of voters in the sense that proposals of candidates may be valued by some voters and disliked by others. None of these modifications alter the qualitative results of the model (all proofs are found in Appendix A.3.2). We end the section with a discussion of the desirability of democratic selection where we contrast a result of Maskin and Tirole (2004) with our findings.

4.7.1. Non-informative equilibria

In the main part of the paper, we focus on the effects of power dispersion in informative equilibria. However, changes in power concentration may also allow to move from a non-informative to an informative equilibrium, thereby increasing welfare.

Proposition 4.3 establishes a dependence between the level of office motivation θ and the characteristics of the unique divine equilibrium. This equilibrium is informative if and only if θ is below the threshold $\bar{\theta}(\rho)$. If instead $\theta \geq \bar{\theta}(\rho)$, the unique divine equilibrium is non-informative: the reform incentive function is larger than zero for all types of candidates. Even the least able candidate proposes a reform, although it will fail with certainty ($\alpha = 0$). However, $\bar{\theta}(\rho)$ is strictly decreasing in ρ. Thus, it is possible to move from non-informative equilibria to informative ones by implementing political institutions that induce more power dispersion.

Proposition 4.6. *For any level of θ, there exists $\bar{\rho}(\theta) > 0$ such that the unique divine equilibrium is informative if and only if power is sufficiently dispersed, $\rho < \bar{\rho}(\theta)$. All informative equilibria strictly welfare-dominate non-informative equilibria.*

For the last statement, note that in any informative equilibrium low ability candidates with $a_i \in [0, \alpha)$ refrain from proposing inefficient reforms. Hence, these candidates receive less power, while more power is allocated to welfare enhancing candidates with $a_i \geq \alpha$, who provide a positive welfare contribution.

4.7.2. Continuous policy space

Until now, we have assumed that candidates can either propose a reform or the status quo. However, many policy decisions are inherently continuous and politicians can choose "how much" of a reform they propose to be implemented. Suppose that the action space of the candidates is $x_i \in [0, 1]$, with $x_i = 0$ being the status quo and $x_i = 1$ representing a complete reform. As before, the welfare contribution of candidate i is given by $x_i(a_i - c)$.

Given the continuous policy space, it is possible to interpret the realized outcome as a compromise between the candidates' agendas instead of a lottery between the proposals. Then, candidates with larger amount of power $f(v_i, \rho)$ are able to enforce larger parts of their agenda, while candidates with less power only slightly influence the political decision.

Proposition 4.7. *Let the action space of candidates be continuous with $x_i \in [0, 1]$. There is a unique divine equilibrium, which is outcome equivalent to the one resulting for a binary action space. As a consequence, Propositions 1 to 5 hold.*

Since reform incentives are still monotonically increasing in the ability of the candidates, so is the magnitude of reform they propose. The divinity criterion yields that only complete reforms or the status quo are played in equilibrium. To see this, first note that complete reforms are always proposed in a divine equilibrium. If this was not the case, the divinity criterion would require that a deviation to a complete reform would have to be attributed to the most able candidate. Given these beliefs, a complete reform would yield a profitable deviation for the most able candidates. As a consequence, all agents with ability above c choose to propose a complete reform instead of proposing only a share of a reform, since they profit from the higher welfare contribution as well as from the higher office utility generated by this proposal. Thus, voters associate a negative welfare contribution with any intermediate reform proposal. Hence, such a proposal leads to a smaller vote share than the status quo. For candidates below c, an intermediate reform proposal also leads to a smaller welfare contribution than the status quo. Overall, only the two extremes of the action space are played in a divine equilibrium.

Consequently, restricting the action space of candidates to the status quo policy or a complete reform does not impose a loss of generality. The unique divine equilibrium is outcome equivalent in the sense that strategies are equivalent and only out-of-equilibrium beliefs may distinguish the equilibria. All proofs carry over to this setup.

4.7.3. Limited commitment

The assumption of full commitment is widely used to ensure tractability of models (see, e.g., Persson and Tabellini 2003). However, it may seem too restrictive that politicians can never change or adapt their agenda. In our setting, candidates with ability lower than c have an incentive to withdraw a reform proposal when they gain power. A straightforward way to introduce limited commitment into the model is to assume that, with probability $\lambda > 0$, the environment changes after the election and politicians may deviate from their proposal. This could be due to an unexpected shock in the policy field, a major event in another policy field, etc. With probability $1 - \lambda$, on the contrary, they have to carry out their proposal.

Proposition 4.8. *Suppose policy proposals are binding with probability λ. Then Propositions 1 to 5 continue to hold.*

This form of limited commitment does not change the main results of the paper. In essence, it increases incentives to propose a reform for low ability candidates, since they may be able to withdraw their proposal after the election. However, this only affects the level of equilibrium cutoffs and not the qualitative results.

The welfare effect of reduced commitment is ambiguous. On the one hand, all candidates with ability $a_i < c$ withdraw their reforms with probability λ, thereby increasing welfare. On the other hand, as limited commitment diminishes the negative welfare contribution of a reform proposal for low ability candidates, more inefficient reforms are proposed. Thus, reform proposals become less informative to the voters, and high-ability candidates receive less political power. The worse selection of politicians as well as the more inefficient reform proposals per se represent a negative effect on welfare.

4.7.4. Heterogeneous preferences

In political philosophy as well as public debate, a major virtue of power dispersion is seen in the political representation of minorities and the prevention of a tyranny of the majority. For example, James Madison argues in the federalist papers #51 that "the rights of the minority will be insecure" without proper checks and balances (Madison, 1788a). So far, our analysis has abstracted from this aspect of political institutions in order to emphasize effects of power dispersion that are independent of minority rights.

To incorporate heterogeneity in voters' policy preferences into our model, we may assume that voters differ in their valuation of a reform rather than in their candidate preferences.[14] In particular, voter k receives a payoff of μ_k if a reform is successfully implemented. Let the preference parameter μ_k be symmetrically distributed according to the pdf $\xi(\mu)$ and the cdf $\Xi(\mu)$ with full support on some interval $\left[\underline{\mu}, \bar{\mu}\right]$. We assume that the mean preference is larger than the reform cost c, while $\underline{\mu} \in (0, c)$. This implies that a majority of voters is in favor of the reform as long as it is adopted by a sufficiently able candidate, while a minority strictly prefers the status quo.

Proposition 4.9. *If the voters have heterogeneous policy preferences according to distribution $\Xi(\mu)$, Propositions 1 to 5 continue to hold.*

Essentially, the proofs for all previous results hold whenever the expected vote share of a reforming candidate i is increasing in the average ability of candidates that propose a reform, i.e., in the equilibrium cutoff α_i. The basic model can be seen as the special case with a degenerate distribution function with $\mu_k = 1$ for all voters.

Given these heterogeneous policy preferences, our model allows to reconsider Madison's conjecture. Increasing power dispersion leads to higher amounts of power for candidates proposing the status quo, which is the minority's preferred option. As a consequence, the status quo is proposed more often yielding an additional increase in the minority's welfare.

[14] Our model also allows for additional (ideological) heterogeneity. Let the reforms advocated by both candidates be targeted towards a different group of voters and let μ_{ki} denote the payoff to voter k from a successful reform by candidate i. If both parameters share the unconditional distribution $\Xi(\mu)$ defined above, Proposition 4.9 continues to hold for any correlation between μ_{k1} and μ_{k2}. With negative correlation, the candidates' reform proposals differ strongly or are even diametrically opposed (as in a stylized left-right policy space).

Lemma 4.3. *In any informative equilibrium, the utility of each minority voter k with $\mu_k \leq c$ is strictly decreasing in the concentration of political power.*

The quote above suggests that the Founding Fathers of the United States were interested in the protection of minority rights per se. Formally, this objective can be captured by introducing inequality aversion into the welfare function, using a strictly increasing, strictly concave and twice continuously differentiable weighting function w:

$$W_{IA} = \int_{\underline{\mu}}^{\bar{\mu}} w(V(\mu_k, \rho))\xi(\mu_k)d\mu_k.$$

In this function, $V(\mu_k, \rho)$ represents the expected utility of a voter with preference μ_k. Following Atkinson (1973) and Hellwig (2005), the relative curvature of w can be referred to as a measure of inequality aversion. Compared to the inequality-neutral welfare function, W_{IA} puts higher weights on voters with low expected utility.

Proposition 4.10. *Any welfare function W_{IA} with inequality aversion is maximized at a lower level of power concentration than the inequality-neutral function W.*

Intuitively, power-dispersing institutions reduce the discretion of the election winner, who is chosen by the majority. The expected utility of the majority of voters is hence reduced while the minority is better off. The utility of the minority is valued strongly by an inequality averse constitutional designer. Thus, he will choose to disperse power more strongly than if he was inequality-neutral.[15]

4.7.5. Desirability of democratic selection

Our model also allows to investigate whether democratic selection leads to higher welfare than a non-democratic system in which one agent is drawn randomly and receives unlimited political power. A similar question is addressed by Maskin and Tirole (2004), who investigate whether decision-making power should be allocated to accountable "politicians" or nonaccountable "judges". While a non-democratic system in our model does not allow for sorting of candidates, it does not provide the decision-maker with incentives to choose inefficient policies, either. Clearly, this is an extremely positive view on non-democratic systems, which reflects the lack of agency problems in our model, once a candidate is in office.

In our framework, a non-democratic system is mathematically equivalent to the limit case of complete dispersion of power. With $\rho = 0$, the allocation of power is independent

[15] Note that W_{IA} is maximized at a strictly lower level than W for any $\theta < \bar{\theta}$. For the opposite case, even constitutional designers with small degrees of inequality aversion will prefer to concentrate power completely.

of the election result and each candidate sets policy with probability one half. Thus, it follows directly from Proposition 4.5 that political selection is always desirable.

Corollary 4.1. *For any $\theta < \infty$, democratic elections with optimally chosen power concentration ($\rho = \rho^*$) provide strictly higher welfare than non-democratic systems with random assignment of political power.*

Note that this contrasts to the result of Maskin and Tirole (2004) who find that, under certain circumstances, political decisions should rather be delegated to "judges" than to "politicians". In contrast to our setup, accountability in their model refers to whether the public official is subject to reelection after taking a political decision. More importantly, however, Maskin and Tirole only consider the polar case of politicians with fully concentrated political power, while we allow for continuous changes in the degree of power concentration.[16]

4.8. Conclusion

We have investigated effects of variations in the level of power concentration on the behavior of politicians in political campaigns and the implied welfare changes if candidates are office-motivated and privately informed about their ability. Increasing the concentration of power causes two effects. On the one hand, it has a positive *empowerment effect* because more power is given to election winners, who provide higher welfare in expectation. On the other hand, it also has a negative *behavioral effect*. With a stronger concentration of political power, low-ability candidates have a stronger incentive to mimic more able ones by committing to risky reforms. This limits the voters' capacities to select high ability politicians. The optimal level of power concentration balances both effects.

Furthermore, we have identified a negative relation between the extent of office motivation and the optimal level of power concentration. If politicians care mainly about welfare, power concentration yields strictly positive effects. In the case of strong office motivation, on the contrary, welfare is maximized by institutions that divide power between election winner and loser.

In the empirical part, we have confronted these findings with data for eighteen established democracies. Our findings are in line with the theoretical predictions: There is a significant positive interaction effect between office motivation and power dispersion. For the lowest levels of office motivation, power-dispersing institutions come along with

[16] If one restricts attention to the two polar cases of fully concentrated and completely dispersed power, however, one of the main results of Maskin and Tirole (2004) can be replicated in our model. Then, non-democratic systems perform better if candidates are mainly office-motivated.

significantly lower economic growth, while we find a positive correlation for countries with higher levels of office motivation.

A. Appendices

A.1. Appendix to Chapter I

A.1.1. Proofs of propositions and lemmas

Proof of Lemma 2.1: Suppose strategy profile s with strategies s_A^x, s_P constitutes a Bayesian equilibrium. Denote by D_1 the set of all defaults being played in equilibrium and by D_2 the nonempty set of out of equilibrium defaults. Take any $d_i \in D_1$ and $d_j \in D_2$.

$$s_P : \rho_l \to (.., p(d_i) = \alpha_1, .., p(d_j) = 0, ..) \, , \rho_h \to (.., p(d_i) = \alpha_2, .., p(d_j) = 0, ..)$$

Define a new strategy of the principal:

$$\hat{s_P} : \rho_l \to \left(.., p(d_i) = \frac{\alpha_1}{2}, .., p(d_j) = \frac{\alpha_1}{2}, ..\right) , \rho_h \to \left(.., p(d_i) = \frac{\alpha_2}{2}, .., p(d_j) = \frac{\alpha_2}{2}, ..\right)$$

Moreover, construct new strategies $\hat{s_A^x}$ for the agents such that $\hat{s_A^x}(d_j, \sigma) = \hat{s_A^x}(d_i, \sigma) = s_A^x(d_i, \sigma)$. Leaving everything else fixed $\hat{s_A^x}, \hat{s_P}$ also constitute a Bayesian equilibrium. First note that $p(\theta|d_i, \sigma) = p(\theta|d_j, \sigma)$ for both θ. According to Bayes' rule these are also equivalent to the conditional probabilities following d_i in the original equilibrium. Hence, agents' maximization problem is identical and the best responses do not change. Given their strategy the $\hat{s_P}$ must also be a best response, since messages d_i, d_j induce the same action and the principal is indifferent between them. By subsequently adding all out of equilibrium defaults to the set of equilibrium defaults we can construct an output equivalent equilibrium without any out of equilibrium defaults.

Proof of Lemma 2.2: Since we consider informative equilibria, there are at least two different defaults in the sense that $p(\theta_h|d_k) > p(\theta_h|d_f)$. Given the signal ρ_h, we show that it is always more profitable to play d_k for the principal. The incentive function of the principal i.e the expected payoff difference from d_k versus d_f is:

$$\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(s_A^x(\sigma, d_k))|\rho_h, x]df(x) - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(s_A^x(\sigma, d_f))|\rho_h, x]df(x) =$$

$$\int_{\frac{1}{2}}^{1} p(\sigma_l, \theta_h|\rho_h, x)(\mu(U_A(\theta_h, s_A^x(\sigma_l, d_k)) - U_A(\theta_h, s_A^x(\sigma_l, d_f))) + b(s_A^x(\sigma_l, d_k)) - b(s_A^x(\sigma_l, d_f)))$$
$$+ p(\sigma_h, \theta_l|\rho_h, x)(\mu(U_A(\theta_l, s_A^x(\sigma_h, d_k)) - U_A(\theta_l, s_A^x(\sigma_h, d_f))) + b(s_A^x(\sigma_h, d_k)) - b(s_A^x(\sigma_h, d_f)))$$
$$+ p(\sigma_h, \theta_h|\rho_h, x)(\mu(U_A(\theta_h, s_A^x(\sigma_h, d_k)) - U_A(\theta_h, s_A^x(\sigma_h, d_f))) + b(s_A^x(\sigma_h, d_k)) - b(s_A^x(\sigma_h, d_f)))$$
$$+ p(\sigma_l, \theta_l|\rho_h, x)(\mu(U_A(\theta_l, s_A^x(\sigma_l, d_k)) - U_A(\theta_l, s_A^x(\sigma_l, d_f))) + b(s_A^x(\sigma_l, d_k)) - b(s_A^x(\sigma_l, d_f)))df(x).$$

The differences of the selfish parts $b(z)$ of the utility function are clearly larger or equal to zero, since the agents' best responses to d_k are weakly higher zs. We can, thus, concentrate on the utility difference of the agents. The integral of these is definitely larger than zero if this is fulfilled for all x. Consider the difference for an arbitrary x with slightly rewritten conditional probabilities:

$$p(\sigma_l|\rho_h, x)p(\theta_l|\sigma_l, \rho_h, x)(U_A(\theta_l, s_A^x(\sigma_l, d_k)) - U_A(\theta_l, s_A^x(\sigma_l, d_f)))$$
$$+ p(\sigma_h|\rho_h, x)p(\theta_l|\sigma_h, \rho_h, x)(U_A(\theta_l, s_A^x(\sigma_h, d_k)) - U_A(\theta_l, s_A^x(\sigma_h, d_f)))$$
$$+ p(\sigma_l|\rho_h, x)p(\theta_h|\sigma_l, \rho_h, x)(U_A(\theta_h, s_A^x(\sigma_l, d_k)) - U_A(\theta_h, s_A^x(\sigma_l, d_f)))$$
$$+ p(\sigma_h|\rho_h, x)p(\theta_h|\sigma_h, \rho_h, x)(U_A(\theta_h, s_A^x(\sigma_h, d_k)) - U_A(\theta_h, s_A^x(\sigma_h, d_f))).$$

The principal's expectation about the difference in payoffs for the agent is larger than the agents own prediction, which in turn is given by

$$p(\sigma_l|\rho_h, x)p(\theta_l|\sigma_l, d_h, x)(U_A(\theta_l, s_A^x(\sigma_l, d_k)) - U_A(\theta_l, s_A^x(\sigma_l, d_f)))$$
$$+ p(\sigma_h|\rho_h, x)p(\theta_l|\sigma_h, d_h, x)(U_A(\theta_l, s_A^x(\sigma_h, d_k)) - U_A(\theta_l, s_A^x(\sigma_h, d_f)))$$
$$+ p(\sigma_l|\rho_h, x)p(\theta_h|\sigma_l, d_h, x)(U_A(\theta_h, s_A^x(\sigma_l, d_k)) - U_A(\theta_h, s_A^x(\sigma_l, d_f)))$$
$$+ p(\sigma_h|\rho_h, x)p(\theta_h|\sigma_h, d_h, x)(U_A(\theta_h, s_A^x(\sigma_h, d_k)) - U_A(\theta_h, s_A^x(\sigma_h, d_f))).$$

This is clearly larger or equal to zero, because agents maximize their expected utility. The integral can only attain zero, if the actions following d_k and d_f are equal for all agents. Since there is a continuum of agents with full support over the interval $[\frac{1}{2}, 1]$, this can never be the case. We conclude that default d_f is played with probability zero whenever the principal receives a high signal. Hence, every default played in response to a high signal inhibits the same information. All other defaults reveal that the principal received a low signal.

Proof of Proposition 2.1: For $\mu \to \infty$, the preferences of the principal and the agents are fully aligned. Hence, full information is the only possible equilibrium if there is information transferred. Since preferences are continuous in μ, there exists a $\bar{\mu}$ such that for all $\mu \geq \bar{\mu}$ full information transmission is the only informative equilibrium. Furthermore, this equilibrium clearly Pareto-dominates the babbling equilibrium, which always exists. If μ is equal to zero, however, the principal always chooses the default inducing the highest z. As a consequence, defaults cannot transfer any information. With the same argument as above there exists a $\underline{\mu}$ such that for all $\mu \leq \underline{\mu}$ defaults inhibit no informational content in equilibrium.

Proof of Proposition 2.2: In the following we show that c^{pd} is increasing in μ along the path of Pareto-efficient equilibria. To simplify notation, we write D_i for any default from the set D_i. Note first that $\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_l))|\rho_h, x]df(x)$ is constant in c^{pd} as explained in Lemma 2.3. In any mixed equilibrium, the principal is indifferent between all defaults given she received a low signal

$$\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x) - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_l))|\rho_l, x]df(x) = 0.$$

The change of c^{pd} with increasing μ can be derived by implicit differentiation of this equilibrium condition.

$$\frac{dc^{pd}}{d\mu} = \\ -\frac{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x) - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_l))|\rho_l, x]df(x)\right)/\partial\mu}{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x) - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_l))|\rho_l, x]df(x)\right)/\partial c^{pd}} = \\ -\frac{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x) - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_l))|\rho_l, x]df(x)\right)/\partial\mu}{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x)\right)/\partial c^{pd}}$$

The numerator of this fraction is smaller than zero, if

$$\begin{aligned}\int_{\frac{1}{2}}^{1} & p(\sigma_l, \theta_l|\rho_l, x)(U_A(\theta_l, s_A^x(\sigma_l, D_h)) - U_A(\theta_l, s_A^x(\sigma_l, D_l)))) \\ &+ p(\sigma_h, \theta_l|\rho_l, x)(U_A(\theta_l, s_A^x(\sigma_h, D_h)) - U_A(\theta_l, s_A^x(\sigma_h, D_l))) \\ &+ p(\sigma_l, \theta_h|\rho_l, x)(U_A(\theta_h, s_A^x(\sigma_l, D_h)) - U_A(\theta_h, s_A^x(\sigma_l, D_l))) \\ &+ p(\sigma_h, \theta_h|\rho_l, x)(U_A(\theta_h, s_A^x(\sigma_h, D_h)) - U_A(\theta_h, s_A^x(\sigma_h, D_l)))df(x) \leq 0\end{aligned}$$

This follows, if the inequality holds for all x. Since defaults convey less information than the signal of the principal and since agents maximize their utility, this holds, because

$$\begin{aligned}& p(\sigma_l|\rho_l, x)p(\theta_l|\sigma_l, D_l, x)(U_A(\theta_l, s_A^x(\sigma_l, D_h)) - U_A(\theta_l, s_A^x(\sigma_l, D_l)))) \\ &+ p(\sigma_h|\rho_l, x)p(\theta_l|\sigma_h, D_l, x)(U_A(\theta_l, s_A^x(\sigma_h, D_h)) - U_A(\theta_l, s_A^x(\sigma_h, D_l))) \\ &+ p(\sigma_l|\rho_l, x)p(\theta_h|\sigma_l, D_l, x)(U_A(\theta_h, s_A^x(\sigma_l, D_h)) - U_A(\theta_h, s_A^x(\sigma_l, D_l))) \\ &+ p(\sigma_h, \rho_l, x)p(\theta_h|\sigma_h, D_l, x)(U_A(\theta_h, s_A^x(\sigma_h, D_h)) - U_A(\theta_h, s_A^x(\sigma_h, D_l))) \leq 0.\end{aligned}$$

Consequently, the numerator is negative, yielding

$$\frac{dc^{pd}}{d\mu} > 0 \Leftrightarrow \frac{\partial \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, D_h))|\rho_l, x]df(x)}{\partial c^{pd}} > 0.$$

Lemma 2 states that the equilibrium with the highest information transmission rate is Pareto-efficient. Suppose that the latter derivative is negative. As a consequence, the payoff from a low default is larger than the one from a high default for all transmission rates larger than c^{pd}. In particular, this also holds for $c = 1$, implying that there exists an equilibrium with full information transmission. This is a contradiction to the assumption that we are in a mixed Pareto-efficient equilibrium. There can, thus, only exist mixed Pareto-efficient equilibria with positive derivative and $\frac{dc^{pd}}{d\mu} \geq 0$. Since agents are Bayesian updater, this leads to weakly stronger responses by the agents, i.e, $s_A^x(\sigma_i, D_h)$ is increasing and $s_A^x(\sigma_i, D_l)$ is decreasing in μ.

Proof of Proposition 2.3: We assumed that the partial derivative with respect to z of $U_A(z, \theta_l)$ is smaller than the corresponding derivative of $U_A(z, \theta_h)$ for all z. Hence, higher zs are optimal if the agent puts more probability weight on θ_h. Since agents are Bayesian updater, they weigh their private signal stronger, if it conveys more information. Hence, for any given equilibrium, agents are more susceptible to the default, if they have less private information.

A.1.2. Additional figures and tables

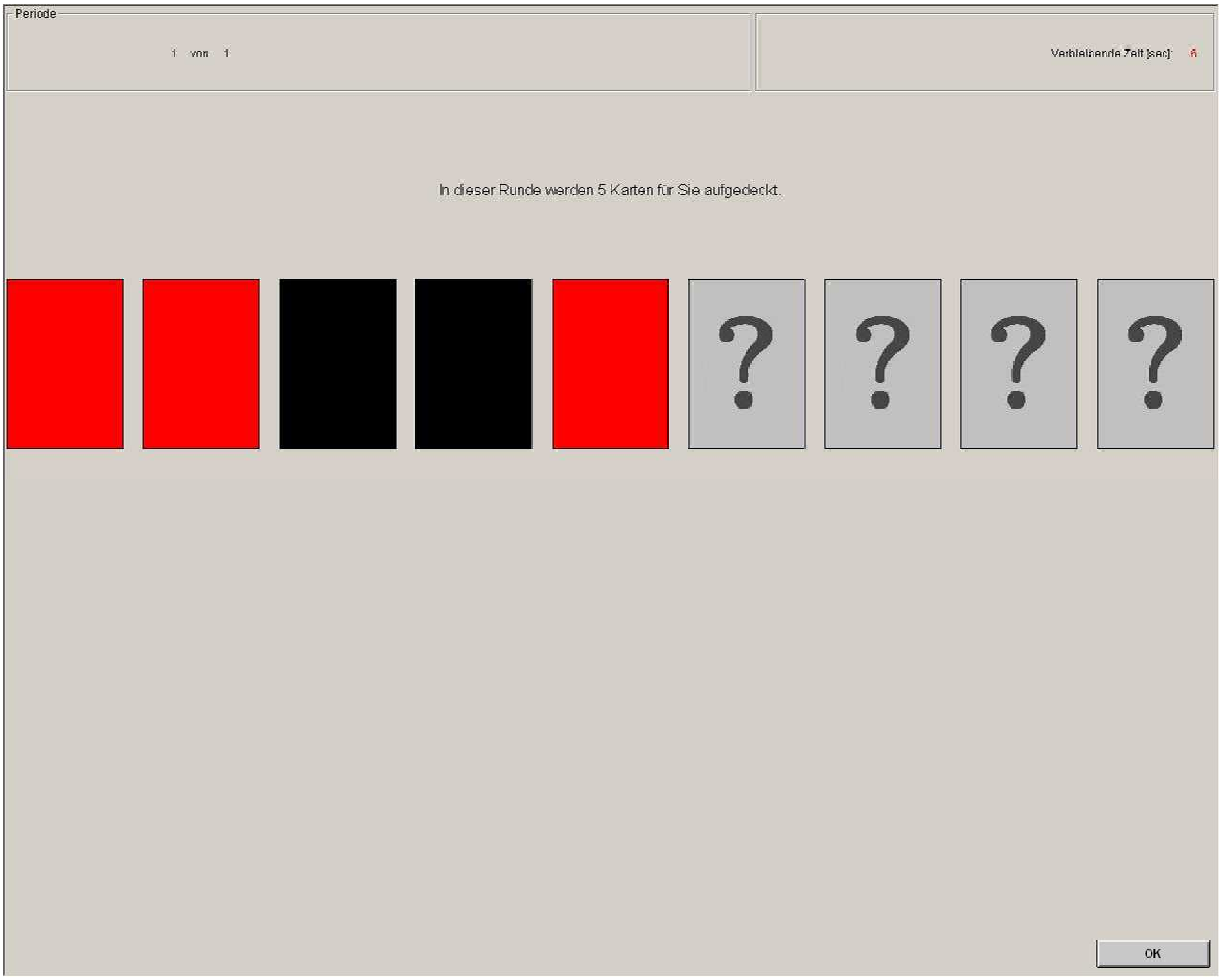


Figure A.1.: *Agent's information screen. Example in which two black and three red cards are revealed.*

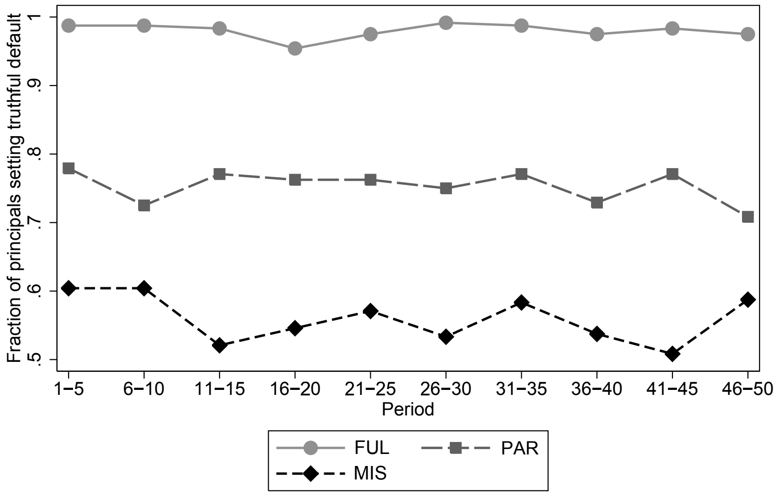


Figure A.2.: *Frequency with which principals set default according to their signal. Average values for 5-period intervals in FUL, PAR, and MIS.*

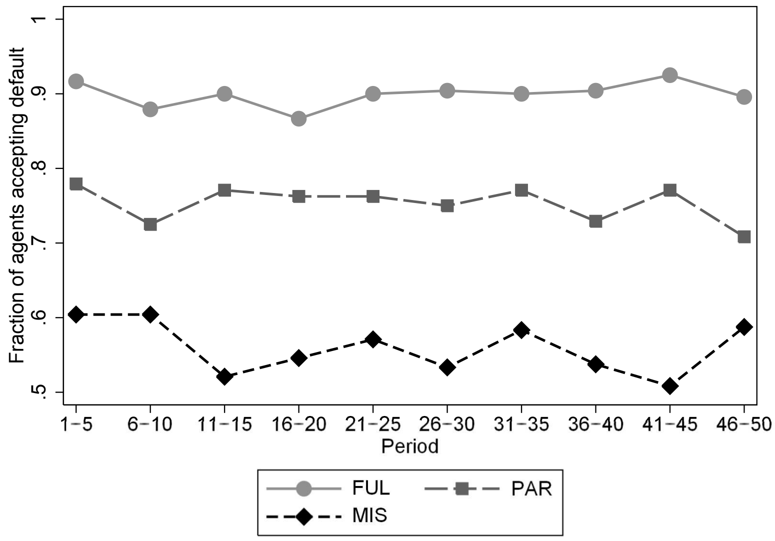


Figure A.3.: *Frequency with which agents follow default option. Average values for 5-period intervals in FUL, PAR, and MIS.*

Dependent variable: Frequency with which principal chooses black default after black signal (MIS treatment)			
	(1)	(2)	(3)
HHS	.341**		.314*
	(.121)		(.143)
CRT		-.859	-.381
		(3.221)	(3.290)
Math grade		2.926	2.621
		(3.354)	(3.751)
Const.	12.573***	6.870	6.927
	(1.273)	(11.294)	(13.021)
N	48	48	48
R^2	.098	.502	.134

Table A.1.: *Determinants of principals' behavior in MIS. OLS estimations. The reported standard errors (in parentheses) account for potential clustering on the matching-group level.*

A.1.3. Instructions of the experiment (PAR Treatment)

Welcome to today's decision experiment!

Please read the following information carefully. Everything that you need to know to participate in this experiment is explained below. Should you have any difficulties in understanding these instructions please raise your hand. We will answer your questions at your cubicle.

For your arrival on time you receive an initial endowment of 4 euros. During the experiment, you can earn further money by earning points. The amount of points that you earn during the experiment depends on your decisions and the decisions of other participants. All points that you earn during the experiment will be added up and converted into euros at the end of the experiment. The exchange rate is:

100 Points = 1 euro

The experiment consists of several periods. In each period you have to make decisions, which you feed into the computer. There will be 50 periods in total. At the end of the experiment, the amount of money that you earned during the experiment and your initial endowment will be paid out in cash. Please hand in all documents you received from us when you collect your payment at the end of the experiment.

Please note that communication between participants is strictly prohibited during the entire experiment. In addition, we would like to point out that you may only use the computer functions which are required for the experiment. Violations of these rules will lead to an exclusion from the experiment. In case you have any questions we shall be glad to assist you.

Short Overview of the Experimental Procedures

The experiment consists of several periods. In each period you have to make decisions which you feed into the computer. **There will be 50 periods in total**.

In the experiment participants will be divided into three groups. The participants of the three groups will be named player 1A, player 1B and player 2 during the entire experiment. There will be **6 participants of type 1A, 6 participants of type 1B and 12 participants of type 2**. Your role will be drawn randomly at the beginning of the experiment, and presented on your screen. Your role remains the same during the entire experiment.

In each period of the experiment groups consisting of one player 1A and one player 2 **or** one player 1B and one player 2 will be formed. For each period **new groups of two players will be formed randomly**. The probability that your group in a certain period consist of one player 1A and one player 2 is thus 50%. The probability that your group in a certain period consist of one player 1B and one player 2 is also 50%. Player 2 learns only at the very end of a given period **whether he was in a group with a player of type 1A or 1B in that period**. The experiment is conducted fully anonymously. This means that you and the other participants never get to know with whom you were matched during the experiment. Only the other group member in a given period will get to know your choices in that period. All of the other participants will learn nothing about your decisions.

In each period of the experiment player 2 has to make a decision about which color appears more often in a set of cards. **In each period there will be a set of 9 cards. Each of the 9 cards can be either red or black.** Player 1A or player 1B receives information about the number of red and black cards in the current set and **makes a preselection** for player 2. Player 2 also receives information about the number of red and black cards and makes the **final decision for the group**. He can either confirm the preselection of the other group member, or make a different decision.

In each period a **new set of cards will be generated randomly for each group**. The probability that a certain card of the set is red or black is 50%. **That is, for each card it is equal likely to be red or black.** Hence, in each period it will be independently and randomly drawn how many red or black cards appear and the order of the cards. At the beginning of a period the probability that there are more red or more black cards is thus the same. The number of red and black cards does not depend on the cards of the previous period or on the cards of another group.

The final decision of player 2 determines the earnings of both members of the group:

- **Player 1A** earns 50 points if the final decision is ”more red cards”. He earns 0 points if the final decision is ”more black cards”.
- **Player 1B** earns 50 points if player 2 chooses the color that appears more frequently in the actual set of cards. He earns 0 points if player 2 chooses the wrong color.
- **Player 2** earns 50 points if his decision is the color that appears more frequently in the actual set of cards. He earns 0 points if the final decision is the wrong color.

The Experimental Procedures in Detail

Each period consists of two stages. On the first stage a new set of cards will be generated for each group. Afterwards the players get incomplete information about which color appears more often in their current set of cards. On the second stage, player 1A or player 1B make a preselection on the color of the cards. Player 2 gets to know the preselection and makes the final decision for the group. He can confirm the preselection or make a different decision. The decision of player 2 is the final decision for the entire group. That is, only the decision of player 2 determines the earnings of both group members.

Player 1 (that is, player 1A or 1B) receives his information about the amount of red and black cards in the current set of cards via a signal on his screen. There are two possible **signals which can be displayed on the screen**: either the signal for player 1 is ”more red cards” or the signal is ”more black cards”. In each group only player 1 gets this signal.

The probability of a correct signal for player 1 is 80%. That is, player 1 gets to know the color that actually appears more often in the current card set of his group in on average 80 out of 100 cases. On average he gets a wrong signal about the color that actually appears more often in the set of cards in 20 out of 100 cases. Only at the end of a given period, player 1 can determine whether he got a correct or incorrect signal.

Whether a signal is correct or incorrect in a certain period will be randomly and independently determined in each period. That is, the probability that player1 gets the correct signal is always exactly 80% and independent of the correctness of her signals in previous periods. The probability of a correct signal is also independent of the information player 2 or other players of type 1A or 1B receive in a certain period. None of the other participants ever gets to know which signal player 1 received in a certain period.

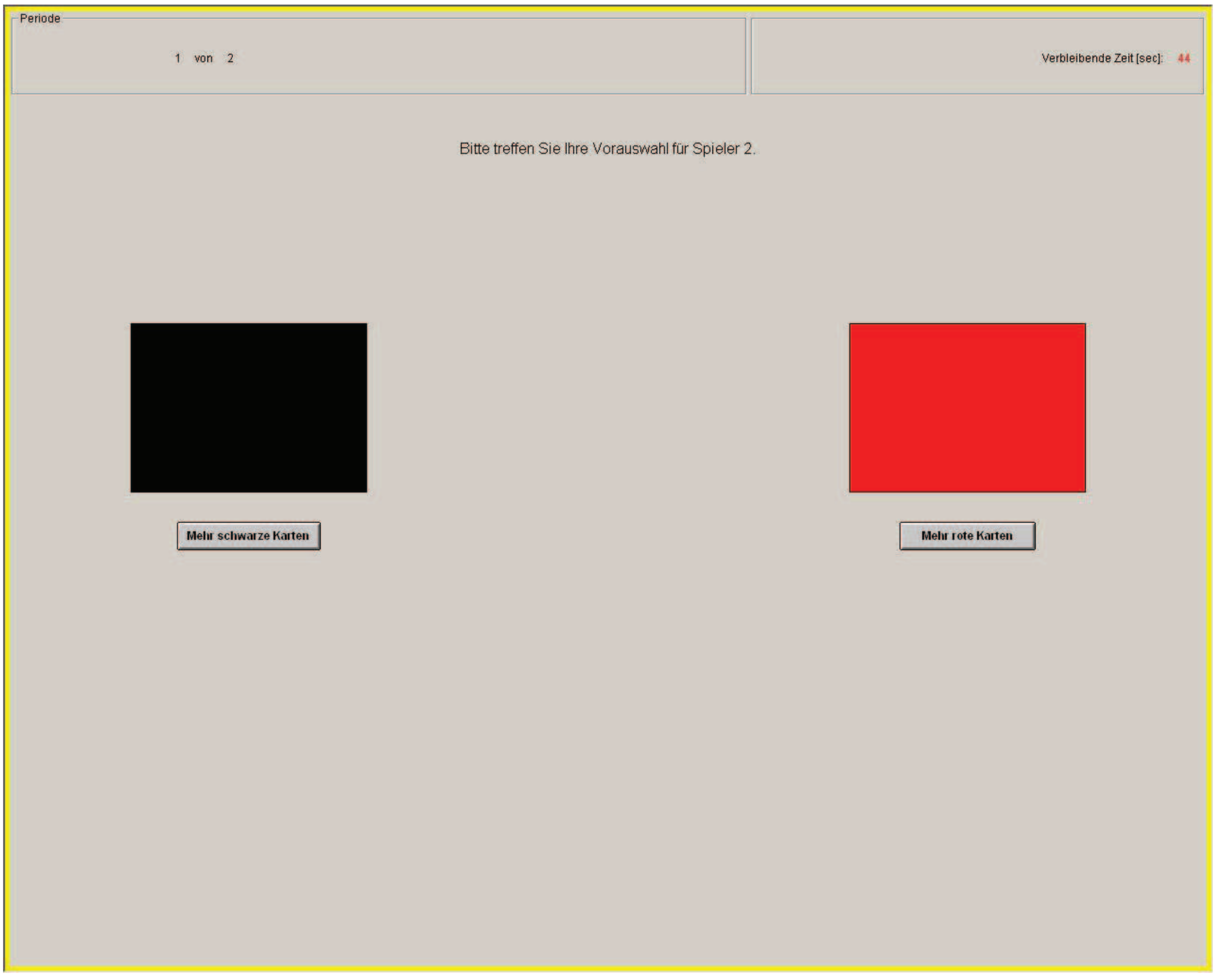
Player 2 receives his information about the number of red and black cards **by the uncovering of a part of the set of 9 cards.** In each period either **two** or **five** cards will be uncovered for player 2. The probability that two cards will be uncovered in a certain period is 50%. The probability that five cards will be uncovered in a certain period is also 50%. That is, in **on average 50 out of 100 cases two cards will be uncovered, and in 50 out of 100 cases five cards will be uncovered**.

Whether player 2 gets to know two or five cards in a certain period will be drawn randomly and independently across periods. The number of cards a certain player 2 gets to know is also independent of the number of cards that other players of type 2 get to know. Finally, the draw whether player 2 gets displayed two or five cards is independent of the correctness of the signal of the other group member. All other participants never get to know how many cards were uncovered for player 2 in a certain period.

Once player 1 and player 2 have seen their information, the second stage of the period begins. **On the second stage player 1 (that is, player 1A or 1B) makes a preselection for player 2**. Player 1 can choose between two options for the preselection. She can either choose "more black cards", or "more red cards". To make her decision, the following screen is displayed to her. She chooses the preselection by clicking the respective button.

Afterwards, player 2 is informed about the preselection which the other group member made for him. The following screen is shown to player 2, on which the preselection of player 1 will be displayed (instead of the black boardered field in the middle of the screen, a red or black field will be displayed):

Subsequently, player 2 makes the final decision for the entire group. He can confirm the preselection of the other group member or change the selection. **If player 2 confirms the preselection**, this is the final decision for the group. **If player 2 decides to change the selection** he sees a new screen. On this screen, he can choose either "more black cards" or "more red cards". At the time of his decision, player 2 doesn't know whether he is in a group with a player of type 1A or 1B. **The decision of player 2 is the final decision for both group members**. The final decision of player 2 determines



the earnings for the entire group.

How are incomes calculated?

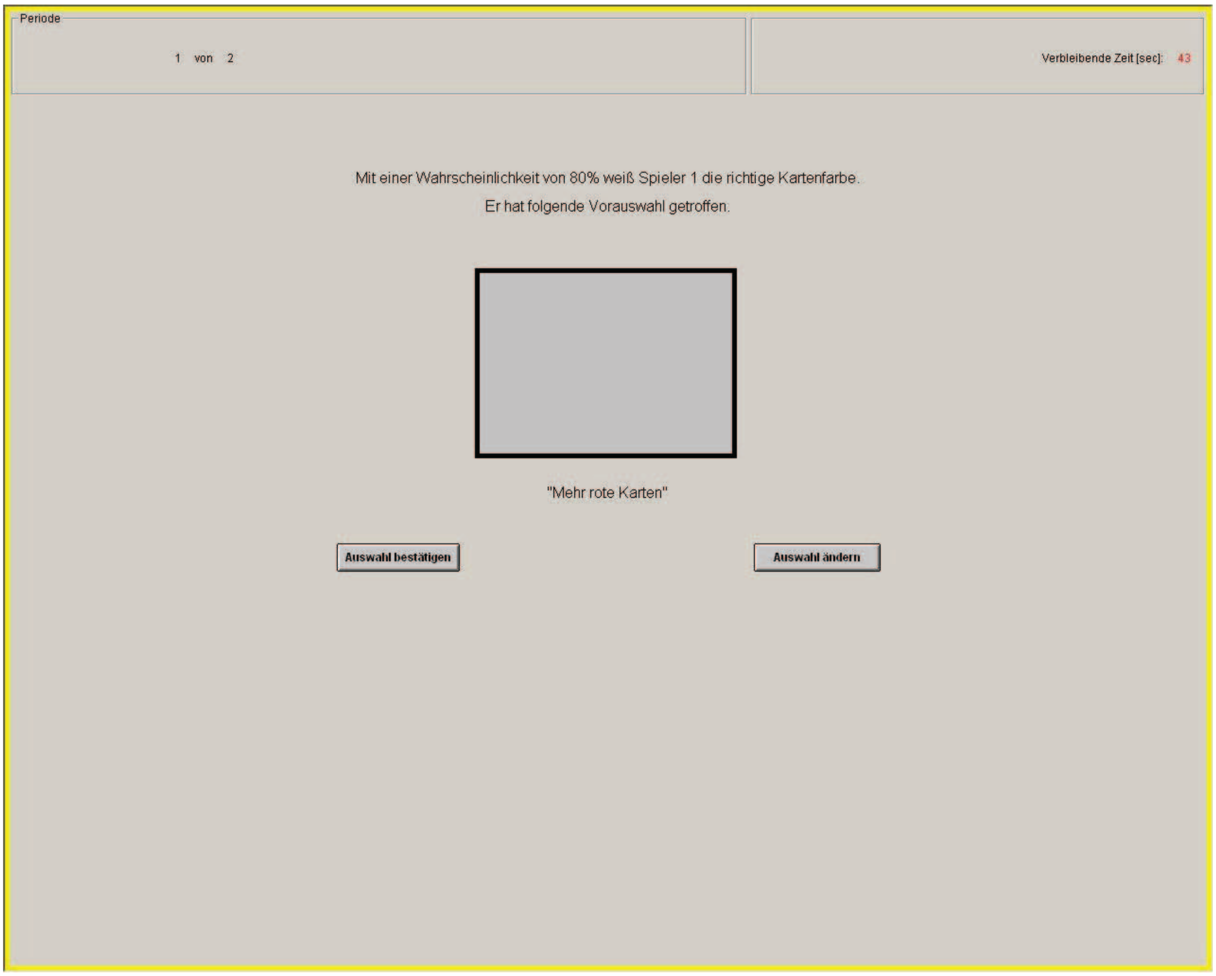
The earnings of player 1A, player 1B and player 2 in a certain period are calculated as follows.

Player 1A:

- When the final decision is **"more red cards"**, player 1A earns an income of **50 points**.
- When the final decision is **"more black cards"**, player 1A earns an income of **0 points**.

Player 1B:

- When the final decision is **equal to the color that actually appears more frequently in the set of cards**, player 1B earns an income of **50 points**.
- When the final decision is **equal to the color that actually appears less frequently in the set of cards**, player 1B earns an income of **0 points**.



Player 2:

- When the final decision is **equal to the color that actually appears more frequently in the set of cards**, player 2 earns an income of **50 points**.
- When the final decision is **equal to the color that actually appears less frequently in the set of cards**, player 2 earns an income of **0 points**.

At the end of a period, the complete set of cards from this period will be displayed to both group members. You are also informed about the earnings of both group members in this period. Subsequently a new period begins, for which new groups of two participants of type 1A or 1B and 2 will be formed randomly.

In case you have any questions, please raise your hand.

A.2. Appendix to Chapter II

A.2.1. Theory appendix

In this appendix, we provide a model that formalizes the intuition for how contractual incompleteness can cause involuntary unemployment and labor market segmentation.

Model Set-up: We denote by $N + U$ the mass of workers and by $\frac{N}{2}$ the mass of firms. Workers and firms interact for T periods in discrete time. All firms can offer up to two vacancies in a given period. Maximum employment is thus N, and excess supply of workers implies that the minimum level of unemployment is U.[1] In any period, firms offer contracts to either one or two workers. If effort is verifiable, the contract determines both the (upfront) wage and the effort level. If efforts are only observable but not third-party verifiable, workers have discretion over their level of effort. We assume that workers are homogeneous in their productivity and can choose between n different effort levels $e \in \{e_1, ..., e_n\}$, where $e_i < e_{i+1}$ and $c(e_i) = c_i$ denotes the increasing cost of effort.

There are two types of workers: a fraction λ of the population is fair (f) and a fraction of $1 - \lambda$ is selfish (s). The utility of an employed selfish worker in any period is the difference of the received wage and the effort cost:

$$u_s = w - c(e).$$

Fair workers, in contrast, face an additional psychological cost or benefit, $g(w - c(\hat{e}))$, if they fulfill the contract. Utility for fair workers is thus denoted as:

$$u_f = \begin{cases} w - c(e) + g(w - c(\hat{e})) & \text{if the worker fulfills the contract } (e = \hat{e}) \\ w - c(e) & \text{if the worker breaches the contract } (e < \hat{e}) \end{cases}$$

We assume that $g(\cdot)$ is strictly increasing in the rent offered by the firm, and $g(\tilde{x}) = 0$ for some $\tilde{x} \geq 0$. Thus, more generous contract terms make shirking increasingly less attractive for a fair worker, and $g(\cdot)$ changes from negative (a cost of contract fulfillment) to positive (an additional benefit of contract fulfillment), when offered rents exceed a benchmark level of fairness, $\tilde{x}$. This captures in a simple and tractable way the central theme of all reciprocity-based fairness models that fair types reward kind actions and punish unkind actions (e.g., Rabin 1993, Falk and Fischbacher 2006), where in our setting kindness of the firm is captured by the size of offered rents. Denote by $g^{-1}(\cdot)$ the inverse function of $g(\cdot)$, which exists and is well defined due to the monotonicity of $g(\cdot)$.

Firms are characterized by the following production technology, with output increasing in the level of effort:

$$0 \leq f(e_1) = z_1 \leq \cdots \leq f(e_n) = z_n, 0 \leq f(2e_1) = z_{n+1} \leq \cdots \leq f(2e_n) = z_{2n},$$

[1] While a level of $U > 0$ is implemented in the experiment, all results also hold for $U = 0$.

z_1 to z_n denote the output of a firm that employs one worker who exerts e_1 to e_n, and z_{n+1} to z_{2n} are the corresponding output levels of two-worker firms.[2] Furthermore, we require the production technology to exhibit a weak form of decreasing returns to scale (Part 1 of Assumption A.1) and to be efficient (Part 2 of Assumption A.1). The latter means that in the final period even in the case of incomplete contracts the wage needed to induce an extra unit of effort by a fair agent is smaller than the induced gain in output.

Assumption A.1. *Let $z_i < z_j$ be output levels with corresponding effort input $e_i < e_j$. Then,*

1. $\dfrac{z_j - z_i}{e_j - e_i} > \dfrac{z_{n+j} - z_{n+i}}{2(e_j - e_i)}$
2. $z_{n+j} - 2(g^{-1}(c_j) + c_j) > z_{n+i} - 2(g^{-1}(c_i) + c_i) > z_i - (g^{-1}(c_i) + c_i) > 0$

We further assume that output translates directly into firm revenue, all firms have access to the same production technology, and maximize total profits (i.e., revenue - overall wage costs). At the end of each period, firms decide whether to renew the contract with their worker(s). As a simplification we assume that a firm which renews the contract of its worker(s) also keeps its size constant. We denote by $b_k^t \in [0, 1]$ with $k \in \{1w, 2w\}$ the probability that a one-worker firm ($1w$) or two-worker firm ($2w$) separates from a worker if he exerts the desired effort level in period t. Purely for expositional reasons, we assume that firms always separate from workers who deviate from the contractually stipulated effort level (a_k^t=1).[3] Workers and firms discount the future at a rate $r \in [0, 1]$.

Equilibrium with verifiable effort: As a first step, we solve for the market equilibrium when effort is contractible. Consider the last period of the game. Selfish workers accept a contract offer if the wage at least covers the stipulated effort costs. Fair workers will accept a contract if

$$w - c(e) + g(w - c(\hat{e})) \geq 0.$$

Given our assumptions about $\tilde{x}$, firms need to pay (possibly low, but) positive monetary rents to induce fair workers to accept an offer. Since the production function is efficient, firms always offer a contract rather then staying out of the market. The optimal terms of the contract depend on the subjective probability of facing a fair worker, denoted by

[2]Note that we rule out the possibility that firms hire two workers and elicit different effort levels from each worker. This corresponds to assuming that a single firm does not offer both bad and good jobs. However, allowing for this possibility would not change the existence of the segmentation equilibrium that is our focus, although existence would be for a more restricted range of parameters.

[3]Endogenizing a_k^t would not change the characteristics of feasible equilibria. In particular, no-shirking conditions for separation rates $0 < a_k^t < 1$ can be derived analogously to conditions (A.1) and (A.3) below.

$\hat{\lambda}$. If, for instance, $\hat{\lambda} = 1$, a firm expects to face a fair worker with certainty and thus pays wages equal to the sum of effort and the psychological costs. The efficiency of the production technology implies that employment of two workers with maximal effort exertion is optimal in this case. Thus, there exists a cutoff $\tilde{\lambda}$ such that for all beliefs $\hat{\lambda} \geq \tilde{\lambda}$, firms use a homogeneous contracting strategy of hiring two workers, that involves paying wages to cover both types of costs, and workers exert the maximum effort level.

Maintaining $\hat{\lambda} > \bar{\lambda}$, in the second-to-last period backward induction implies that fair workers accept the same offer as before, since they do not expect any future utility rents. In contrast, selfish workers may anticipate a positive rent in the next period. Nevertheless, they do not accept any offer involving negative current rents, because this would reveal their type. If they did, the firm would have an incentive to renegotiate in the next period, and offer rents equivalent to the value of unemployment which is zero in the last period. However, selfish workers always accept the contract needed to employ fair types, which ensures positive current and future rents. For any stipulated level of effort, the rent needed to hire selfish workers is therefore weakly larger than in the last period, while the rent for fair types is the same. This makes it even less attractive than in the last period for firms to offer a contract that only attracts selfish workers, and thus firms offer the same contract terms as in the last period.

Assuming that there are enough fair types in the population such that $\lambda \geq \tilde{\lambda}$, this argument holds for all previous periods. This implies for all periods that (i) firms offer a wage-effort combination that induces maximal effort provision, (ii) firms always hire two workers and the level of unemployment is thus minimal, and (iii) all workers accept posted contracts.

Equilibria with non-verifiable effort: If effort is non-verifiable, the type of equilibrium depends on the shape of firms' production technology and the psychological cost function for fair workers. Since our main interest is to illustrate how contractual incompleteness can give rise to equilibria involving endogenous unemployment and market segmentation, we concentrate on the existence of these segmentation equilibria. As a first step we show that the endgame of the model features positive continuation values. Second, we characterize the no-shirking conditions for both worker types in pre-final periods. Third, assuming positive continuation values from period t, we pin down the conditions leading to a segmentation equilibrium in period $t-1$.

Endgame and continuation value: Selfish workers always choose to shirk in the final period. In contrast, fair workers fulfill their contract if their effort cost is smaller than their psychological return from fulfilling the contract. Anticipating this behavior of workers, firms either stay out of the market or offer wages such that fair workers are indifferent between working and shirking. The latter is profitable if there exists at least one effort level e_i such that either a one- or a two-worker firm expects positive payoffs from inducing e_i, given its belief $\hat{\lambda}$ that it faces a fair worker.

Assumption A.2.

$$\exists \quad e_i : \qquad \lambda z_i - (g^{-1}(c_i) + c_i) > 0 \quad or \quad \lambda^2 z_{n+i} + 2(1-\lambda)\lambda z_i - 2(g^{-1}(c_i) + c_i) > 0.$$

Assumption A.2 ensures that there are enough fair types such that firms are willing to offer contracts with positive wages in the one-shot version of the game.[4] This generates positive continuation values for workers in the pre-final period, if there has not been any screening in previous periods. Note that all firms use homogeneous contracting strategies that involve positive rents in the final period.[5] Hence, unemployment is involuntary from a worker's perspective. We denote the value of a job for a worker of type j who is employed by a firm of type k in period t by $V^t_{k,j}$. The value of unemployment in period for a worker of type j is denoted by $V^t_{u,j}$. We set all continuation values to zero in period $T+1$ and denote by L^t_{1w}, L^t_{2w} the number of jobs in one-worker and two-worker firms in period t.

No-shirking conditions for workers in pre-final periods: We start the analysis of pre-final periods by characterizing worker behavior. Workers trade off the short-run gains of low effort costs due to shirking against potential long-run costs due to higher risk of dismissal and unemployment. Let w^{t-1}_k and $\hat{e}^{t-1}_k$ denote the wage and desired effort level, offered by a firm of type k in period $t-1$. For a fair worker, the no-shirking condition in period $t-1$ is then:

$$w^{t-1}_k - c^{t-1}_k + g(w^{t-1}_k - c^{t-1}_k) + (1-r)\left[(1-b^{t-1}_k)V^t_{kf} + b^{t-1}_k V^t_{uf}\right] \geq w^{t-1}_k + (1-r)V^t_{uf} \tag{A.1}$$

$$\Rightarrow w^{t-1}_k \geq g^{-1}\left[c^{t-1}_k + (1-r)(b^{t-1}_k - 1)(V^t_{kf} - V^t_{uf})\right] + c^{t-1}_k. \tag{A.2}$$

A fair worker's utility in case of contract fulfillment (left hand side of equation (A.1)) consists of four components. The worker earns the current period's wage w^{t-1}_k, bears the

[4]Strictly speaking, fair types are not necessary for a final-period rent if workers can generate positive output without incurring effort costs (i.e., if $c_1 = 0$ and $z_1, z_{n+1} > 0$, as was the case in our experimental setting). In this case, firms could profitably offer minimal, but positive worker rents in the final period, which in turn opens up the possibility for "reputation equilibria", even when all workers are selfish. Empirically, final-period rents are substantially above this level, and many workers exert non-minimal effort in the final period; thus, an equilibrium based on fair types is better supported by the data (see our discussion in section 3.4.1).

[5]Since all firms are of the same type in the last period, we require that either firm type may rehire a worker from the pre-final period, using the final-period contract terms.

cost of effort c_k^{t-1}, experiences psychological utility $g(\cdot)$, and receives the continuation value conditional on contract fulfillment. In case of shirking (right hand side of (A.1)), the worker saves the effort cost and experiences no psychological utility. Furthermore, the current firm does not renew the worker's contract in the next period. Hence, the worker only receives the value of unemployment V_{uf}^t in the next period, which compromises the likelihood of finding a job of either type, and the likelihood of remaining unemployed in that period.

Selfish workers are not subject to the psychological cost, and since wages are paid before efforts are revealed, their effort choice is independent of the current period's wage. Selfish workers thus exert effort in period $t-1$ if:

$$w_k^{t-1} - c_k^{t-1} + (1-r)\left[(1-b_k^{t-1})V_{ks}^t + b_k^{t-1}V_{us}^t\right] \geq w_k^{t-1} + (1-r)V_{us}^t \tag{A.3}$$
$$\Leftrightarrow V_{ks}^t - V_{us}^t \geq \frac{c_k^{t-1}}{(1-r)(1-b_k^{t-1})}.$$

Denote by $B^t(e_i, V_{us}^t)$ the set of all vectors of separation rates (b_{1w}^t, b_{2w}^t) such that selfish types are willing to exert effort e_i in both types of firms, for a given value of unemployment V_{us}^t. Note that B^t may be empty for some effort levels, if future rents cannot compensate selfish types for the respective effort costs in the given period, as it is for instance the case in the final period of the game. Moreover, define the set $\bar{B}^t = B^t(e_n, \max_{L_t} V_{us}^t)$, which constitutes the set of separation rates for which the highest effort level is implementable even in the case of minimal unemployment threat.

In what follows, we derive a necessary and sufficient condition for separation rates to be in $\bar{B}^t$. For a worker of type j in period t the value of unemployment is given by:

$$\begin{aligned} V_{uj}^t = & \frac{(b_{1w}^{t-1}-1)L_{1w}^{t-1} + L_{1w}^t}{U + (1+b_{1w}^{t-1})L_{1w}^{t-1} + b_{2w}^{t-1}L_{2w}^{t-1}} V_{1w,j}^t + \frac{(b_{2w}^{t-1}-1)L_{2w}^{t-1} + L_{2w}^t}{U + (1+b_{1w}^{t-1})L_{1w}^{t-1} + b_{2w}^{t-1}L_{2w}^{t-1}} V_{2w,j}^t \\ & + \left[1 - \frac{(b_{1w}^{t-1}-1)L_{1w}^{t-1} + L_{1w}^t}{U + (1+b_{1w}^{t-1})L_{1w}^{t-1} + b_{2w}^{t-1}L_{2w}^{t-1}} - \frac{(b_{2w}^{t-1}-1)L_{2w}^{t-1} + L_{2w}^t}{U + (1+b_{1w}^{t-1})L_{1w}^{t-1} + b_{2w}^{t-1}L_{2w}^{t-1}}\right] V_{uj}^{t+1}(1-r). \end{aligned}$$

The equation illustrates the three components of the value of unemployment in period t: finding a new job in either a one-worker firm, or a two-worker firm, or remaining unemployed in that period. V_{uj}^t is endogenously determined in equilibrium. Depending on the contract renewal strategies of either firm type in period $t-1$, it is more or less likely to get hired in a corresponding job in period t. The total derivative with respect to the number of one-worker jobs L_{1w}^{t-1}, shows that V_{uj}^t is monotone in L_{1w}^{t-1}. Hence, V_{uj}^t is smaller than the maximum of its boundary values:

$$\lim_{L_{1w}^{t-1}\to\frac{N}{2}} V_{uj}^t = \underbrace{\frac{(b_{1w}^{t-1}-1)\frac{N}{2}+L_{1w}^t}{U+(1+b_{1w}^{t-1})\frac{N}{2}}V_{1w,j}^t + \frac{L_{2w}^t}{U+(1+b_{1w}^{t-1})\frac{N}{2}}V_{2w,j}^t + \frac{U+L_{1w}^t}{U+(1+b_{1w}^{t-1})\frac{N}{2}}V_{uj}^{t+1}(1-r)}_{=H_{1,j}(b_{1w}^{t-1},b_{2w}^{t-1})}$$

$$\lim_{L_{1w}^{t-1}\to 0} V_{uj}^t = \underbrace{\frac{L_{1w}^t}{U+b_{2w}^{t-1}N}V_{1w,j}^t + \frac{(b_{2w}^{t-1}-1)N+L_{2w}^t}{U+b_{2w}^{t-1}N}V_{2w,j}^t + \frac{U+L_{1w}^t}{U+b_{2w}^{t-1}N}V_{uj}^{t+1}(1-r)}_{=H_{2,j}(b_{1w}^{t-1},b_{2w}^{t-1})}.$$

H_1 and H_2 represent the value of unemployment for maximal and minimal level of unemployment, respectively. Since the maximal value of unemployment is, thus, well defined for each combination of separation rates, so is the set $\bar{B}^t$. A necessary and sufficient condition for a vector of separation rates to be in $\bar{B}^t$ is thus:

$$V_{ks}^t - \max\left\{H_{1,s}(b_{1w}^{t-1},b_{2w}^{t-1}), H_{2,s}(b_{1w}^{t-1},b_{2w}^{t-1})\right\} \geq \frac{c_n}{(1-r)(1-b_k^{t-1})}, \qquad k \in \{1w,2w\}.$$

Sufficient conditions for a segmentation equilibrium in pre-final periods: In the following we derive conditions for segmentation equilibria to arise, given separation rates that induce selfish workers to work. As discussed in the text, the data indicate that selfish types are indeed willing to work in pre-final periods of the experiment, consistent with such an equilibrium.

Condition 1. *There exist* $(b_{1w}, b_{2w}) \in \bar{B}^t$ *such that for all* $c_i \leq c_j$*:*

$$(a) \quad \Gamma_1^{t-1}(b_{1w}, b_{2w}, c_j, c_i) < z_{n+i} - z_j < \Gamma_2^{t-1}(b_{1w}, b_{2w}, c_j, c_i)$$
$$(b) \quad (1-b_k)N \leq L_k^t$$

Γ_1^{t-1} and Γ_2^{t-1} will be defined in the course of the proof and depend on the fairness consideration of workers. They represent the difference in wage payments in one- and two-worker firms for the tightest labor market (Γ_1^{t-1}) and the least tight labor market (Γ_2^{t-1}). The essence of part (a) of Condition 1 is that the production function is (i) "sufficiently concave" such that the output differential between one-worker and two-worker firms cannot become too large and (ii) steep enough such that a one-worker strategy inducing high effort does not dominate a two-worker strategy with lower effort levels for all possible labor market conditions. Part (b) of Condition 1 is purely technical: it guarantees that the number of jobs in period t is larger or equal to the number of workers who have their contract renewed.

Lemma A.1. *Suppose Condition 1 is fulfilled, the continuation values satisfy* $(1-r)V_{uf}^{t+1} < V_{2w,f}^t, V_{1w,f}^t$*, and* $\hat{\lambda} = \lambda$*. Then there exist an equilibrium with a segmented labor market in period* $t-1$ *that exhibits the following properties:*

1. *Effort levels* $e_{1w}^{t-1} \geq e_{2w}^{t-1}$ *are realized in one- and two-worker firms, respectively, with wages* $w_{1w}^{t-1} \geq w_{2w}^{t-1}$ *such that fair workers are indifferent between working and shirking.*

2. *Fair and selfish workers exert effort for the given wages and shirk if they get paid less.*

3. *There is a number* $L_{1w}^{t-1} > 0$ *of one-worker firms, a number* $L_{2w}^{t-1} > 0$ *of two-worker firms, and a number of* $L_{1w}^{t-1} + U$ *unemployed agents.*

4. *Workers who do not exert the stipulated effort level or who are known to be selfish are fired with certainty, those who exert the contractually stipulated effort have separation rates of* $(b_{1w}^{t-1}, b_{2w}^{t-1}) \in B^{t-1}$.

5. *From a fair worker's perspective, high-effort jobs in one-worker firms yield higher rents than jobs in two-worker firms, which in turn yield higher rents than unemployment.*

To prove the lemma we first characterize firms' optimal wage-effort schedules for given behavior of workers and given separation rates. In a second step, we show that there is an intermediate number of one-worker firms and corresponding separation rates, such that the derived wage-effort schedules for one-worker and two-worker firms are equally profitable for firms. This gives rise to a segmentation equilibrium, if the offered wage-effort schedules are incentive compatible for workers, which we show in the last step.

In any period, firms decide first on the wage and stipulated effort level. For any level of effort a firm pays wages to set fair workers indifferent between working and shirking. If a firm offered a contract with a lower wage, all workers would shirk, thereby decreasing firms profits. An offer of a higher wage inducing the same effort clearly also diminishes firm profits. We denote by $e_{1w}^{t-1}, e_{2w}^{t-1}$ the profit-maximizing levels of effort given the value of unemployment and implied wage payments.[6] Effort in one-worker firms needs to be higher than in two-worker firms. Otherwise firms could profitably deviate, due to the decreasing returns to scale production function. To see this, suppose the opposite $e_{1w}^{t-1} < e_{2w}^{t-1}$. From the optimal behavior of the firms we know:

$$f(e_{1w}^{t-1}) - w_{1w}^{t-1} \geq f(e_{2w}^{t-1}) - w_{2w}^{t-1} \qquad \text{and} \qquad f(2e_{2w}^{t-1}) - 2w_{2w}^{t-1} \geq f(2e_{1w}^{t-1}) - 2w_{1w}^{t-1}$$
$$\Rightarrow 2(f(e_{2w}^{t-1}) - f(e_{1w}^{t-1})) \leq f(2e_{2w}^{t-1}) - f(2e_{1w}^{t-1}).$$

This is a contradiction to Assumption A.1.

[6] If firms are indifferent between two or more levels of induced effort, we assume that there is a tie breaking rule that is homogeneous across firms.

From the fair workers' no-shirking condition (A.2), and the observation that future rents of employed workers exceed those of unemployed agents for at least one firm type, we know that there are wage-effort schedules such that the wage is below $g^{-1}(c_i) + c_i$. Since the production technology is efficient (Assumption A.1), firms will offer the most profitable contract rather than leaving the market. The decision of firms then boils down to deciding whether to employ only one worker or two workers. A one-worker strategy is more profitable if:

$$
\begin{aligned}
& z_{1w}^{t-1} - w_{1w}^{t-1} > z_{2w}^{t-1} - 2w_{2w}^{t-1} \\
\Leftrightarrow \quad & 0 < z_{1w}^{t-1} - (g^{-1}(c_{1w}^{t-1} + (1-r)(b_{1w}^{t-1} - 1)(V_{1w,f}^{t} - V_{uf}^{t}) + c_{1w}^{t-1}) \\
& - z_{2w}^{t-1} + 2(g^{-1}(c_{2w}^{t-1} + (1-r)(b_{2w}^{t-1} - 1)(V_{2w,f}^{t} - V_{uf}^{t})) + c_{2w}^{t-1}) \equiv \Delta
\end{aligned}
$$

Whether this is the case depends on the difference in outputs between one-worker firms (z_{1w}^{t-1}) and two-worker firms (z_{2w}^{t-1}), and the tightness of the labor market, which determines V_{uf}^{t}. Δ is continuous in V_{uf}^{t}, therefore a shift in the sign of this inequality leads to at least one level for the value of unemployment such that firms are indifferent between both strategies.[7] This reversal in the sign of Δ is given if:[8]

$$
\begin{aligned}
& \Gamma_1^{t-1}(b_{1w}, b_{2w}, c_{1w}^{t-1}, c_{2w}^{t-1}) \equiv 2g^{-1}\left(c_{2w}^{t-1} + (1-r)(b_{2w}^{t-1} - 1)(V_{2w,f}^{t} - H_{1,f})\right) + 2c_{2w}^{t-1} \\
& - g^{-1}\left(c_{1w}^{t-1} + (1-r)(b_{1w}^{t-1} - 1)(V_{1w,f}^{t} - H_{1,f})\right) - c_{1w}^{t-1} \\
& < z_{2w}^{t-1} - z_{1w}^{t-1} \\
& < 2g^{-1}\left(c_{2w}^{t-1} + (1-r)(b_{2w}^{t-1} - 1)(V_{2w,f}^{t} - H_{2,f})\right) + 2c_{2w}^{t-1} \\
& - g^{-1}\left(c_{1w}^{t-1} + (1-r)(b_{1w}^{t-1} - 1)(V_{1w,f}^{t} - H_{2,f})\right) - c_{1w}^{t-1} \equiv \Gamma_2^{t-1}(b_{1w}, b_{2w}, c_{1w}^{t-1}, c_{2w}^{t-1})
\end{aligned}
$$

Condition 1 ensures that a combination of separation rates from the set $\bar{B}^{t-1}$ exists such that this is fulfilled. Hence, there is a number L_{1w}^{t-1} such that firms are indifferent between both strategies. This gives rise to a segmentation equilibrium in which L_{1w}^{t-1} firms employ only one worker at a high level of effort. These firms do not fill their second vacancy and thus the equilibrium features endogenous unemployment. Note that the separation rates in any segmentation equilibrium must be larger than zero. If they are zero, firms renew all their contracts and there are no vacancies in the next period, independently of the fraction of one-worker firms. The value of unemployment would then be independent of the current fraction of firms that ration jobs ($H_{1,j} = H_{2,j}$). Hence, the profitability of firm strategies would not depend on the labor market conditions $\left(\Gamma_1^{t-1} = \Gamma_2^{t-1}\right)$, and one firm size would dominate the alternative throughout.

[7] If the optimal induced effort choice changes in the course of varying the number of one-worker firms from zero to $\frac{N}{2}$ the segmentation equilibria may feature three different levels of effort.

[8] There is also a segmentation equilibrium if $\lim_{L_{1w}^{t-1} \to \frac{N}{2}} \Delta > 0 > \lim_{L_{1w}^{t-1} \to 0} \Delta$. The following analysis is also valid for this case.

After firms have observed the effort choice of the worker they decide whether to renew the contract. In the Lemma we consider the case in which firms have no additional information about the type of agent they employ compared to those that are unemployed ($\hat{\lambda} = \lambda$). Hence, there is always an "equally good" worker available, implying that firms are indifferent between dismissing the worker and renewing the contract. All separation rates are thus incentive compatible.

Turning to the workers, fair types by construction choose to exert the stipulated effort level. Furthermore, selfish workers comply to the contract, since the separation rates are in $\bar{B}^{t-1}$. Hence, there exists an equilibrium fulfilling properties 1 to 4 from Lemma A.1.

Part 5 of Lemma A.1 implies that the segmentation of the labor market into one- and two-worker firms has strong consequences for workers. First, a fraction of workers is unemployed, and this unemployment is involuntary. This follows since an unemployed agent only receives the discounted value of unemployment from the next period which is less than what an agent in either job receives (cf. equation (A.1)). Second, from a fair worker's perspective, there are two types of jobs: "primary-sector" jobs that pay high rents for high efforts, and "secondary-sector" jobs with lower rent payments and lower efforts. This difference arises since workers have some discretion about their effort level. The higher effort levels in one-worker firms imply higher wages, which in turn yields higher rents by the no shirking condition of fair types:

$$V_{1w,f}^{t-1} \geq V_{2w,f}^{t-1} \qquad \Leftrightarrow \qquad w_{1w}^{t-1} + (1-r)V_{uf}^{t} \geq w_{2w}^{t-1} + (1-r)V_{uf}^{t}.$$

This concludes the proof of Lemma A.1.

Proposition A.1. *If Condition 1 is fulfilled for all $t < T$, there exists an equilibrium with a segmented labor market and involuntary unemployment in all periods $t < T$.*

Since Condition 1 is satisfied for all t, there is an equilibrium with a segmented labor market in all periods, if (i) $(1-r)V_{uf}^{t+1} < V_{2w,f}^{t}, V_{1w,f}^{t}$ and (ii) $\hat{\lambda} = \lambda$ hold for all t. The first follows immediately from the no-shirking condition (A.2), since wages are positive. Moreover, workers behave homogeneously in every period. As a consequence, screening is not possible and $\hat{\lambda} = \lambda$ in all periods. This implies the existence of a segmentation equilibrium featuring involuntary unemployment in all periods.

A.2.2. Additional figures and tables

	Dependent variable:							
	work effort e_i				1 if $e_i < \hat{e}_i$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wage (w_i)	0.121***	0.117***	0.118***	0.114***	-0.020***	-0.018***	-0.019***	-0.018***
	(0.010)	(0.010)	(0.010)	(0.012)	(0.003)	(0.003)	(0.003)	(0.003)
Desired effort ($\hat{e}_i$)	0.248***	0.251***	0.257***	0.235***	0.096***	0.096***	0.094***	0.098***
	(0.066)	(0.066)	(0.064)	(0.075)	(0.017)	(0.017)	(0.016)	(0.018)
Wage coworker (w_j)	-0.004	-0.001	-0.003	-0.008	0.002	0.002	0.002	0.003
	(0.011)	(0.011)	(0.010)	(0.012)	(0.003)	(0.003)	(0.003)	(0.003)
Desired effort coworker ($\hat{e}_j$)	-0.003	-0.012	-0.011	0.019	0.017	0.019	0.017	0.015
	(0.064)	(0.064)	(0.059)	(0.068)	(0.018)	(0.018)	(0.016)	(0.018)
Relational contract		0.400	0.239	-0.186		-0.136**	-0.068	0.011
		(0.279)	(0.287)	(0.322)		(0.067)	(0.066)	(0.074)
Market period			0.042	0.030			-0.015***	-0.014**
			(0.028)	(0.032)			(0.005)	(0.006)
Final period			-1.605***	-1.415**			0.363***	0.360**
			(0.586)	(0.716)			(0.133)	(0.152)
Constant	0.771***	0.744***	0.538	1.184***	0.159**	0.166**	0.271***	0.103
	(0.243)	(0.230)	(0.365)	(0.445)	(0.073)	(0.071)	(0.088)	(0.105)
N	452	452	452	452	452	452	452	452
Worker fixed effects	no	no	no	yes	no	no	no	yes

Table A.2.: *Influence of co-worker wage on effort provision.* *** *indicates significance on the 1-percent level,* ** *significance on the 5-percent level,* * *significance on the 10-percent level. All models are estimated with random effects at the firm and session level.*

A.2.3. Instructions of the experiment (IC Treatment)

In what follows, we present a translation of the instructions for buyers (i.e., employers) in the IC treatment. The instructions for workers in this treatment had a similar structure. The instructions of participants in the C treatment differed only in the description of the second stage (i.e., the work phase).

Instructions for Buyers

You are now taking part in an economic experiment. Please read the following instructions carefully. Everything that you need to know to participate in this experiment is explained below. Should you have any difficulties in understanding these instructions please raise your hand. We will answer your questions at your cubicle.
At the beginning of the experiment you receive an initial endowment of **8 euros**. During the experiment you can increase your income by earning **points**. The amount of points that you earn during the experiment depends on your decisions and the decisions of other participants.
All points that you earn during the experiment will be converted into euros at the end of the experiment. The exchange rate is:

1 Point = 4 cents

At the end of the experiment, the amount of money that you earned during the experiment will be paid out in cash.
The experiment consists of several periods. In each period you have to make decisions which you enter in a computer. There will be 18 periods in total.
Please note that communication between participants is strictly prohibited during the entire experiment. In addition we would like to point out that you may only use the computer functions which are required for the experiment. Violations of these rules will lead to exclusion from the experiment. In case you have any questions we shall be glad to assist you.
Prior to the experiment, the 24 participants were divided into 2 groups: buyers and sellers. There are 7 buyers and 17 sellers in the experiment.
You will be a buyer for the entire duration of the experiment. All participants have received an identification number which they will keep for the entire experiment. You can find your identification number on the documentation sheet in front of you.

Short Overview of the Experimental Procedures

In each period of the experiment every buyer can trade a product with zero, one, or two sellers. The seller earns profits when he sells the product at a price which exceeds his production costs. The buyer earns profits when the price she pays for the product is less than her valuation of the product. The production costs of the traded product as well as the buyer's valuation of the product depend on the quality of the product. In addition, the value of the product for the buyer depends on the number of products bought. Two products of a certain quality are worth more, but not twice as much as one product of the same quality.

The experiment lasts for 18 periods. In each period, procedures are as follows.

1. Each period commences with a **trading phase** that lasts for 200 seconds. During this phase buyers can submit trade offers which can be accepted by sellers. When submitting an offer a buyer has to specify three variables:

- Which price she offers to pay
- which product quality she desires
- and finally, which sellers she wants to submit the offer to. Buyers can submit two types of offers: private offers and public offers. **Private offers are submitted to one specific seller** and can only be accepted by that seller. **Public offers are submitted to all sellers** and can be accepted by any seller.

As a buyer you can submit as many offers as you like in each period. Once submitted, offers can be accepted at any time. **Each seller can at most conclude one trade agreement in each period. Each buyer can at most conclude two trade agreements per period.** As there are 7 buyers and 17 sellers in total, some sellers will not trade in each period.

2. After the trading phase, every seller who concluded a trade agreement has to determine which product quality he provides to his buyer. **The seller does not have to provide the product quality desired by the buyer.**

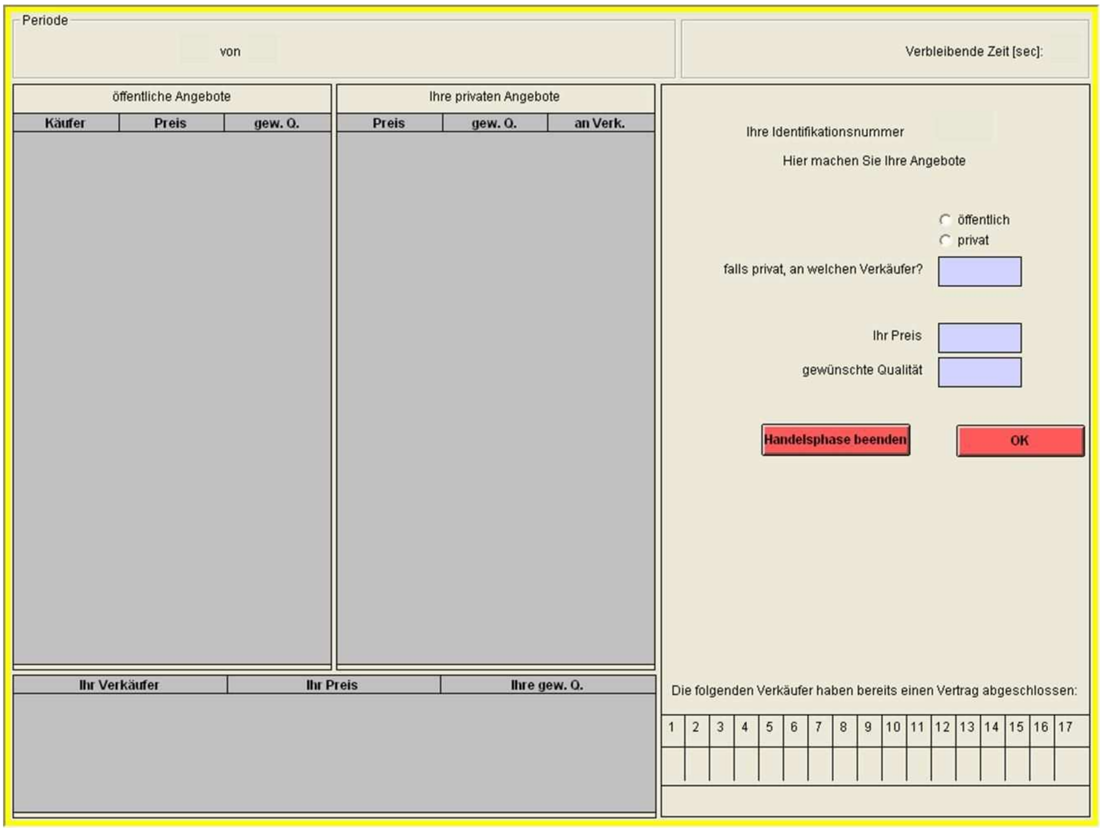
Once every seller has chosen his product quality, earnings of all participants for the given period are determined. Subsequently, the next period starts. Earnings of all 18 periods will be summed up at the end of the experiment, converted into euros and paid out in cash together with your initial endowment.

The Experimental Procedures in Detail

There are 7 buyers and 17 sellers in the experiment. You are a **buyer** for the entire duration of the experiment. During the experiment you enter your decisions in a computer. In the following, we describe in detail how you make your decisions in each period.

1. The Trading Phase

Each period begins with a trading phase. During the trading phase the buyers can conclude trade agreements with the sellers. In order to do so, **each buyer can submit as many trade offers as she wishes.** In each trading phase you will see the following screen.



In the top left corner of the screen you will see the current period of the experiment. In the top right corner of the screen you will see the time remaining in this trading phase, displayed in seconds. **The trading phase in each period lasts 200 seconds.** When this time is up the trading phase is over, and no further offers can be submitted or accepted in this period.

Once you see the screen displayed, the trading phase starts. As a buyer, you now have the opportunity to submit trade offers to the sellers. In order to do so you have to enter three variables on the right hand side of the screen:

(A) First, you have to specify whether you want to submit a public or private offer:

Public offers

Public offers are communicated to all participants in the market. All sellers see all public offers on their screens. A public offer can therefore be accepted by **any seller**. As a buyer, you will also see all public offers submitted by the other buyers. If you want to submit a public offer, please click on the field “public” using the mouse.

Private offers

Private offers are submitted to **one seller** only. Only this seller will be informed about the offer and only this seller can accept the trade offer. No other seller or buyer will be informed about the offer. If you want to submit a private offer, please click on the field “private” using the mouse. In the next field, you have to specify **to which seller you want to submit the offer.** Each of the 17 sellers has an identification number (Seller1, Seller2, ..., Seller17). Each seller keeps his identification number for the entire duration of the experiment. To submit an offer to a specific seller, please enter the number of that seller (e.g. “4” for Seller4).

(B) Once you have specified who you want to submit an offer to, you have to determine **which price you offer.** You enter the price into the field “your price”. The price you offer must not be below 0 or above 100:

$$\mathbf{0 \leq Price \leq 100}$$

(C) Finally, you have to specify which product quality you desire. You enter this in the field “desired quality”. **Your desired quality** cannot be lower than 1 or higher than 10.

$$\mathbf{1 \leq Desired\ quality \leq 10}$$

After you have fully specified your trade offer, you have to click the “OK” button to submit it. As long as you have not clicked “OK” you can still change your offer. After you click “OK” the offer will be displayed to all sellers to whom it has been submitted.

On the left side of your screen you see the heading “public offers”. All public offers in the current trading phase will be displayed here—your public offers as well as the public offers of all other buyers. You can see which buyer submitted the offer, which price she offered and which quality she desired. All buyers also have an identification number that they keep throughout the experiment (Buyer1, Buyer2, ..., Buyer7).

In the middle column of the screen, under the header “your private offers” you will see all private offers that you have submitted in the current trading phase. Here you can see to which sellers you made an offer, which price you offered and which quality you desired.

As long as none of your offers has been accepted by a seller, you as a buyer can submit as many private and public offers as you wish in a given period.

Each offer that you submit can be accepted at any time during the trading phase.

As soon as one of your offers has been accepted, you are informed which seller accepted the offer and which of your offers has been accepted. In the bottom left corner of your screen the identification number of the seller who accepted the offer will be displayed, together with your offered price and your desired quality. At the same time all your other offers will be automatically canceled.

You can then decide whether you want to conclude another trade agreement. **Each buyer can conclude zero, one, or two trade agreements in each period.** If you want to conclude another trade agreement you can submit further offers to the sellers. As long as none of your offers has been accepted by a seller you can offer as many private and public offers for the second trade as you wish.

If you do not want to conclude another trade agreement you can press the button “finish trading phase”. This reduces the length of the trading phase in case no other buyer wants to submit further offers. By pressing the button, the offers you have already submitted will be automatically canceled and you can not submit further offers. Trade agreements which were already accepted by a seller of course persist. In addition, you will continue to see the screen of the trading phase until it is over.

Each seller can conclude at most one trade agreement in a given period. You will be constantly informed which sellers have not yet concluded a trade agreement. In the table with the title “The following sellers have already concluded a contract” you can see 17 fields. Once a seller has accepted an offer, a “+” will appear in the field below his identification number. You cannot submit private offers to a seller who has already accepted an offer.

The trading phase is over as soon as one of the following occurs: 200 seconds have elapsed, or all buyers have concluded two trade agreements, or the remaining buyers have signalled that they do not want to conclude trade agreements anymore by pressing the button “finish trading phase”.

No buyer is obliged to submit trade offers, and no seller is obliged to accept a trade offer.

2. Determination of actual product quality

After the trading phase, all sellers who have concluded a trade agreement determine which product quality they will supply to their buyer. First, the sellers see again the price and the desired quality on a new screen. If you have concluded two trade agreements in a given period your sellers can also see the price and the desired quality of your other seller. The sellers then decide independently which actual product quality to choose for their product. **The product quality which you desired in your trade offer is not binding for your seller(s).** Your seller can choose exactly the quality you desired, but he can also choose a higher or lower product quality. The product quality that your seller chooses has to be an integer between 1 and 10:

$$\mathbf{1 \leq Actual\ product\ quality \leq 10}$$

While the sellers determine the actual product quality, we will ask you on a separate screen to specify which quality(ies) you expect him (them) to supply. In addition we ask you to state how sure you are about this expectation.

How are incomes calculated?

Your income:

If you **have not concluded a trade agreement** during a trading phase you earn an income of **0 points** in this period.

If you have concluded **one** trade agreement, your income depends on the price you paid and the product quality your seller supplied to you. Your income equals 10 times the actual product quality minus the price you paid. Your income thus amounts to:

Your income = 10*Actual product quality - Price

If you have concluded **two** trade agreements, your income depends on which prices you paid to the sellers and which product qualities were supplied to you. The value of the products **in total** can be higher for you if you conclude two trade agreements, but the value of **each individual** product is lower in this case.

In other words, two products of a certain quality are worth more to you than one product of the same quality; but they are not worth twice as much as one product of the same quality. If you buy **one** product you earn 10 times the chosen product quality. If you buy **two** products you earn 7 times the quality of the first product and 7 times the quality of the other product. Of course, when you buy two products you also have to pay two prices. Your income if you conclude two trade agreements is thus calculated as follows:

Your income = 7*Actual product quality Product 1
+ 7*Actual product quality Product 2 - Price 1 - Price 2

An example: If you conclude **one** trade agreement and the actual product quality is 8, your income is 80 minus the price you paid. If you conclude **two** trade agreements and both actual product qualities are 8 your income is 112 (=7*8+7*8) minus both prices. If, for instance, one actual product quality is 8 and the other quality is 1, your income is 63 (= 7*8 + 7*1) minus both prices.

As you can see from the above formula your income generally increases in the product quality actually supplied by the seller(s). At the same time your income is higher, the lower the price(s) you have to pay for the product(s).

Income of your seller:

If a seller has not concluded a trade agreement during a trading phase he earns an income of 0 points in this period.

If a seller has accepted a trade offer his income will equal the price he receives minus the production costs he incurs for supplying the product The income of your seller is determined as follows:

Income of your seller = Price - Production Costs

The production costs of a seller are higher, the higher the quality of the product he chooses. The production costs for each product quality are displayed in the table below:

Quality	1	2	3	4	5	6	7	8	9	10
Production costs	0	1	2	4	6	8	10	12	15	18

The income of your seller increases in the price he receives. Furthermore, his income is higher, the lower the product quality he supplies.
The incomes of all buyers and sellers are determined in the same way. **Each buyer can therefore calculate the income of his seller(s) and each seller can calculate the income of her buyer.** In addition, buyers and sellers are informed of the identification number of their trading partner in a given period.
Please note that buyers and sellers can also incur losses in each period. These losses have to be covered from your initial endowment or from earnings in other periods.
You will be informed about your income and the income of your seller on a separate **"income screen"**. On the screen the following information will be displayed:

- Which seller(s) you traded with
- Which price(s) you paid
- Your desired quality(ies)
- The actual product quality(ies) supplied by your seller(s)
- The income of your seller(s) in this period
- Your income in this period.

Please enter all of the information from this screen into the documentation sheet on your desk. After the income screen has been displayed, the period is finished. Thereafter the trading phase of the following period starts. Once you have finished reading the income screen please click on the "OK" button.
The sellers also see an income screen which displays the above information. They see the ID of their buyer, the price, the desired and actual product quality as well as their own

income, your income and—if you have concluded two trade agreements—the income of your other seller.

The experiment will not start until all participants are completely familiar with all procedures. In order to make sure that this is the case we ask you to solve the exercises below.

In addition we will conduct a **trial period of the trading phase** to familiarize you with using the computer. This trial phase will not be added to the results of the experiment, and will not be remunerated. After the trial phase, the experiment which lasts for 18 periods will start.

A.3. Appendix to Chapter III

A.3.1. Proofs of propositions and lemmas main part

Proof to Proposition 4.1:

Efficient policy choice: Since voters can directly observe candidates' abilities as well as their policies, voters have that the belief system $\sigma = a$ and are able to fully anticipate the difference in payoffs. The vote share of Candidate 1 is given by:

$$v_1(x_1, x_2, \sigma) = 1 - \Omega(x_2(\hat{a_2}(x_2) - c) - x_1(\hat{a_1}(x_1) - c)) = 1 - \Omega(x_2(a_2 - c) - x_1(a_1 - c)).$$

Candidate 1 chooses x_1, taking into account his opponent's strategy X_2, to maximize

$$U_1(x_1, a_1) = \int_0^1 \phi(a_2) f(v_1(x_1, x_2, \sigma), \rho)\,(\theta + x_1(a_1 - c))\, da_2$$

As f is strictly increasing in v_1, which in turn is strictly increasing in the difference in welfare contributions, candidate 1 is only interested in maximizing his welfare contribution. Clearly, the dominant strategy is given by

$$X_1(a_1) = \begin{cases} x_i = 0 & \text{for} \quad a_i < c \\ x_i = 1 & \phantom{\text{for}} \quad a_i \geq c. \end{cases}$$

Positive welfare effect of increasing power concentration: To simplify notation, denote the welfare contribution of player i by $\pi_i(x_i, \sigma) = x_i(\hat{a}_i(x_i) - c)$. We suppress the dependence on the belief system and the action if possible without creating confusion. Moreover, let $g(\pi_1 - \pi_2, \rho)$ be the expected power share above one half

$$f(v_1(x_1, x_2, \sigma), \rho) - \frac{1}{2} = g(\pi_1 - \pi_2, \rho).$$

For any candidate there are two cases. He can either face an opponent with a reform proposal or one that proposed the status quo. The ex ante welfare is the weighted average of these two alternatives. Given the optimal behavior identified above, welfare in the full information case is given by

$$\begin{aligned} &W(\rho) = \\ &\int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[\frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[\frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1 \\ &+ \int_0^1 \int_{a_1}^1 \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[\frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[\frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1 \\ &= 2 \int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[\frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[\frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1. \end{aligned}$$

The derivative of the welfare function with respect to power concentration ρ is given by

$$\frac{dW}{d\rho} = 2 \int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) \left[\pi(X_1(a_1), \sigma) - \pi(X_2(a_2), \sigma) \right] \frac{\partial g(\pi_1 - \pi_2, \rho)}{\partial \rho} da_2 da_1 > 0.$$

As $a_1 > a_2$ under the integral, the payoff difference $\pi_1 - \pi_2$ and $g(\pi_1 - \pi_2, \rho)$ are throughout positive, so that we have $\frac{\partial g(\pi_1 - \pi_2, \rho)}{\partial \rho} > 0$ due to the properties of the function $f(v, \rho)$.

Proof of Lemma 4.1: We only need to deal with the case of a politician with ability lower than c, since candidates with ability above c always choose to reform. Candidate 1 chooses to reform if and only if:

$$\begin{aligned}
&prob(x_2 = 1)f(v_1(x_1 = 1, x_2 = 1, \sigma), \rho)(\theta + a_1 - c) \\
&+ (1 - prob(x_2 = 1))f(v_1(x_1 = 1, x_2 = 2, \sigma), \rho)(\theta + a_1 - c) \\
&- prob(x_2 = 1)f(v_1(x_1 = 0, x_2 = 1, \sigma), \rho)\theta - (1 - prob(x_2 = 1))\theta\frac{1}{2} > 0.
\end{aligned}$$

It can easily be seen that this reform incentive function is strictly monotonously increasing in the individual ability a. The same argument holds for Candidate 2. Thus, the optimal strategy of each candidate will always be of the cutoff type.

Proof of Proposition 4.2: Depending on the parameter values there may exist equilibria with cutoff being equal to one. As argued in the text these cannot be divine. In the following we show that there exists a unique equilibrium exhibiting cutoffs different from 1 and that this equilibrium features symmetric cutoffs.

Symmetry of cutoffs: Using the insight from above, we can write the incentive function of player 1 as:

$$\begin{aligned}
R_1(a, \alpha_1, \alpha_2, \rho) =&(1 - \Phi(\alpha_2))f(\rho, v_1(x_1 = 1, x_2 = 1, \sigma))(\theta + a - c) \\
&+ \Phi(\alpha_2)f(\rho, v_1(x_1 = 1, x_2 = 0, \sigma))(\theta + a - c) \\
&- (1 - \Phi(\alpha_2))f(\rho, v_1(x_1 = 0, x_2 = 1, \sigma))\theta - \Phi(\alpha_2)\frac{\theta}{2}.
\end{aligned}$$

Using the definition of $g(\pi_1 - \pi_2, \rho)$, this simplifies to

$$\begin{aligned}
&R(a, \alpha_1, \alpha_2) = (1 - \Phi(\alpha_2)) \left[\left(g(\pi_1 - \pi_2, \rho) + \frac{1}{2} \right) (\theta + a - c) \right] \\
&+ \Phi(\alpha_2) \left[\left(g(\pi_1, \rho) + \frac{1}{2} \right) (\theta + a - c) \right] - (1 - \Phi(\alpha_2)) \left(\frac{1}{2} - g(\pi_2, \rho) \right) \theta - \Phi(\alpha_2)\theta\frac{1}{2}.
\end{aligned}$$

In equilibrium, the reform incentive is zero for the cutoff-type α_1.

$$\begin{aligned}
&R_1(\alpha_1, \alpha_1, \alpha_2) = 0 \\
&\Leftrightarrow \frac{\theta \left[\Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2)) \left[g(\pi_1 - \pi_2, \rho) + g(\pi_2, \rho) \right] \right]}{c - \alpha_1} = \\
&\frac{1}{2} + \Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2))g(\pi_1 - \pi_2).
\end{aligned}$$

Subtracting the corresponding equation for R_2, we get

$$\frac{\theta\left[\Phi(\alpha_2)g(\pi_1,\rho)+(1-\Phi(\alpha_2))\left[g(\pi_1-\pi_2,\rho)+g(\pi_2,\rho)\right]\right]}{c-\alpha_1}$$
$$-\frac{\theta\left[\Phi(\alpha_1)g(\pi_2,\rho)+(1-\Phi(\alpha_1))\left[-g(\pi_1-\pi_2,\rho)+g(\pi_1,\rho)\right]\right]}{c-\alpha_2}=$$
$$\Phi(\alpha_2)g(\pi_1,\rho)+(1-\Phi(\alpha_2))g(\pi_1-\pi_2,\rho)-\Phi(\alpha_1)g(\pi_2,\rho)+(1-\Phi(\alpha_1))g(\pi_1-\pi_2,\rho)$$
$$\Leftrightarrow\left[\frac{\theta\Phi(\alpha_2)}{c-\alpha_1}-\frac{\theta(1-\Phi(\alpha_1))}{c-\alpha_2}-\Phi(\alpha_2)\right]g(\pi_1,\rho)$$
$$-\left[\frac{\theta\Phi(\alpha_1)}{c-\alpha_2}-\frac{\theta(1-\Phi(\alpha_2))}{c-\alpha_1}-\Phi(\alpha_1)\right]g(\pi_2,\rho)$$
$$+\underbrace{\left[(1-\Phi(\alpha_2))\left(\frac{\theta}{c-\alpha_1}-1\right)+(1-\Phi(\alpha_1))\left(\frac{\theta}{c-\alpha_2}-1\right)\right]g(\pi_1-\pi_2,\rho)}_{>0}=0.$$

If $\alpha_1=\alpha_2$, this condition is trivially fulfilled. Assuming wlog $\alpha_1>\alpha_2$, the equality above can only be satisfied, if

$$\left[\frac{\theta\Phi(\alpha_2)}{c-\alpha_1}-\frac{\theta(1-\Phi(\alpha_1))}{c-\alpha_2}-\Phi(\alpha_2)\right]g(\pi_1,\rho)<$$
$$\left[\frac{\theta\Phi(\alpha_1)}{c-\alpha_2}-\frac{\theta(1-\Phi(\alpha_2))}{c-\alpha_1}-\Phi(\alpha_1)\right]g(\pi_2,\rho).$$

However, we have $\pi_1>\pi_2$ by assumption, which implies $g(\pi_1,\rho)>g(\pi_2,\rho)$. Furthermore, we can show that the factor before $g(\pi_1,\rho)$ is larger than the one before $g(\pi_2,\rho)$:

$$\frac{\theta}{c-\alpha_1}\Phi(\alpha_2)-\frac{\theta}{c-\alpha_2}(1-\Phi(\alpha_1))-\Phi(\alpha_2) > \frac{\theta}{c-\alpha_2}\Phi(\alpha_1)-\frac{\theta}{c-\alpha_1}(1-\Phi(\alpha_2))-\Phi(\alpha_1)$$
$$\Leftrightarrow\frac{\theta}{c-\alpha_1}+\Phi(\alpha_1) > \frac{\theta}{c-\alpha_2}+\Phi(\alpha_2).$$

The last inequality is clearly fulfilled, generating a contradiction. Thus, the reform incentive functions R_1 and R_2 can never attain zero simultaneously for different cutoffs and there are only symmetric equilibria.

Existence: Let π denote the difference in welfare contributions between a reform and a status quo proposal. Making use of the symmetric cutoffs, the incentive function simplifies to:

$$R(\alpha,\rho)=\left[\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)\right](\alpha-c)+\theta g(\pi,\rho)=0.$$

Note that $R(1,\rho)$ is always positive if $\alpha=1$. If $R(0,\rho)<0$, the reform incentive is equal to zero at least once, due to the continuity and there exists an interior equilibrium. If $R(0,\rho)\geq 0$, it is an equilibrium that all candidates choose to reform. Hence, there is at least one equilibrium.

Proof of Proposition 4.3: Next, we establish uniqueness. The derivative of the incentive function with respect to α is:

$$\frac{\partial R}{\partial\alpha}=\underbrace{(\theta+(\alpha-c)\Phi(\alpha))g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha}}_{A}+\underbrace{\left(\frac{1}{2}+(\Phi(\alpha)+(\alpha-c)\phi(\alpha))g(\pi,\rho)\right)}_{B}.$$

The reform incentive function yields that A is always larger than zero in equilibrium for the cutoff type. B is also larger than zero, due to Assumption 4.1. The reform incentive is thus increasing at any cutoff. Consequently, there can only be one cutoff value.

Since the incentive function is equal to zero in any informative equilibrium, we use implicit differentiation to prove that there is a unique $\tilde{\theta}(\rho)$. The cutoff α is given by the maximum of zero and the value implying a reform incentive equal to zero. If there is an informative equilibrium the derivative is given by

$$\frac{d\alpha}{d\theta} = -\frac{g(\pi,\rho)}{(\theta+(\alpha-c)\Phi(\alpha))g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha}+\left(\frac{1}{2}+(\Phi(\alpha)+(\alpha-c)\phi(\alpha))g(\pi,\rho)\right)} < 0.$$

The denominator is positive (see above), as is numerator. Thus, the derivative is strictly negative in any interior equilibrium. Moreover, the reform incentive implies that $\alpha \to c$ if $\theta \to 0$. Hence, there is a unique $\tilde{\theta} > 0$, such that the unique equilibrium is informative if $\theta < \tilde{\theta}$.

Proof of Proposition 4.4: Again we use implicit differentiation to evaluate the derivative.

$$\frac{d\alpha}{d\rho} = -\frac{\frac{\partial R}{\partial\rho}}{\frac{\partial R}{\partial\alpha}} = -\frac{(\theta+(\alpha-c)\Phi(\alpha))g_\rho(\pi,\rho)}{(\theta+(\alpha-c)\Phi(\alpha))g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha}+\left(\frac{1}{2}+(\Phi(\alpha)+(\alpha-c)\phi(\alpha))g(\pi,\rho)\right)} < 0.$$

While the numerator is unambiguously positive, the positive sign of the denominator follows from Assumption 4.2. Hence, the overall effect is negative.

Proof of Lemma 4.2: Using the symmetry in equilibrium, the welfare can be simplified considerably.

$$\frac{W(\rho)}{2} = \underbrace{\int_\alpha^1 \phi(a)(a-c)da}_{z(\alpha)} \left(\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)\right).$$

Note that there is a direct effect on welfare, since the function $g(\pi,\rho)$ depends on ρ and an indirect effect since ρ changes the strategies of the politicians. Hence, we evaluate the total derivative of $W(\rho)$:

$$\frac{dW}{d\rho} = \frac{\partial W}{\partial\rho} + \frac{\partial W}{\partial\alpha}\frac{d\alpha}{d\rho}.$$

In the following we denote by $D > 0$ the denominator of the derivative of α with respect to ρ.

$$\begin{aligned}
\frac{dW}{d\rho} =& \Phi(\alpha)z(\alpha)g_\rho(\pi,\rho)+ \\
&+\left\{(c-\alpha)\phi(\alpha)\left(\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)\right)+z(\alpha)\left(\phi(\alpha)g(\pi,\rho)+\Phi(\alpha)g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha}\right)\right\}\frac{d\alpha}{d\rho} \\
=&\left\{\Phi(\alpha)z(\alpha)\left[\theta+(\alpha-c)\Phi(\alpha)\right]g_\pi\frac{d\pi}{d\alpha}+\Phi(\alpha)z(\alpha)\left[\frac{1}{2}+(\Phi(\alpha)+(\alpha-c)\phi(\alpha))g(\pi,\rho)\right]\right. \\
&-\left[(c-\alpha)\phi(\alpha)\left(\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)\right)+z(\alpha)\left(\phi(\alpha)g(\pi,\rho)+\Phi(\alpha)g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha}\right)\right] \\
&\left.[\theta+\Phi(\alpha)(\alpha-c)]\right\}\frac{g_\rho(\pi,\rho)}{D} \\
=&\frac{g_\rho(\pi,\rho)}{D}\left\{\Phi(\alpha)z(\alpha)\left[\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)+(\alpha-c)\phi(\alpha)g(\pi,\rho)\right]\right. \\
&\left.-\left[\phi(\alpha)\theta g(\pi,\rho)+z(\alpha)\phi(\alpha)g(\pi,\rho)\right]\frac{c-\alpha}{2g(\pi,\rho)}\right\} \\
=&\frac{g_\rho(\pi,\rho)}{D}\left\{\Phi(\alpha)z(\alpha)\left(\frac{1}{2}+\Phi(\alpha)g(\pi,\rho)\right)-\phi(\alpha)(c-\alpha)\left[\frac{\theta}{2}+z(\alpha)\left(\frac{1}{2}+\Phi(\alpha)\right)\right]\right\} \\
=&\frac{g_\rho(\pi,\rho)}{D}\left\{\Phi(\alpha)\frac{W(\rho)}{2}-\phi(\alpha)(c-\alpha)\left[\frac{\theta}{2}+\frac{W(\rho)}{2}\right]\right\} \\
=&\frac{g_\rho(\pi,\rho)}{2D}\left\{\Phi(\alpha)W(\rho)-\phi(\alpha)(c-\alpha)\left(\theta+W(\rho)\right)\right\}
\end{aligned}$$

In any equilibrium, the term in brackets has to equal zero, since its factor is always positive. Rearranging, we get the following necessary and sufficient condition for extreme values of the welfare function:

$$h(\rho)\equiv\frac{\Phi(\alpha)}{\phi(\alpha)(c-\alpha)}-\left(1+\frac{\theta}{W(\rho)}\right)=0.$$

Next, we prove that function h has at most one root in ρ, i. e., the welfare function attains at most one maximum. Assumption 4.2 is a sufficient condition for the first term to be decreasing in ρ and, thus, increasing in α. In any extreme value of the welfare function, the second term is constant in ρ. Thus, h is decreasing in each root and so is the term in brackets. As $h(\rho)$ is continuous in ρ, this implies that the welfare function has at most one interior maximum and no interior minimum, i. e., it is strictly quasi-concave.

Proof of Proposition 4.5: In the next step, we show how the derived maximum shifts with changes in θ. For $\theta \to 0$ we get from the equilibrium condition $\alpha = c$. The derivative of the welfare function at $\theta = 0$ is given by:

$$\left.\frac{dW(\rho)}{d\rho}\right|_{\theta=0}=\frac{g_\rho(\pi,\rho)}{D}\Phi(\alpha)W(\rho).$$

This is positive. Hence, for $\theta \to 0$ the optimal institution embodies full concentration of power. Due to continuity, this is also true for an interval around 0. Finally, we show that the optimal ρ decreases monotonically in θ.

$$\frac{d\rho^*}{d\theta}=-\frac{\frac{dh(\rho)}{d\theta}}{\left.\frac{dh(\rho)}{d\rho}\right|_{\rho=\rho^*}}.$$

As argued before the term in the denominator is negative. With respect to the numerator, note that the equilibrium cutoff α is decreasing in θ, $\frac{d\alpha}{d\theta} = -\frac{g(\pi,\rho)}{D} < 0$. Consequently, the same is true for welfare, $\frac{dW}{d\theta} = \frac{\partial W}{\partial \alpha}\frac{d\alpha}{d\theta} < 0$. Hence, h is monotonically decreasing in θ. In total, we conclude that $\frac{d\rho^*}{d\theta} < 0$.

A.3.2. Proofs of propositions and lemmas extensions

Proof of Proposition 4.6: The cutoff $\bar{\theta}(\rho)$ is defined by $\bar{\theta} = \frac{c}{2g(\pi,\rho)}$. At this point a candidate with ability equal to zero is indifferent between proposing a reform or the status quo. $\bar{\theta}$ is decreasing in ρ. For $\rho \to 0$, we get that $g(\pi,\rho) \to 0$ implying $\bar{\theta} \to \infty$. Hence, for any given θ there is a ρ such that only informative equilibria can exist. Due to the monotonicity of $\theta(\bar{\rho})$, there is exactly one cutoff $\bar{\rho}(\theta)$, such that for all $\rho < \bar{\rho}(\theta)$ the unique divine equilibrium must be informative.

Proof of Proposition 4.8: For the case of limited commitment, the proofs of Proposition 1-5 need to be considered one by one. We shorten the argumentation, whenever it is equivalent or very similar to the case with full commitment. The proof of Proposition 4.1 does not rely on full commitment and thus conveys to the new setting.

Proof of Lemma 4.1 with limited commitment: We only need to deal with the case of a politician with ability lower than c, since candidates with ability above c always choose to reform. Limited commitment changes the payoff from entering office from $\theta+a-c$ to $\theta+\lambda(a-c)$. The rest of the proof is equivalent to the case with full commitment.

Proof of Proposition 4.2 with limited commitment: We just need to prove symmetry of cutoffs. The proof with regard to the classification of equilibria is identical to the case with full commitment. In equilibrium, the reform incentive with limited commitment simplifies to:

$$
\begin{aligned}
&R_1(\alpha_1,\alpha_1,\alpha_2) = \\
&\theta\left[\Phi(\alpha_2)g(\pi_1,\rho) + (1-\Phi_2)\left[g(\pi_1-\pi_2,\rho) + g(\pi_2,\rho)\right]\right] + \\
&\lambda(\alpha_1 - c)\left[\frac{1}{2} + \Phi(\alpha_2)g(\pi_1,\rho) + (1-\Phi(\alpha_2))g(\pi_1-\pi_2)\right] = 0 \\
&\Leftrightarrow \frac{\theta\left[\Phi(\alpha_2)g(\pi_1,\rho) + (1-\Phi(\alpha_2))\left[g(\pi_1-\pi_2,\rho) + g(\pi_2,\rho)\right]\right]}{\lambda(c-\alpha_1)} = \\
&\frac{1}{2} + \Phi(\alpha_2)g(\pi_1,\rho) + (1-\Phi(\alpha_2))g(\pi_1-\pi_2).
\end{aligned}
$$

Subtracting the corresponding equation for the second player and proceeding as in the proof with full commitment, we obtain

$$\Leftrightarrow \left[\frac{\theta\Phi(\alpha_2)}{\lambda(c-\alpha_1)} - \frac{\theta(1-\Phi(\alpha_1))}{\lambda(c-\alpha_2)} - \Phi(\alpha_2)\right] g(\pi_1,\rho) +$$
$$\underbrace{\left[(1-\Phi(\alpha_2))\left(\frac{\theta}{\lambda(c-\alpha_1)}-1\right) + (1-\Phi(\alpha_1))\left(\frac{\theta}{\lambda(c-\alpha_2)}-1\right)\right] g(\pi_1-\pi_2,\rho)}_{>0} =$$
$$\left[\frac{\theta\Phi(\alpha_1)}{\lambda(c-\alpha_2)} - \frac{\theta(1-\Phi(\alpha_2))}{\lambda(c-\alpha_1)} - \Phi(\alpha_1)\right] g(\pi_2,\rho).$$

If $\alpha_1 = \alpha_2$, this condition is trivially fulfilled. Assuming wlog $\alpha_1 > \alpha_2$, the equality above implies that

$$\left[\frac{\theta\Phi(\alpha_2)}{\lambda(c-\alpha_1)} - \frac{\theta(1-\Phi(\alpha_1))}{\lambda(c-\alpha_2)} - \Phi(\alpha_2)\right] g(\pi_1,\rho) < \left[\frac{\theta\Phi(\alpha_1)}{\lambda(c-\alpha_2)} - \frac{\theta(1-\Phi(\alpha_2))}{\lambda(c-\alpha_1)} - \Phi(\alpha_1)\right] g(\pi_2,\rho).$$

However, we have $\pi_1 > \pi_2$. Moreover, we can show that

$$\frac{\theta}{\lambda(c-\alpha_1)}\Phi(\alpha_2) - \frac{\theta}{\lambda(c-\alpha_2)}(1-\Phi(\alpha_1)) - \Phi(\alpha_2) \quad > \quad \frac{\theta}{\lambda(c-\alpha_2)}\Phi(\alpha_1) - \frac{\theta}{\lambda(c-\alpha_1)}(1-\Phi(\alpha_2)) - \Phi(\alpha_1)$$
$$\Leftrightarrow \frac{\theta}{\lambda(c-\alpha_1)} + \Phi(\alpha_1) \quad > \quad \frac{\theta}{\lambda(c-\alpha_2)} + \Phi(\alpha_2).$$

Thus, the reform incentive functions R_1 and R_2 can never simultaneously attain zero for $\alpha_1 > \alpha_2$, implying that there can only be symmetric equilibria.

Existence: The reform incentive function simplifies to

$$R(\alpha,\rho) = \left[\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right]\lambda(\alpha-c) + \theta g(\pi,\rho) = 0.$$

Note that it is always positive if $\alpha = 1$. If $R(0,\rho) < 0$, the reform incentive is equal to zero at least once, due to the continuity and there exists an interior equilibrium. If $R(0,\rho) \geq 0$, it is an equilibrium that all candidates choose to reform. Hence, there is at least one equilibrium.

Proof of Proposition 4.3 with limited commitment: Next, we establish uniqueness. The derivative with respect to α is:

$$\frac{\partial R}{\partial \alpha} = \underbrace{(\theta + (\alpha-c)\lambda\Phi(\alpha))g_\pi(\pi,\rho)\frac{\partial \pi}{\partial \alpha}}_{A} + \underbrace{\lambda\left(\frac{1}{2} + (\Phi(\alpha) + (\alpha-c)\phi(\alpha))g(\pi,\rho)\right)}_{B}.$$

The remainder of the proof is equivalent to the case with perfect commitment.

Proof of Proposition 4.4 with limited commitment: We use implicit differentiation to prove the proposition:

$$\frac{d\alpha}{d\rho} = -\frac{\frac{\partial R}{\partial \rho}}{\frac{\partial R}{\partial \alpha}} = -\frac{(\theta + \lambda(\alpha-c)\Phi(\alpha))g_\rho(\pi,\rho)}{(\theta + (\alpha-c)\lambda\Phi(\alpha))g_\pi(\pi,\rho)\frac{\partial\pi}{\partial\alpha} + \lambda\left(\frac{1}{2} + (\Phi(\alpha) + (\alpha-c)\phi(\alpha))g(\pi,\rho)\right)} < 0.$$

While the numerator is unambiguously positive, the positive sign of the denominator follows from Assumption 4.1.

Proof of Proposition 4.5 with limited commitment: The welfare can be simplified to:

$$\frac{W(\rho)}{2} = \underbrace{\left(\lambda \int_{\alpha}^{1} \phi(a)(a-c)da + (1-\lambda) \int_{c}^{1} \phi(a)(a-c)da\right)}_{z(\alpha)} \left(\frac{1}{2} + \Phi(\alpha) g(\pi,\rho)\right).$$

The total derivative of the welfare function can be simplified along the same lines as with full commitment and yields the same equilibrium condition

$$h(\rho) = \frac{\Phi(\alpha)}{\phi(\alpha)(c-\alpha)} - \left(1 + \frac{\theta}{W(\rho)}\right) = 0.$$

Thus, the rest of the proof is equivalent.

For the second step, we have to show how the unique maximum changes with θ. For $\theta \to 0$ we get from the reform incentive $\alpha = c$ and

$$\left.\frac{dW(\rho)}{d\rho}\right|_{\theta=0} = \frac{g_\rho(\pi,\rho)}{D} \Phi(\alpha) \lambda W(\rho).$$

This is positive. Hence, for $\theta \to 0$ the optimal institution is full power concentration. Due to continuity we get that this is also true for an interval around 0. Since $h(\rho,\alpha)$ does not change with limited commitment we again refer the reader to the proofs for full commitment to see that the optimal ρ is monotonically decreasing in θ.

Proof of Proposition 4.9: In all proofs of Appendix A.3.1 we use only one important feature of the vote share $v_i(x_1, x_2, \sigma)$. This is, the vote share is weakly increasing in the expected abilities of the candidates and thus in the difference of welfare contributions. In the following we show that this still holds for the case of heterogeneous policy preferences. All other proofs do not change. In the new setting, voter i votes for candidate 1 if:

$$x_1(\mu_{1i}\hat{a_1} - c) \geq x_2(\mu_{2_i}\hat{a_2} - c)$$

If both propose a reform, the vote share for candidate one is:

$$v_1(x_1 = 1, x_2 = 1, \sigma) = \int_k^l \int_k^{\frac{\mu_1 \hat{a_1}}{\hat{a_2}}} \xi(\mu_2)\xi(\mu_1).$$

The derivative with respect to $\hat{a}_1$ is positive, since we assume the mean of the preferences to be larger than zero. If candidate 1 faces a status quo proposing opponent, his expected vote shares is

$$v_1(x_1 = 1, x_2 = 0, \sigma) = \int_{\frac{c}{\hat{a_1}}}^{l} \xi(\mu_1).$$

The derivative with respect to the expected ability of Candidate 1 is again positive. In the third case, where Candidate 1 proposes the status quo, the vote share does not depend on the expected ability, since the payoff does neither. Hence, the expected overall vote share of candidate i is weakly increasing in his expected competence and, thus, his welfare contribution.

Proof of Lemma 4.3: In an informative equilibrium, the expected utility of voter k with reform preference μ_k is given by

$$V(\mu_k,\rho) = 2\int_{\alpha}^{1} \phi(a)(\mu_k a - c)da \left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right).$$

It is strictly increasing in μ_k, and negative for any $\mu_k \leq c$. Its derivative with respect to power concentration follows as

$$\begin{aligned}
\frac{dV(\mu_k,\rho)}{d\rho} &= 2\Phi(\alpha)\frac{dg}{d\rho}\int_{\alpha}^{1}\phi(a)(\mu_k a - c)da \\
&\quad +2\left[\left(\phi(\alpha)g + \Phi(\alpha)\frac{dg}{d\rho}\right)\int_{\alpha}^{1}\phi(a)(\mu_k a - c)da \right.\\
&\quad \left. -\phi(\alpha)(c-\mu\alpha)\left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right)\right]\frac{d\alpha}{d\rho} \\
&= 2\left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right)(\Phi(\alpha) + \phi(\alpha - c))\frac{dg}{d\rho}\int_{\alpha}^{1}\phi(a)(\mu_k a - c)da \\
&\quad +2\left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right)\phi(\alpha)(c-\mu\alpha)\frac{d\alpha}{d\rho} \\
&= \frac{g_\rho}{D}\left\{[\Phi(\alpha) - \phi(\alpha)(c-\alpha)]\left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right)\int_{\alpha}^{1}\phi(a)(\mu_k a - c)da \right.\\
&\quad \left. -\phi(\alpha)(c-\mu\alpha)\frac{\theta}{2}\right\}
\end{aligned}$$

For any $\rho \leq \rho^*(\theta)$, the term $\Phi(\alpha) - \phi(c-\alpha)$ is positive by Proposition 4.5 and Assumption 4.2. Thus, the expected utility of every voter with $\mu_k < c$ is strictly decreasing in ρ on this interval. By a similar argument as used in Lemma 4.2, it can be shown that $V(\mu_k,\rho)$ has at most one minimizer. For the limit case $\rho \to \infty$, however, we find that $\frac{dV}{d\rho} \leq 0$. In this limit, we have $g(\pi,\rho) = 1$, which implies $\theta \geq [1 + \Phi(\alpha)](c-\alpha)$ and a negative sign of the bracket in the last line above. Thus, $V(\mu_k,\rho)$ is monotonically decreasing in ρ.

Proof of Proposition 4.10: The proof consists of two steps. First we show that there exists at least one maximum for some $\rho < \rho^*$, second we ensure that there can never be a maximum for any $\rho \geq \rho^*$. Note that the expected utility $V(\mu_k,\rho)$ is increasing in the reform preference μ_k. Due to the strict concavity of w, this directly implies that $w'(V(\mu_k,\rho)) > w'(V(\mu'_k,\rho))$ for any $\mu_k < \mu'_k$. Moreover, the marginal effect of ρ on the expected utilities is

$$\frac{d^2V(\mu_k,\rho)}{d\rho\, d\mu_k} = \frac{2g_\rho(\pi,\rho)}{D}\left[\left(\frac{1}{2} + \Phi(\alpha)g(\pi,\rho)\right)(\Phi(\alpha) + \phi(\alpha)(\alpha - c))\int_{\alpha}^{1} a\phi(a)da + \frac{\theta}{2}\alpha\phi(\alpha)\right].$$

Take any welfare function of an inequality averse society:

$$W_{IA}(\rho) = \int_{\underline{\mu}}^{\bar{\mu}} w(V(\mu_k,\rho))\xi(\mu_k)d\mu_k$$

Its derivative with respect to ρ is

$$\frac{dW_{IA}(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} w'(V(\mu_k,\rho))\frac{dV(\mu_k,\rho)}{d\rho}\xi(\mu_k)d\mu_k.$$

For the case of $\rho' < \rho^*$, the cross derivative $\frac{d^2V(\mu_k,\rho)}{d\rho\ d\mu_k}$ is larger than zero. All terms of it are always positive except for $(\Phi(\alpha) + \phi(\alpha - c))$. This however is positive for all $\rho' \leq \rho^*$ (see Proposition 4.5 and Assumption 4.2). The positive cross derivative yields

$$\frac{dW_{IA}(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} w'(V(\mu_k,\rho))\frac{dV(\mu_k,\rho)}{d\rho}\xi(\mu_k)d\mu_k < \frac{dW(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} \frac{dV(\mu_k,\rho)}{d\rho}\xi(\mu_k)d\mu_k.$$

The derivatives of the expected utility are smaller for voters with smaller μ. Exactly these utilities are weighted more strongly in the case of inequality aversion, since $w'(V(\mu_k,\rho)) > w'(V(\mu_k',\rho))$. Hence, the derivative of the welfare function at ρ^* is negative and there exists at least one local maximum for a $\rho' < \rho^*$.

Now consider the case of $\rho' > \rho^*$, where $\frac{dW}{d\rho} < 0$. If the cross derivative is still positive for any $\rho' > \rho^*$, we have that $\left.\frac{dW_{IA}}{d\rho}\right|_{\rho'} < \left.\frac{dW}{d\rho}\right|_{\rho'} < 0$ and there can not be a maximum at ρ'. Suppose that the cross derivative is negative at ρ'. From Lemma 4.3 we know that the marginal effect of ρ is negative for voters with $\mu < c$. As a consequence of the negative cross derivative, the marginal effect is negative for all agents. Thus, the derivative of the welfare function is certainly negative. Overall there cannot be any maximum in the range $[\rho^*, \infty)$.

A.3.3. Additional figures and tables

Description and sources of variables

Growth in real GDP per capita	Average growth rate. Self calculated based on GDP in 2000 US\$. World Bank.
GDP per capita	Denominated in thousand year 2000 US\$. World Bank.
Office motivation	International Social Survey Panel, module "Citizenship 1", 2004.
Power dispersion	Lijphart's index for executive-parties dimension. Lijphart (1999).
Checks and balances	Number of veto players. Keefer and Stasavage (2003).
Plurality electoral system	Dummy variable. Beck et al. (2001).
Electoral system	Type of electoral system, 9 minor categories. IDEA (2004).

Country list

Australia	Austria	Canada	Denmark
Finland	France	Germany	Ireland
Israel	Japan	Netherlands	Norway
Portugal	Spain	Sweden	Switzerland
United Kingdom	United States		

Summary of variables

	Mean	Std. dev.	Min	Max	Poss. values
Power dispersion	0.31	0.98	−1.21	1.77	[-2,2]
Office motivation	3.37	0.37	2.61	4.20	[1,5]
GDP p.c.	26.98	7.69	11.55	39.83	
GDP p.c. growth (2004-2011)	0.68	0.74	−0.61	2.40	
GDP p.c. growth (1991-2004)	2.08	1.07	0.56	5.59	

Correlation between variables

	Power dispersion	Office motivation	GDP p.c.	GDP p.c. growth (2004-2011)
Office motivation	−0.20	1		
	(0.43)			
GDP p.c.	0.20	−0.58	1	
	(0.43)	(0.01)		
GDP p.c. growth (2004-2011)	0.44	0.072	−0.15	1
	(0.07)	(0.78)	(0.56)	
GDP p.c. growth (1991-2004)	−0.27	−0.10	−0.021	−0.48
	(0.28)	(0.69)	(0.93)	(0.05)

Pearson's correlation coefficient, p-values in parentheses

Bibliography

Abadie, A. and S. Gay (2006). The impact of presumed consent legislation on cadaveric organ donation: A cross-country study. *Journal of Health Economics 25*, 599–620.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hyothesis and unemployment. *Quarterly Journal of Economics 105*, 255–283.

Albrecht, J. and S. Vroman (1992). Dual labor markets, efficiency wages, and search. *Journal of Labor Economics 10*(4), 438–461.

Armingeon, K. (2002). The Effects of Negotiation Democracy: A Comparative Analysis. *European Journal of Political Research 41*, 81 – 105.

Armingeon, K., R. Careja, P. Potolidis, M. Gerber, and P. Leimgruber (2011). *Comparative Political Data Set III 1990-2009*. Institute of Political Science, University of Berne.

Ashton, M. and K. Lee (2009). The hexaco-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment 91*, 340–345.

Atkinson, A. (1973). How progressive should income tax be? In M. Parkin and R. Nobay (Eds.), *Essays on Modern Economics*. London: Longman.

Bandiera, O., I. Barankay, and I. Rasul (2011). Field experiments with firms. *Journal of Economic Perspectives 25*(3), 63–82.

Banks, J. S. and J. Sobel (1987). Equilibrium selection in signaling games. *Econometrica 55*(3), 647 – 661.

Barro, R. J. (1973). The Control of Politicians: An Economic Model. *Public Choice 14*, 19–42.

Barro, R. J. (1991, May). Economic growth in a cross section of countries. *The Quarterly Journal of Economics 106*(2), 407–43.

Bartling, B., E. Fehr, and K. M. Schmidt (2012). Screening, competition, and job design: Economic origins of good jobs. *American Economic Review 102*(2), 834–864.

Beck, T., G. Clarke, A. Groff, P. Keefer, and P. Walsh (2001, September). New tools in comparative political economy: The database of political institutions. *World Bank Economic Review 15*(1), 165–176.

Besley, T. (2005, Summer). Political selection. *Journal of Economic Perspectives 19*(3), 43–60.

Besley, T. and M. Smart (2007, April). Fiscal Restraints and Voter Welfare. *Journal of Public Economics 91*(3-4), 755–773.

Bewley, T. (1999). *Why Wages Don't Fall During a Recession.* Cambridge, MA.: Harvard University Press.

Blume, L., J. Müller, S. Voigt, and C. Wolf (2009, April). The economic effects of constitutions: replicating-and extending-persson and tabellini. *Public Choice 139*(1), 197–225.

Board, S. and M. Meyer-ter-Vehn (2011). Relational contracts in competitive labor markets. *Working Paper, UCLA.*

Brown, C. and J. Medoff (1989). The employer size-wage effect. *Journal of Political Economy 97*(5), 1027–1059.

Brown, C. L. and A. Krishna (2004). The skeptical shopper: A metacognitive account for the effects of default options on choice. *Journal of Consumer Research 31*(3), 529–539.

Brown, M., A. Falk, and E. Fehr (2004, 5). Relational contracts and the nature of market interactions. *Econometrica 72*(3), 747–780.

Brown, M. and M. Serra-Garcia (2012). Debt enforcement and relational contracting. *Working Paper, University of Munich.*

Brown, M. and C. Zehnder (2007). Credit reporting, relationship banking, and loan repayment. *Journal of Money, Credit and Banking 39*(8), 1883–1918.

Bulow, J. I. and L. H. Summers (1986). A theory of dual labor markets with application to industrial policy, discrimination, and keynesian unemployment. *Journal of Labor Economics 4*(3), 376–414.

Cain, D. M., G. Loewenstein, and D. Moore (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *The Journal of Legal Studies 34*(1), 1–25.

Callander, S. (2008, 07). Political motivations. *Review of Economic Studies 75*(3), 671–697.

Callander, S. and S. Wilkie (2007). Lies, Damned Lies, and Political Campaigns. *Games and Economic Behavior 60*(2), 262 – 286.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review 151*, 1211–1254.

Campbell III, C. M. and K. S. Kamlani (1997). The reasons for wage rigidity: Evidence from a survey of firms. *Quarterly Journal of Economics 112*(3), 759–789.

Caplin, A. and D. Martin (2011). A testable theory of imperfect perception. *Working Paper, NYU*.

Caplin, A. and D. Martin (2012). Defaults and attention: The drop out effect. *Working Paper, NYU*.

Cappelli, P. and K. Chauvin (1991). An interplant test of the efficiency wage hypothesis. *Quarterly Journal of Economics 106*(3), 769–787.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. *The Quarterly Journal of Economics 124*(4), 1639–1674.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

Charness, G. and P. Kuhn (2007). Does pay inequality affect worker effort? experimental evidence. *Journal of Labor Economics 25*, 693–723.

Charness, G. and P. J. Kuhn (2011). Lab labor: What can labor economists learn from the lab? In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 229–330. Amsterdam: Elsevier.

Choi, J., D. Laibson, and B. C. Madrian (2004). Plan design and 401(k) savings outcomes. *National Tax Journal 57*(2), 275–298.

Cohn, A., E. Fehr, and L. Götte (2012). Fair wages and effort: Evidence from a field experiment. *Working Paper, University of Zurich*.

Crawford, V. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1432–1451.

Doeringer, P. B. and M. J. Piore (1971). *Internal Labor Markets and Manpower Analysis*. Lexington, MA: Heath Lexington Books.

Downs, A. (1957). *"An Economic Theory of Democracy"*. New York: Harper.

Enikolopov, R. and E. Zhuravskaya (2007, December). Decentralization and political institutions. *Journal of Public Economics 91*(11-12), 2261–2290.

Erat, S. and U. Gneezy (2012). White lies. *Management Science forthcoming*.

Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior 54*(2), 293–315.

Fehr, E. and S. Gächter (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives 14*, 159–181.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics 108*, 437–460.

Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition and Cooperation. *The Quarterly Journal of Economics 114*, 817–868.

Feld, L. P. and S. Voigt (2003, September). Economic growth and judicial independence: cross-country evidence using a new set of indicators. *European Journal of Political Economy 19*(3), 497–527.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1), 5 – 25.

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

Fischbacher, U. and F. Heusi (2008). Lies in disguise—an experimental study on cheating. *Working Paper, University of Zurich*.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives 19*(4), 25–42.

Fudenberg, D. and E. Maskin (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica 54*(3), 533–554.

Gächter, S., D. Nosenzo, and M. Sefton (2012). The impact of social comparisons on reciprocity. *Scandinavian Journal of Economics 114*, 1346–1367.

Glaeser, E. L. (2005). Paternalism and psychology. *Harvard Institute of Economic Research, Discussion Paper 2097*.

Greiner, B. (2003). An Online Recruitment System for Economic Experiments. In K. Kremer and V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, pp. 79–93. Gï¿½ttingen: Ges. fï¿½r Wiss. Datenverarbeitung.

Hellwig, M. F. (2005). A utilitarian approach to the provision and pricing of excludable public goods. *Journal of Public Economics 89*, 1981–2003.

Huck, S., A. Seltzer, and B. Wallace (2011). Deferred compensation in multiperiod labor contracts: An experimental test of lazear's model. *American Economic Review 101*(2), 819–843.

Ichino, A. and R. Riphahn (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association 3*(1), 120–143.

IDEA, I. (2004). *Handbook of Electoral System Design*. International Institute for Democracy and Electoral Assistance.

Johnson, E. J. and D. Goldstein (2003). Do defaults save lives? *Science 302*, 1338–1339.

Johnson, E. J., J. Hershey, J. Meszaros, and H. Kunreuther (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty 7*, 35–51.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991). The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives 5*(1), 193–206.

Kartik, N. and R. P. McAfee (2007, June). Signaling Character in Electoral Competition. *American Economic Review 97*(3), 852–870.

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation. In S. Fischer (Ed.), *NBER Macroeconomic Annual*. Cambridge, MA: MIT Press.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2009). Governance matters VIII : aggregate and individual governance indicators 1996-2008. World Bank Policy Research Working Paper No. 4978.

Keefer, P. and D. Stasavage (2003, August). The limits of delegation: Veto players, central bank independence and the credibility of monetary policy. *American Political Science Review 97*(3), 407–423.

King, A. (2002). Do leaders' personalities really matter? In A. King (Ed.), *Leaders' Personalities and the Outcomes of Democratic Elections*, pp. 1–42. Oxford: Oxford University Press.

Levav, J., M. Heitmann, A. Herrmann, and S. S. Iyengar (2010). Order in product customization decisions: Evidence from field experiments. *Journal of Political Economy 118*(2), 274–299.

Lijphart, A. (1999). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.

Lijphart, A. (2012). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries* (2 ed.). New Haven: Yale University Press.

Linardi, S. and C. Camerer (2010). Can relational contracts survive stochastic interruptions? *Social Science Working Paper No. 1340, California Institute of Technology*.

Lindbeck, A. and J. W. Weibull (1987). Balanced-Budget Redistribution as the Outcome of Political Competition. *Public Choice 52*(3), 273–297.

Lizzeri, A. and N. Persico (2001, March). The Provision of Public Goods under Alternative Electoral Incentives. *American Economic Review 91*(1), 225–239.

Lizzeri, A. and N. Persico (2005, December). A Drawback of Electoral Competition. *Journal of the European Economic Association 3*(6), 1318–1348.

MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica 57*(2), 447–480.

MacLeod, W. B. and J. M. Malcomson (1998). Motivation and markets. *American Economic Review 88*(3), 388–411.

Madison, J. (1788a). Federalist No. 51, The Structure of the Government Must Furnish the Proper Checks and Balances Between the Different Departments. accessed through `http://thomas.loc.gov/home/histdox/fed\_51.html`.

Madison, J. (1788b). Federalist No. 57, The Alleged Tendency of the New Plan to Elevate the Few at the Expense of the Many Considered in Connection with

Representation. accessed through http://thomas.loc.gov/home/histdox/fed_57.html.

Madrian, B. C. and D. F. Shea (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economics 116*, 1149–1187.

Majumdar, S. and S. W. Mukand (2004, September). Policy Gambles. *American Economic Review 94*(4), 1207–1222.

Marshall, M. G. and K. Jaggers (2010). Polity IV Project: Dataset Users' Manual.

Maskin, E. and J. Tirole (2004, September). The Politician and the Judge: Accountability in Government. *American Economic Review 94*(4), 1034–1054.

McKenzie, C. R., M. J. Liersch, and S. R. Finkelstein (2006). Recommendations implicit in policy defaults. *Psychological Science 17*(5), 414–420.

Pancer, S. M., S. D. Brown, and C. W. Barr (1999). Forming impressions of political leaders: A cross-national comparison. *Political Psychology 20*(2), 345–368.

Persson, T. and G. Tabellini (2003). *The Economic Effects of Constitutions*. Cambridge: The MIT Press.

Persson, T. and G. Tabellini (2004, March). Constitutional rules and fiscal policy outcomes. *American Economic Review 94*(1), 25–45.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*(5), 1281–1302.

Saint-Paul, G. (1996). *Dual Labor Markets – A Macroeconomic Perspective* (1 ed.). Cambridge, MA: MIT Press.

Sala-i-Martin, X. (1994, April). Cross-sectional regressions and the empirics of economic growth. *European Economic Review 38*(3-4), 739–747.

Sala-i-Martin, X. (1997, May). I just ran two million regressions. *American Economic Review 87*(2), 178–83.

Samuelson, W. and R. Zeckhauser (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty 1*(1), 7–59.

Schwartz, A. and R. E. Scott (2003). Contract theory and the limits of contract law. *Yale Law Journal 113*, 541–619.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review 74*(3), 433–444.

Smart, M. and D. M. Sturm (2006, December). Term Limits and Electoral Accountability. CEP Discussion Papers dp0770, Centre for Economic Performance, LSE.

Stokes, D. E., A. Campbell, and W. E. Miller (1958). Components of electoral decision. *The American Political Science Review 52*(2), pp. 367–387.

Sunstein, C. R. (2002). Switching the default rule. *New York University Law Review 77*, 106–134.

Sunstein, C. R. (2012a). Impersonal default rules vs. active choices vs. personalized default rules: A triptych. *Working Paper*.

Sunstein, C. R. (2012b). The storrs lectures: Behavioral economics and paternalism. *Yale Law Journal* (forthcoming).

Tannenbaum, D. (2011). Information asymmetries in policy defaults: Some defaults speak louder than others. *Working Paper, University of California, Irvine*.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American Economic Review 93*, 175–179.

Tsebelis, G. (2002). *Veto Players. How Political Institutions Work*. Princeton University Press/Russell Sage Foundation.

Voigt, S. (2011). Positive constitutional economics ii – a survey of recent developments. *Public Choice 146*(1), 205–256.