Incentives and Information as Driving Forces of Default Effects[†]

Steffen Altmann[*], Armin Falk[‡], Andreas Grunewald[§]

February 4, 2015

Abstract

The behavioral relevance of non-binding defaults is well established. While most research has focused on decision makers' responses to a given default, we argue that this individual decision making perspective is incomplete. Instead, a comprehensive understanding of default effects requires to take account of the strategic interaction between default setters and decision makers. We analyze theoretically and empirically which defaults emerge in such interactions, and under which conditions defaults are behaviorally most relevant. Our analysis demonstrates that the alignment of interests between default setters and decision makers, as well as their relative level of information are key drivers of default effects. In particular, default effects are more pronounced if the interests of the default setter and decision makers are more closely aligned. Moreover, decision makers are more likely to follow default options the less they are privately informed about the relevant decision environment.

JEL classification: D03, D18, D83, C92

Keywords: default options, behavioral economics, strategic communication, laboratory experiment

[†] Financial Support from Volkswagen Foundation and the Deutsche Forschungsgesellschaft (SFB/TR 15) is gratefully acknowledged.

[*]University of Copenhagen and IZA, Øster Farimagsgade 5, building 26, 1353 København K, Denmark. Email: steffen.altmann@econ.ku.dk

[‡]University of Bonn and IZA, Adenauerallee 24-42, 53113 Bonn, Germany. Email: armin.falk@uni-bonn.de

[§] University of Bonn, Kaiserstr. 1, 53113 Bonn, Germany. Email: gruni@uni-bonn.de

1 Introduction

A substantial body of empirical research has shown that non-binding default options can strongly affect consumption and saving decisions. Default effects have been documented, for instance, in decisions on retirement saving (Madrian and Shea 2001, Beshears *et al.* 2008, Chetty *et al.* 2014), participation in workplace training (Borghans and Golsteyn 2014), consumers' choices of insurance coverage, product specifications, and utility contracts (Johnson *et al.* 1993, Levav *et al.* 2010, Ebeling 2013), or consent to postmortem organ donations (Johnson and Goldstein 2003). To explain why default effects arise in these contexts, the literature has mainly focused on the preferences and characteristics of the decision maker, relating individuals' tendency to stick to defaults to factors such as status-quo biases, quasi-hyperbolic discounting, or limited attention (e.g., Samuelson and Zeckhauser 1988, Carroll *et al.* 2009, Caplin and Martin 2013).

In this paper, we argue that this individual decision making perspective is not sufficient to fully understand the behavioral consequences of defaults. Instead, default effects commonly arise in situations where default setters—firms, organizations, or governmental agencies—deliberately specify the default option. Defaults are thus the outcome of an inherently strategic interaction between the default setter and a population of decision makers. As a consequence, the characteristics of both the default setter and the decision makers might influence how defaults are being set, and to what extent they affect people's choices. By embracing their strategic nature, it also becomes apparent that defaults can constitute a *mode of communication* between default setters and decision makers. To formalize our arguments, we therefore apply a well-established theoretical framework of strategic communication to the question of how default effects arise. Specifically, we adopt a simplified version of Crawford and Sobel (1982) in which we incorporate two-sided asymmetric information, to capture the notion that default setters as well as decision makers may differ in their knowledge of what is the optimal choice for the decision maker.

Analyzing default effects from the perspective of strategic communication directly reveals that the alignment of interests between default setters and decision makers, as well as their relative knowledge of the decision environment can have important consequences for default specifications and the strength of default effects in equilibrium. We demonstrate that both factors indeed play a key role for the emergence of default effects. Specifically, we show that in any Pareto-efficient equilibrium, defaults are set in a way that is more informative for decision makers if the interests between default setters and decision makers are more closely aligned. At the same time, the reaction of rational agents to a default option depends on its informative value, and is therefore itself related to the default setter's level of benevolence. Moreover, our theoretical analysis indicates that default effects should differ systematically across subgroups of the population: when defaults are at odds with individuals' own information, decision markers who are less knowledgeable about the choice environment are more likely to stick to defaults. In contrast, defaults do not affect choices of individuals who are particularly well-informed about the decision.

The idea that defaults can be behaviorally relevant since they convey information and are perceived as an implicit recommendation by the default setter has been informally discussed in the literature on default effects (e.g., Madrian and Shea 2001, Johnson and Goldstein 2003, McKenzie *et al.* 2006). We demonstrate that a well-established theoretical approach can be used to capture this intuition in a unifying formal framework. Doing so also helps to organize the existing empirical findings on default effects. For instance, McKenzie *et al.* (2006) as well as Tannenbaum (2011) provide survey evidence supporting the notion that consumers stick to defaults to follow the default setter's recommendation. On the other hand, Brown and Krishna (2004) have argued that "marketplace metacognition" makes consumers skeptical about defaults set by profit maximizing firms; as a consequence, they should more heavily rely on active choices. Our framework addresses both points of view, showing that the informational content of default options and the extent to which consumers account for this information both depend on the degree to which the interests of default setters and decision makers are aligned.

Our results on differential reactions to defaults, depending on decision makers' own knowledge of the choice environment, are also consonant with a number of stylized facts from the empirical literature. For instance, Bronchetti *et al.* (2013) find that saving defaults have essentially no effect on behavior for individuals with strong prior consumption plans. Similarly, in an experiment with environmental economists, Löfgren *et al.* (2012) observe no impact of defaults in a carbon-offsetting program and attribute this finding to their participants' high level of experience. The observation that less informed agents are more prone to follow defaults is also made by Brown *et al.* (2011), who find that employees' propensity to stick to the default savings plan is negatively associated with their overall financial literacy. Furthermore, in a survey among the plan participants, 50% of those who remained in the default plan mentioned decision complexity or other information-related issues as a main reason to do so. Finally, Levav *et al.* (2010) observe that default effects in product customization decisions are more pronounced for consumers with a lower self-rated knowledge of the manufacturer's products. On a more general level, our results might also help to understand why defaults sometimes have no aggregate effect on behavior, or why their effects are confined to relatively small sup-groups of the population (e.g., Beshears *et al.* 2010, Altmann *et al.* 2014).

While a considerable body of evidence is thus consistent with the notion that incentives and information matter for default effects, direct causal evidence on the role of this channel has—to the best of our knowledge—remained scarce. In the second part of this paper, we take a step towards closing this gap with the help of a laboratory experiment in which we exogenously vary the alignment of interests as well as the level of information for default setters and decision makers. In the experiment, we study a simple binary-choice paradigm in which decision makers ("agents") have to decide whether a set of nine cards, which can be either red or black, contains more red cards or more black cards. Before making choices, the default setter and the agents receive independent and informative signals on the agents' payoffs from the two available choice options. The default setter selects one of the options as the default and the agents, in turn, can accept the default option or make an active decision. To identify the causal effects of default setters' strategic incentives, we exogenously vary whether preferences of the default setters 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 default setters and agents. In particular, there are always some agents who are better informed than the default setter, and some agents whose information quality is below that of default setters.

Our empirical results underline the importance of strategic incentives and private information for default setters' and decision makers' behavior. First, defaults truthfully reveal default setters' information in 98% of cases in FUL, but only in 75% of cases in PAR and in 56% in MIS. Default setters with more closely aligned interests thus select informative defaults, while defaults specified by default setters with misaligned preferences barely convey any information. Second, agents are substantially more likely to accept defaults that are chosen by more benevolent default setters, 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 default options. At the same time, decisions of agents with superior information are barely affected by defaults in either of the treatments: in particular, well-informed agents almost always opt out when the default is in conflict to their own information. Finally, our results suggest that defaults set by fully or partially benevolent default setters can in fact enhance the aggregate quality of agents' decisions, lending support to a common theme in the literature on "libertarian paternalism" (Thaler and Sunstein 2003, Sunstein 2012). In particular, agents with lower levels of information benefit substantially from accepting default options in the FUL and PAR treatment. At the same time, default options do not distort choices of well-informed agents. However, we also find that the effects on agents' welfare are more mixed under misaligned incentives. In particular, agents with intermediate levels of information tend to follow defaults somewhat too frequently, with detrimental consequences for their payoffs.

Our theoretical framework as well as our empirical setup abstract from a number of factors that have been discussed as potential sources of default effects. These include status-quo effects and loss aversion (Samuelson and Zeckhauser 1988, Kahneman *et al.* 1991), pecuniary and non-pecuniary costs of opting out of defaults (Schwartz and Scott 2003, Thaler and Sunstein 2003), quasi-hyperbolic discounting and procrastination of active decisions (Madrian and Shea 2001, Carroll *et al.* 2009), or limited attention of decision makers (Caplin and Martin 2013). While abstracting from these individual-level motives helps us to isolate the strategic aspects of default effects, this does not mean that we consider them of secondary importance. Instead, the strategic aspects of defaults should influence the overall strength of default effects independently of the other psychological factors that are at play. A loss-averse consumer, for example, should also take his experience and knowledge of the decision environment into account when deciding on whether or not to stick to a default. Similarly, consumers should consider the intentions of default setters in a given choice setting even if they generally tend to postpone active decisions. An important feature of our theoretical framework is that it is flexible enough to be subsequently enriched by these factors.

Because defaults and strategic communication are intimately linked in our setup, our findings also contribute to the existing body of empirical research on communication games. A number of papers in this literature has investigated how communication depends on the alignment of interests (see Dickhaut *et al.* 1995, Gneezy 2005, Cai and Wang 2006, Sanchez-Pages and Vorsatz 2007; see also Crawford 1998 for an overview of the early empirical literature).

Focusing on the case of a fully informed sender who interacts with a single uninformed receiver, these studies have documented two main sets of results. First, it is typically found that closer alignment of interests leads to more information transmission. Second, senders tend to reveal somewhat more information than theoretically predicted, which can be explained by individual-level differences in strategic reasoning (e.g., *level-k* models as in Crawford 2003, Wang *et al.* 2010, or Crawford *et al.* 2013), genuine preferences for honest behavior (Gneezy 2005, Vanberg 2008), or a combination of the two. Our experiment shows that both of these key results extend to the case of a partially informed sender—the default setter—who interacts with a population of agents that differ in the quality of their information.[1] Our data further suggest that excess information revelation in our setting is primarily driven by a subgroup of default setters with relatively strong intrinsic preferences for honesty (as measured by a subscale of the HEXACO personality inventory; see Ashton and Lee 2009). At the same time, however, our findings also point at the relevance of boundedly rational behavior. In particular, subjects with low degrees of cognitive reflection (Frederick 2005) are significantly more likely to follow defaults from default setters with misaligned interests if these are at odds with their own information.

The remainder of the paper is organized as follows. In the next section, we formalize the idea that defaults can be a means of strategic communication between default setters and decision makers. We also derive testable implications for the experiment which we present in Sections 3 and 4. Section 5 concludes.

2 A communication perspective on default effects

In this section, we want to formalize the intuition that the nature of default effects depends on the strategic incentives of default setters and decision makers, as well as the parties' relative level of information about the decision environment. Starting out from the idea that defaults can constitute a mode of communication between default setters and decision makers, our discussion in this section will be guided by the classic framework of strategic communication by Crawford and Sobel (1982). Their work, as well as much of the following literature on strategic communication, however, considers a fully informed sender who interacts with a sin-

[1] Complementary evidence on strategic communication with multiple receivers is provided by Battaglini and Makarov (2014) and Drugov *et al.* (2013). Their focus, however, is on a fully informed sender who interacts with two uninformed receivers that differ in their preferences. In contrast, we focus on heterogeneity in agents' information, while holding the alignment of preferences constant within a given treatment.

gle, non-informed receiver. In contrast, defaults are typically observed in situations in which a single default setter interacts with a population of decision makers. Moreover, a common feature of choice environments with defaults is that decision makers have at least partial information about the decision at hand; the decision-relevant information might, however, be heterogeneously distributed in the population. Finally, it is often argued that the default setter herself might not be fully informed about the decision environment (e.g., Sunstein 2012, Sunstein 2013). To capture these feature, we therefore present a simplified version of Crawford and Sobel (1982) in which we incorporate two-sided asymmetric information, as well as a population of decision makers that is characterized by heterogeneous information quality.[2] All proofs can be found in Section A.1 in the appendix.

2.1 Model setup

In our model, one default-setter (i.e., the "sender" in the communication game) interacts with a population of rational decision makers (the "agents" or "receivers"). The decision makers choose an action z in order to maximize their utility $U_A(z, \theta)$, which is strictly concave in z for all θ. Agents' preferences over the choice alternatives depend on the state of the world, θ, which takes value θ_h or θ_l with equal probability. More precisely, we assume that the partial derivative of $U_A(z, \theta_l)$ with respect to z is smaller than the corresponding derivative of $U_A(z, \theta_h)$ for all z. Higher zs are thus optimal for agents that put more probability weight on θ_h. Decision makers choose between a finite number of options $z \in \{z_1, \ldots, z_m\}$ with $z_i < z_{i+1}$, where the choice set is rich enough such that the maximizer of $U_A(z, \theta_l)$ differs from the one of $U_A(z, \theta_h)$.[3]

[2] A number of other papers have recently studied different aspects of strategic communication in settings with two-sided asymmetric information (Chen 2009, Moreno de Barreda 2013, Lai 2014). Chen (2009) shows that equilibrium communication in such settings can be non-monotone, i.e., senders may communicate if the state is "extreme" or "intermediate" instead of "high" vs. "low". Moreover, she finds that the receiver cannot credibly reveal her signal when communicating to the sender first. Moreno de Barreda (2013) and Lai (2014) demonstrate that enhanced private information of the receiver can lead to a decrease in the level of communication and potentially even reduce the decision maker's welfare, relative to a benchmark with a fully informed sender and a single uninformed receiver. In contrast, our focus is on the interaction between a partially informed sender (the default setter) and a heterogeneous population of receivers (the decision makers). Moreover, we want to allow for situations in which a fraction of the population can be strictly better informed about the decision than the default-setting institution.

[3] Letting the agents choose from a continuous action space does not alter our theoretical results. However, default options are frequently observed in choice environments that entail only a finite number of alternatives.

While the true state of the world is ex ante unobservable, the default setter and the decision makers receive incomplete private information about it. The quality of agents' information may differ within the population. In particular, decision makers are distributed according to a "knowledge function" $f(x)$, 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 distributions $p(\sigma_l|\theta_l) = p(\sigma_h|\theta_h) = x$. The default setter also has private information about the optimal decision for the agents. Her signal is drawn independently from the ones of the agents and is denoted by $\rho \in \{\rho_l, \rho_h\}$ with signal quality $q = p(\rho_l|\theta_l) = p(\rho_h|\theta_h) \in (\frac{1}{2}, 1)$. Consequently, the default setter is always better informed about the state of the world than some of the decision makers, and worse than others.

While the default setter cannot directly influence the agents' decisions, she can determine a default option $d \in \{d_1, \ldots, d_n\}$ to which agents' are exposed when making their choices. Defaults in our setup are thus a mode of communication through which the default setter may transmit her private knowledge about what is optimal for the agents.[4] The default setter's utility resulting from choice z of any agent in the population is a weighted sum of the agent's utility, $U_A(\theta, z)$, and a term $b(z)$ that captures a potential conflict of interests.

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

For the overall objective of the default setter, we assume for simplicity that she puts equal weight on the utility (received from the decisions) of different agents.[5] The parameter μ is common knowledge and describes the alignment of interests between the default setter and the decision makers. For small μ, the default setter cares only little about the agents' well-being and focuses more strongly on her private interests. For large μ, the preferences of the default setter and the agents are more closely aligned. μ thus captures the notion that defaults are observed in a variety of contexts in which the strategic interests of the default-setting institutions may differ. For example, the alignment of interests may be stronger

For instance, consumers can choose between the default energy contract or a "green" alternative, they can stick to the default car configuration or choose from a set of optional features, etc. (Ebeling 2013, Levav *et al.* 2010).

[4]An interesting follow-up question, which is beyond the scope of this paper, would be to analyze the conditions under which a sender decides to communicate through a default rather than through other possible modes of communication (e.g., direct recommendations, advertisement, frames, or other cheap-talk messages).

[5]It is straightforward to extend the model to incorporate a weighting function with heterogeneous weights for different subgroups of the population.

between a governmental agency and consumers than it is the case for a profit maximizing firm that specifies default product configurations for her customers. To put structure on the default setter's preferences, we assume an upward bias in the sense that $b(z)$ is strictly increasing in z. Hence, the utility-maximizing choice of z for the default setter is weakly higher than the one for the agents.

The game is divided into four stages. In stage 1 nature draws the state of the world, θ, and all private signals. Moreover, all agents and the default setter observe their corresponding signal. In stage 2, the default setter determines the default option d. In stage 3, the default is revealed to the agents, and they decide individually on which choice z to implement. Dependent on the choice and the true state of the world, payoffs for agents and the default setter are realized in stage 4. In this setup, a strategy of the default setter $s_P(\rho)$ specifies the probability of every default to be chosen for all possible signals ρ. A pure strategy $s^x_A(\sigma, d)$ of decision maker x determines for any combination of a private signal σ and a default d the choice of z.[6] As a solution concept we apply perfect Bayesian equilibrium (PBE).

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 replicated by an output equivalent equilibrium without out-of-equilibrium defaults. More precisely, for every PBE there exists another PBE that involves playing all feasible defaults such that the ex ante probabilities of any utility outcome for the agents and the default setter are identical.[7]

Before agents decide on z, the default setter can try to induce a more favorable outcome to herself by appropriately specifying the default option. Whenever the default setter receives a high signal, her preferences are aligned with those of the agents. In this case, the message that most certainly signals the high state of the world maximizes both her selfish interests $b(z)$ and the utility of the decision makers. As a consequence, all defaults that are set with positive probability after a high signal must have the same informational content and are therefore qualitatively identical. The remaining defaults, in contrast, reveal that the default setter received a low signal. According to this distinction we can split the set of

[6] In general, 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 actions has mass zero. Hence, extending the analysis to mixed strategies does not provide any additional insights.

[7] The construction of the output equivalent perfect Bayesian equilibrium can be found in the appendix.

defaults $\{d_0, \ldots, d_n\}$ into two parts. Denote by $\{d_0, \ldots d_{k-1}\} = D_l$ the set of defaults that are exclusively played after a low signal and by $\{d_k \ldots, d_n\} = D_h$ the set of defaults that are played with positive probability after a high signal. For simplicity, we call each default in D_h a "high default" (d_h) and those in D_l "low defaults" (d_l).

Lemma 1. *In equilibrium the default setter uses at most two qualitatively different defaults.*

Using the lemma we are able to characterize the amount of information that is transmitted through the default by a single parameter $c = p(d_0|\rho_l) + \cdots + p(d_{k-1}|\rho_l)$. The higher the transmission rate c, the higher is the likelihood that the default setter truthfully reveals having received a low signal by setting a low default, thereby leading to more information transmission in equilibrium. In the following, we say that an equilibrium is informative if the default has at least some informational content, i.e., if $c > 0$.

Alignment of interests

Intuitively, an agent's decision to follow a default may depend on the default setter's level of benevolence. A default setter with opposing interests may try to use her private information to extract a higher rent for herself. Anticipating this, decision makers should rely less strongly on the defaults specified by such a default setter. In contrast, a default setter whose interests are more closely aligned with those of the agents may specify a default that is more informative for the agents; her defaults should thus be more likely to be followed. The following result corroborates this intuition. Note that, in general, there may emerge multiple equilibria exhibiting different levels of information transmission. In particular, there always exist babbling equilibria in which the default does not convey any information. To refine the set of equilibria, we employ the Pareto-efficiency criterion and focus on the question how an informative Pareto-efficient equilibrium depends on the model's main parameters.

Proposition 1. *For informative Pareto-efficient equilibria, larger μ imply a higher information transmission rate $c(\mu)$. Agents internalize this effect and exhibit more pronounced reactions to defaults the higher is μ, i.e., $s^x_A(\sigma_i, d_h)$ is increasing and $s^x_A(\sigma_i, d_l)$ is decreasing in μ.*

The proposition mirrors the main insight of Crawford and Sobel (1982) for our setting with a binary state space and a heterogeneously informed population of agents. The underlying intuition is straightforward. With increasing μ the default setter's preferences become more closely aligned to those of the agents. As a consequence, her utility of a truthful report increases relative to the utility of concealing her private information. This leads to a higher information transmission rate c. The higher informational content of the default yields, *ceteris paribus*, a stronger incentive for decision makers to adapt their choice towards the default. In particular, in case the default is in conflict with an agent's private signal, the agent weighs the default more heavily for higher μ. Hence, default effects are stronger in a given population if the interests between the population and the default setter are more closely aligned.

Finally, consider the two extreme cases of a fully benevolent ($\mu \to \infty$) and a fully selfish ($\mu = 0$) default setter, for later reference. In the former case, preferences are completely aligned and the default setter incorporates all her private information in the default in any Pareto-efficient equilibrium. In the latter case, in contrast, the default setter conceals her information completely and the default never conveys any information.

The quality of information

Differences in agents' knowledge about the decision environment are a potentially important reason why the strength of defaults effects may differ across subgroups of the population. In our model, such differences in agents' information are captured by the agents' signal strength, x. Intuitively, agents who are less familiar with the choice environment may rely more heavily on the information conveyed through the default option. In contrast, decision makers with particularly accurate private information may rely less on defaults. The following proposition shows that this intuition is indeed correct.

Proposition 2. *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 prone 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 $s_A^x(\sigma_l, d_h)$, 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 signals shift "further" away from the high default and rely more strongly on their own information.

To wrap up, our discussion in this section has shown that the classic framework of strategic communication can successfully be adopted to capture the idea that defaults may constitute a mode of communication between default setters and decision makers. While this intuition has informally been discussed in the literature on default effects (e.g., Madrian and Shea 2001, Johnson and Goldstein 2003, McKenzie *et al.* 2006), our analysis illustrates that the strategic incentives of default setters and decision makers and the parties' relative level of information can influence the specification of default options, the overall strength of default effects, and the impact of defaults on different subgroups of the population. The comparative static predictions of our theoretical analysis are both intuitively plausible and consistent with a number of stylized facts from the empirical literature on defaults (see, e.g., McKenzie *et al.* 2006, Levav *et al.* 2010, Brown *et al.* 2011, Bronchetti *et al.* 2013). In the following sections, we will provide direct causal evidence on the role of incentives and information for default effects, based on a controlled laboratory experiment.

3 Design of the experiment

3.1 The game

In the experiment, we examine the key comparative statics of our theoretical framework regarding the impact of strategic incentives and private information on default effects. An important advantage of the laboratory environment is that we can directly vary both dimensions of interest, while holding all other aspects of the choice environment and the population of decision makers constant. Specifically, we consider two distinct treatment dimensions. First, we exogenously vary the alignment of interests between default setters and decision makers. We consider three different situations in which 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). Second, 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.

As a workhorse for implementing these treatment conditions, we use a simple paradigm in which one default setter interacts with one decision maker. 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 default setter and the agent receive independent signals about the composition of the current set of cards.

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

Agents receive information about the number of red and black cards in the current set, by a subset of cards that is privately revealed to them. 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.[8] 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 1. Since the default setter's signal strength is always 0.8, the agent is informed worse than the default setter in about 56% of cases (Columns 1 and 2 in Table 1); in about 25% of cases he has about the same signal strength (Column 3 in Table 1), and in 19% of cases (Columns 4 and 5), the agent is better informed than the default setter. We can therefore test whether agents with different levels of information about the decision systematically differ in their behavior, holding default setters' information quality and benevolence constant (Proposition 2).

Agent Type	A50	A69	A77	A94	A100
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 1 red	3 black 2 red	2 black 0 red	4 black 1 red	5 black 0 red

Table 1: *Distribution of agents' signals.*

Participants in the experiment do not see Table 1, but the signal-generating mechanism for default setters and agents is common knowledge. In particular, the default setter knows

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

the procedure how the agent is informed, and consequently the distribution of signal qualities. However, she is not informed about the number or colors of the cards that are revealed to the agents in a specific period. Hence, in line with our theoretical framework, the default setter essentially plays against a *population of agents* as depicted in Table 1, although default setters and agents are matched one-to-one in a given period.

After having received her signal, the default setter selects the default option for the current period. She can either specify "more red cards" or "more black cards" as the default. Default setters in the experiment can thus "communicate" with the agents through the default option, and reveal or conceal their private information by specifying the default accordingly. 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 confirm 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 default setter and the agent. If the agent presses "opt out", a new screen pops up on which the agent can (actively) choose between "more red cards" and "more black cards" as her ultimate decision for this period.[9] 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, default setter 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.

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 default setters differed across treatment conditions. In treatment FUL, default setters' payoffs are perfectly aligned with those of the agents. Thus, a default

[9] We neither impose a time limit for default setters 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.

setter 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, default setters 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 induces a default setter who wants to direct the agent towards one particular alternative, without taking the welfare consequences for the agent into account, i.e., a default setter with fully selfish interests. Another theoretically equivalent way to implement misaligned preferences is to assume that default setters receive points if and only if agents do not receive any points.[10] We opted for the first version to keep the setting as understandable for the subjects as possible. To induce partial benevolence, each matching group in our third treatment, PAR, consisted of 50% benevolent and 50% selfish default setters with payoff functions π_P^{FUL} and π_P^{MIS}, respectively. Agents are not informed about which type of default setter they are matched with in a given period. From agents' (ex ante) perspective, this treatment is thus equivalent to interacting with a partially benevolent default setter.

3.3 Parameters and procedures

We conducted a total 12 sessions of the experiment, four each for the FUL, PAR, and MIS treatment. In each of the sessions, we had 12 default setters and 12 agents interacting over 50 periods. Subjects within a session were divided into two matching groups with 12 participants each. Default setters 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.

[10] In both cases, the unique perfect Bayesian equilibrium is the babbling equilibrium in which defaults convey no information.

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 public 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.[11] The experiment started only after all participants had answered several control questions correctly.

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 1. *Default setters' propensity to truthfully reveal their signal through defaults is highest in FUL, intermediate in PAR, and lowest in MIS.*

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

Hypothesis 2. *Agents' aggregate propensity to accept defaults should be highest in FUL, intermediate in PAR, and lowest in MIS.*

This aggregate hypothesis is derived from a more specific sub-hypothesis. Proposition 1 implies that the strength of default effects is weakly increasing in the benevolence of the default setter. Applying Proposition 1 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" (types A50, A69, and A77 in Table 1).[12] In contrast, the model predicts

[11] A translation of the instructions for the PAR treatment can be found in Section A.3 in the appendix.

[12] Since some of the predictions are equal for individual types of agents who have lower signal quality than the default setter, we jointly denote types A50, A69, and A77 from Table 1 as "low-information types" for ease of exposition. Types A94 and A100 are denoted as "high-information types", accordingly.

no treatment difference in default adherence after conflicting signals for "high-information types" (type A94 and A100 in Table 1). Similarly, the application of Proposition 2 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 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 4. *For low-information types, the presence of default options increase the quality of decisions in FUL and PAR and has no influence in MIS. For high-information types, the quality of decisions is not affected by the presence of default options.*

4 Results

In this section, we present the results of the experiment. We first summarize the behavior of default setters 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).

4.1 How do default setters specify defaults?

Figure 1 summarizes default setters' behavior in the different treatments. The figure depicts the average frequency of defaults that correspond to the default setters' private signal about the state of the world. In line with Hypothesis 1, we observe strong treatment differences in the likelihood that defaults truthfully reveal default setters' private signal. Fully benevolent default setters (FUL) almost always reveal their signal (in 98% of cases). In contrast, default setters reveal their private signal only in 75% of cases in PAR[13] values of defaults setters.

[13]For the PAR treatment, we display the average level of information revelation over both types of default setters (i.e., those with misaligned and aligned preferences), since this is the crucial metric from the agents'

Selfish default setters 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 default setters' knowledge of the state of the world. The overall difference in the information quality of

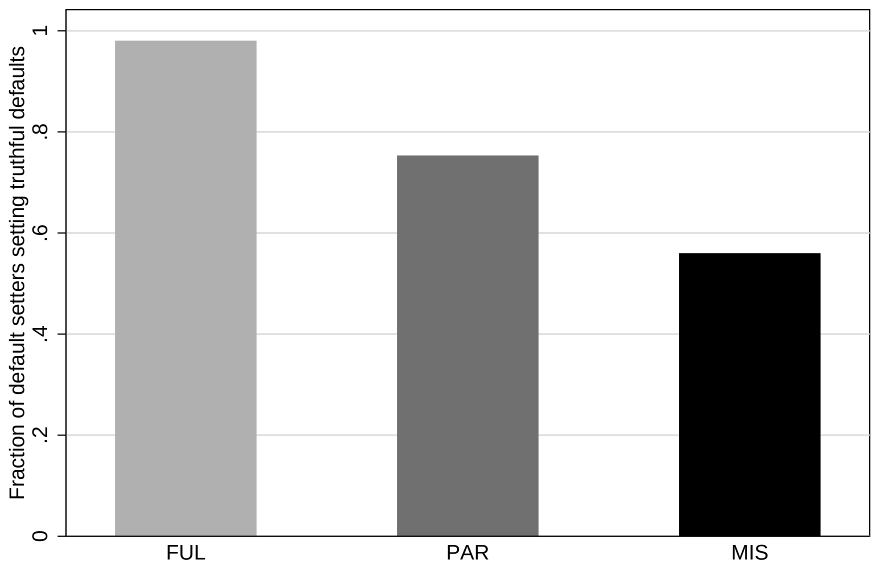


Figure 1: *Frequency of default setters setting default according to their signal. Average values per period in FUL, PAR, and MIS treatment.*

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).[14] Figure A.2 in the appendix shows that default setters' behavior 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, default setters' behavior in the experiment is not only qualitatively in line with our theoretical model, but also matches the quantitative predictions relatively accurately. A noteworthy exception is observed in the MIS treatment. The model predicts that defaults set by fully selfish default setters convey no information in equilibrium. That is, the likelihood of a default setters' signal being red or black should be 50%, independently of the observed default. Default setters' behavior in the experiment comes close to this prediction. However, the actual frequency of informative defaults in MIS is 56%. A similar pattern of over-

perspective.

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

communication for the case of misaligned interests has also been observed in a number of previous experiments on strategic communication (e.g., Gneezy 2005, Cai and Wang 2006, Sanchez-Pages and Vorsatz 2007). There are two broad sets of motives that could account for this deviation from equilibrium predictions. First, the higher informativeness of defaults could be explained by boundedly rational behavior of players (e.g., *level-k* reasoning as in Crawford 2003 or Wang *et al.* 2010). Second, some default setters in MIS might not be fully selfish in the sense that they deliberately provide agents with information about their private signal. This could, for instance, be due to some intrinsic preferences for honesty (Vanberg 2008, Erat and Gneezy 2012, Fischbacher and Heusi 2013), aversion towards payoff inequalities (Fehr and Schmidt 1999), or other forms of social preferences. Put differently, the true preferences of some default setters might not coincide with the monetary incentives induced in the MIS treatment.

Our data allow us to shed some light on the relevance of both channels in our setting. To do so, we focus on default setters in the MIS treatment and relate their default-setting behavior to different individual-level characteristics of the players. As a measure for the truthfulness of defaults, we calculate the relative frequency with which an individual default setter specifies a black default after receiving a black signal.[15] We then estimate a regression model with the default setters' individual truthfulness frequencies as dependent variable and different individual-level characteristics obtained from the post-experimental questionnaire as regressors. As an indicator for the "social preference" channel, 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).[16] As proxies for the "bounded rationality" channel, we include default setters' scores in the Cognitive Reflection Test (Frederick 2005), as well as their final math grade in high school.

Column 1 of Table 2 shows that default setters who score higher on the Honesty-Humility-Scale have a significantly higher likelihood to truthfully reveal a black signal in the MIS treatment. This suggests that heterogeneity in default setters' instrinsic preferences for honesty can at least partially account for the overcommunication that we observe. Column 2

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

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

Dependent variable: Frequency with which default setter 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	.050	.134

Table 2: *Determinants of default setters' behavior in MIS. OLS estimations; reported standard errors (in parentheses) account for potential clustering on the matching-group level. The honesty-humility-scale, HHS, measures an individual's inclination to avoid manipulation of others for personal gain (Ashton and Lee 2009). It ranges from -60 to +40, which higher scores indicating a higher inclination to avoid manipulations. The cognitive-reflection test, CRT, is a three-item questionnaire that measures an individual's inclination to suppress a spontaneous, but wrong answer (Frederick 2005). "Math grade" measures participants' final high-school math grade. Grades range from 1-6 with 1 being the best grade.*

shows that our proxies of cognitive ability also point in the expected direction. Specifically, worse (i.e., higher) math grades as well as lower CRT scores are both associated with a higher propensity to communicate truthfully under misaligned interests. Both effects, however, turn out to be insignificant. When including both sets of measures (Column 3), the indicator for honesty stays (weakly) significant, whereas the measures for cognitive reflection and math abilities remain insignificant. Overall, our evidence thus suggests that the "too informative" defaults observed in MIS might primarily be driven by preferences for honest behavior among a subset of default setters.

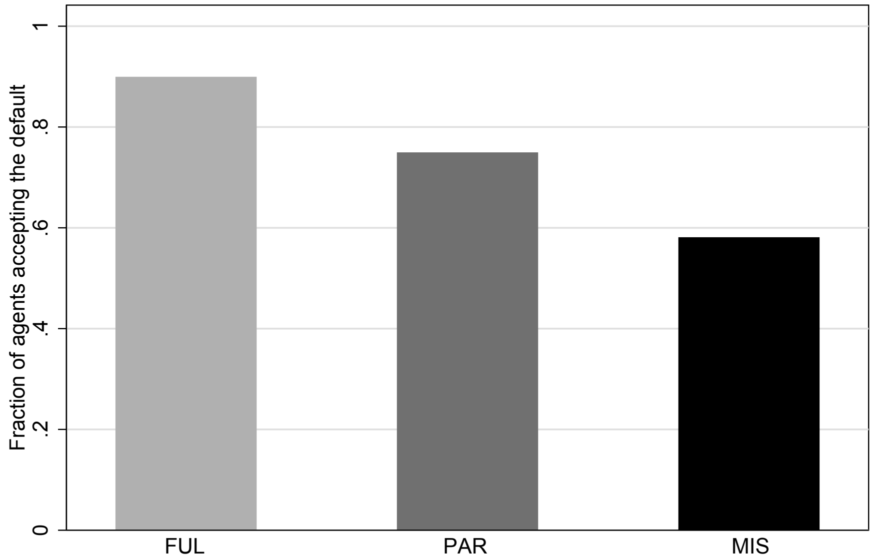


Figure 2: *Fraction of agents accepting the default. Average values in FUL, PAR, and MIS.*

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 depicts the aggregate frequencies of default adherence for agents in FUL, PAR, and MIS. Agents accept the default specified by default setters 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 the appendix 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.[17] The overall differences in agents' default acceptance rates are in line with our theoretical predictions, lending support to Hypothesis 2.

Thus, agents account for differences in the default setters' 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 3 depicts the average frequency of default acceptance

[17] 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).

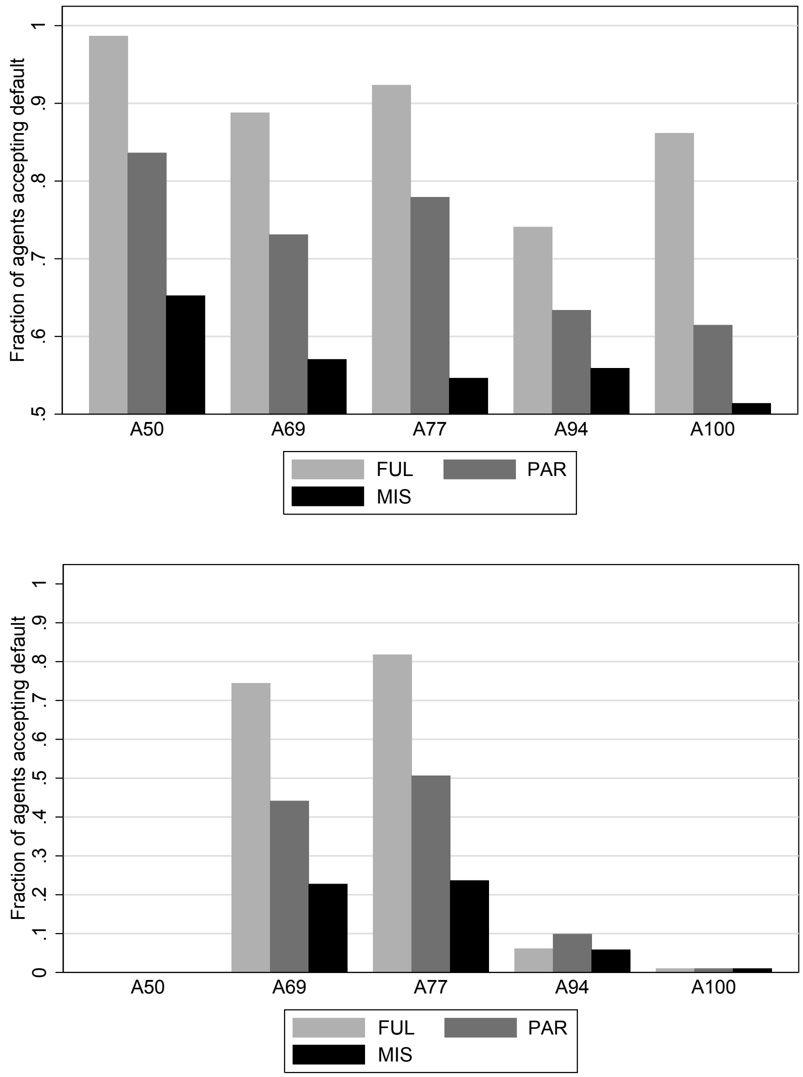


Figure 3: *Fraction of agents accepting the default. Upper panel: average aggregate values for different agent types in FUL, PAR, and MIS. Bottom panel: cases where default is in conflict with agents' private signal; average values for different agent types in FUL, PAR, and MIS.*

for the five different types of agents in the different treatments.[18] Substantiating our previous results, we find that the likelihood of accepting defaults increases in the benevolence of the default setter 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 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

[18]Agents are ordered according to their signal quality. For instance, type "A69" in Figure 3 are agents who received a signal with 69% precision in a given period (Column 2 in Table 1).

two red cards, but a red default).[19] The data depicted in the graph allow us to directly test the sub-hypothesis implied by Proposition 1. In particular, we predicted that low-information types are strictly more likely to accept defaults in FUL than in PAR 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 default setter. The figure indicates that both predictions are borne out by our data. We find that the default adherence rate for low-information types is strictly increasing in the benevolence of the default setter if the default is in conflict with their private signal (Fisher-exact test for low-information types. FUL vs. PAR: $p < .001$, FUL vs. MIS: $p < .001$, PAR vs MIS: $p = .010$). In contrast, 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 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 default setter who is at least partially benevolent. This within-treatment effect is borne 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 latter difference is driven by low-information types who follow conflicting defaults in the MIS treatment in about 20% of cases.

The top panel of Figure 3 shows that the diverging reactions of agents to conflicting defaults also induce overall differences in type-specific default acceptance rates. Agents who are better informed than the default setter (i.e., Type A94 and A100) are less likely to accept defaults than agents with inferior information quality if the default setter is (partially) benevolent. A Wilcoxon signed-rank test confirms that high-information types behave significantly different than low-information types in both the FUL and PAR treatment ($p = .012$, for both FUL and PAR). When being confronted with selfish default setters, 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

[19]Note that type A50 is excluded from the analysis. Since these types have no informative signal (1 red card and 1 black card), there is no conflict between the private signal and the default.

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 case of conflicting signals (bottom panel of Figure 3). Our model predicts that all types of agents completely ignore defaults set by a fully selfish default setter. 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.[20]

However, we have already seen that some default setters 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 the defaults that agents face during the experiment. If agents' choices are a best response to the behavior of default setters, the overall quality of agents' decision should be as least as high as in an environment in which there are no defaults and agents always follow their private signal. The next section explores in more detail whether this is the case.

4.3 Do defaults improve decisions?

In the final step of our empirical analysis, we turn to the question whether defaults improve the overall quality of agents' decisions in our experiment. We also study how the impact of defaults on decision quality depends on the agents' level of information and the default setters' level of benevolence. Figure 4 depicts the percentage change in the agents' decision quality compared to a hypothetical situation in which agents are assumed to always follow their private signals.[21] This scenario mimics a choice environment in which there are no default options and agents can base their decision only on their private signal. Positive values of the difference between actual and hypothetical decision quality in Figure 4 thus

[20] 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.

[21] To construct a measure of decision quality in this situation, we use the sets of cards and the actually realized private signals in the experiment for each agent.

indicate that agents benefited from defaults, while negative values mean that they would have attained higher payoffs in the absence of default options.[22]

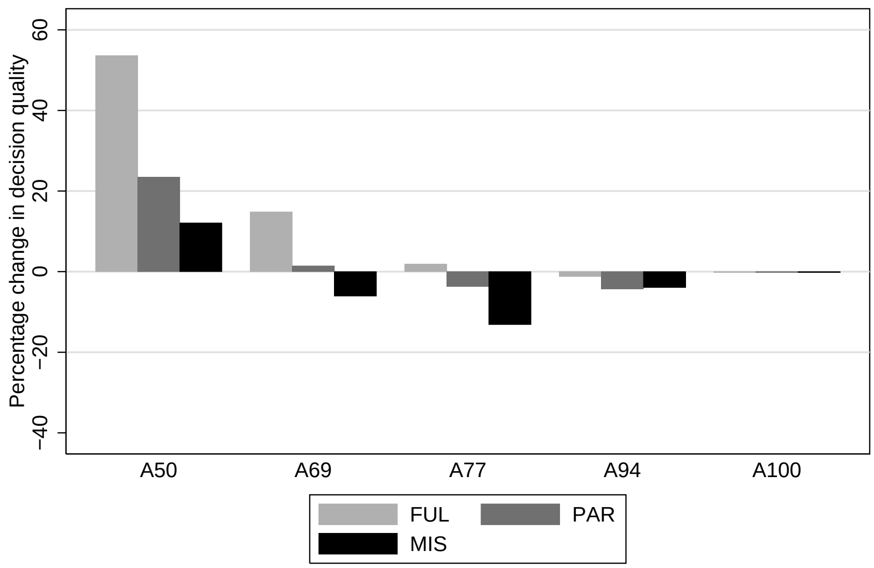


Figure 4: *Percentage change in agents' decision quality compared to a hypothetical situation in which agents always follow their own signal, for different agent types and treatments.*

Figure 4 indicates that the decision quality of high-information types (A94, A100) is hardly affected by the presence of defaults. This holds independently of the alignment of preferences between default setters and decision makers, and it reiterates the observation that these types react appropriately to their high signal quality: whenever the default option is in conflict with their private information, these agents tend to rely on the latter. 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 alignment of interests with the default setter. In the FUL treatment where default setter's and agents' information are fully aligned, we observe generally positive effects of default options, relative to the situation without default specifications. In particular, the agents with the lowest quality of private information (A50 and A69) make use of the informativeness of defaults which are set by the better informed default setter, and thereby attain higher payoffs.

For the PAR and MIS treatment, our findings on the effects of defaults for agents with low information quality are somewhat mixed. First, we observe unambiguously positive

[22]Note that this analysis provides a lower bound on the potentially beneficial effects of defaults in our setup, since it assumes that agents make no mistakes in the decision environment without defaults.

effects on the decision quality of agents with no information (A50). For type A69 who has intermediate information quality, defaults set by partially benevolent or selfish default setters have almost no impact on the decision quality, with slightly positive overall consequences in the PAR treatment, and slightly negative ones in the MIS treatment. Finally, the results for type A77 suggest that agents with information quality similar to the one of the default setter seem to misperceive the informational content of the default and rely to little on their private information. As a consequence, defaults in the MIS and PAR treatment are detrimental for agents' decision quality in this case.[23] Overall, this finding suggests that a sound understanding of the strategic incentives and the relative quality of one's own and the default setters' information are crucial for agents to reap the benefits of informative defaults, without bearing the detrimental consequences of following ill-specified ones.[24]

5 Conclusions

In this paper, we have analyzed two key channels through which non-binding default options can affect behavior—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 aspects have been informally discussed in the literature, we show that a well-established theoretical approach from the literature on strategic communication can be used to analyze these intuitions in a unifying formal framework. Adopting the notion that defaults can be a mode of strategic communication yields two main testable predictions. First, default options should have a stronger impact on behavior when the interests of the default setter and decision makers are more closely aligned. Second, whenever decision makers face a conflict between a default option and their own information about the decision, those with less information are more prone to accept the default. The empirical results from our laboratory experiment lend support to both of these predictions. In particular, the informational content

[23] Performing an analysis similar to the one in Table 2 further reveals systematic individual-level differences in agents' tendency to accept "conflicting" defaults that are set by default-setters with misaligned interests. In particular, agents with lower cognitive-reflection scores are significantly more likely to follow defaults that are at odds with their private signals in the MIS treatment. This suggests that the negative payoff consequences of default options may primarily be borne by decision makers with relatively low levels of cognitive or strategic reasoning (see also Crawford 2003 and Wang *et al.* 2010).

[24] Attaining such understanding might, admittedly, be challenging: findings by Cain *et al.* (2005) indicate that agents sometimes trust the information given by selfish advisers too strongly even if they are informed about potential conflicts of interest.

of defaults and the overall strength of default effects are increasing in the alignment of interests. We also observe that agents with lower information quality are indeed more susceptible to accept default options, if these are in conflict with the agent's individual information.

On a more general level, our analysis suggests that a comprehensive understanding of defaults effects needs to account both for the individual-level factors that may strengthen the impact of defaults (e.g., decision makers' status quo biases or present-biased preferences), as well as for the strategic aspects that are inherent in the interaction between default setters and decision makers. Our framework also lends itself to a number of interesting extensions. These include integrating the possibility that agents can spend cognitive resources for acquiring additional information, the possibility of heterogeneous preferences of agents for a given state of the world, or allowing for imperfect information about the opponent's interests.

Finally, our findings potentially provide some interesting insights for the discussion on "libertarian paternalistic" policy interventions. In particular, most participants in our experiment do take the default setter's strategic incentives and information into account, and condition their acceptance of default options on both factors. However, our empirical results also show that some agents misperceive the informational content of defaults and follow defaults from default setters with misaligned preferences too frequently. These findings suggest that a high level understanding of default setters' information and strategic incentives are crucial to ensure that consumers react appropriately to default options. Consumer protection policies such as mandatory disclosure laws might be a means to help decision makers attain such understanding.

References

Altmann, S., Falk, A., Heidhues, P. and Jayaraman, R. (2014). Defaults and donations: Evidence from a field experiment. *IZA Discussion Paper No. 8680*.

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

Battaglini, M. and Makarov, U. (2014). Cheap talk with multiple audiences: An experimental analysis. *Games and Economic Behavior*, **83**, 147 –164.

Beshears, J., Choi, J. J., Laibson, D. and Madrian, B. C. (2008). The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States.

In S. J. Kay and T. Sinha (eds.), *Lessons from Pension Reform in the Americas*, Oxford: Oxford University Press, pp. 59–87.

—, —, — and — (2010). The limitations of defaults. *Working Paper, Harvard University.*

Borghans, L. and Golsteyn, B. H. H. (2014). Default options and training participation. *Empirical Economics*, **46** (4), 1417–1428.

Bronchetti, E. T., Dee, T. S., Huffman, D. B. and Magenheim, E. (2013). When a nudge isn't enough: Defaults and saving among low-income tax filers. *National Tax Journal*, **66** (3), 609–634.

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

Brown, J. R., Farrell, A. M. and Weisbenner, S. J. (2011). The downside of defaults. *Working Paper, University of Illinois at Urbana-Champaign.*

Cai, H. and Wang, J. T.-Y. (2006). Overcommunication in strategic information transmission games. *Games and Economic Behavior*, **56**, 7–36.

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

Caplin, A. and Martin, D. (2013). Defaults and attention: The drop out effect. *Working Paper, NYU.*

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

Chen, Y. (2009). Communication with two-sided asymmetric information. *Working Paper, Arizona State University.*

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H. and Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from denmark. *The Quarterly Journal of Economics*, **129** (3), 1141–1219.

Crawford, V. (1998). A survey of experiments on communication via cheap talk. *Journal of Economic Theory*, **78**, 286–298.

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

Crawford, V. P. (2003). Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *The American Economic Review*, **93** (1), 133–149.

—, Costa-Gomes, M. A. and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, **51** (1), 5–62.

Dickhaut, J. W., McCabe, K. A. and Mukherji, A. (1995). An experimental study of strategic information transmission. *Economic Theory*, **6**, 389–403.

Drugov, M., Hernan-Gonzalez, R., Kujal, P. and Martinez, M. T. (2013). Cheap talk with two audiences: An experiment. *working paper*.

Ebeling, F. (2013). Non-binding defaults and voluntary contributions to a public good—clean evidence from a natural field experiment. *Working Paper, University of Cologne*.

Erat, S. and Gneezy, U. (2012). White lies. *Management Science*, **58** (4), 723–733.

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

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

— and Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, **11** (3), 525–547.

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

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, **95** (1), 384–394.

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*, Göttingen: Ges. für Wiss. Datenverarbeitung, pp. 79–93.

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

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

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

Lai, E. K. (2014). Expert advice for amateurs. *Journal of Economic Behavior and Organization*, **103**, 1–16.

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

Löfgren, A., Martinsson, P., Hennlock, M. and Sterner, T. (2012). Are experienced people affected by a pre-set default option—results from a field experiment. *Journal of Environmental Economics and Management*, **63**, 66–72.

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

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

Moreno de Barreda, I. (2013). Cheap talk with two-sided private information. *Working Paper, University of Oxford*.

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

Sanchez-Pages, S. and Vorsatz, M. (2007). An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior*, **61**, 86–112.

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

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

— (2013). Impersonal default rules vs. active choices vs. personalized default rules: A triptych. *Working Paper, Harvard University*.

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

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

Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica*, **76** (6), 1467–1480.

Wang, J. T., Spezio, M. and Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *The American Economic Review*, **100** (3), 984–1007.

Appendix

A.1 Proofs of Lemmas and Propositions

Construction of an output equivalent equilibrium

Suppose strategy profile s with strategies s_A^x, s_P constitutes a perfect 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 $d_j \in D_2$. The strategy of the default setter can be written as

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

Define a new strategy of the default setter:

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

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 constitutes a PBE. 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, $\hat{s_P}$ must also be a best response, since messages d_i, d_j induce the same action and the default setter 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 that does not contain any out-of-equilibrium default.

Proof of Lemma 1

If the considered equilibrium is uninformative all defaults inhibit the same information. Next, consider informative equilibria. Hence, there are at least two different defaults with $p(\theta_h|d_k) > p(\theta_h|d_f)$. Given the signal ρ_h, we show that it is always more profitable to play d_k for the default setter. The incentive function of the default setter, i.e., the expected payoff difference from d_k versus d_f is

$$\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(s_A^x(\sigma,d_k))|\rho_h,x]f(x)dx - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(s_A^x(\sigma,d_f))|\rho_h,x]f(x)dx$$
$$= \int_{\frac{1}{2}}^{1} p(\sigma_l,\theta_l|\rho_h,x)(\mu(U_A(\theta_l,s_A^x(\sigma_l,d_k)) - U_A(\theta_l,s_A^x(\sigma_l,d_f))) + b(s_A^x(\sigma_l,d_k)) - b(s_A^x(\sigma_l,d_f)))$$
$$+ p(\sigma_h,\theta_l|\rho_h,x)(\mu(U_A(\theta_l,s_A^x(\sigma_h,d_k)) - U_A(\theta_l,s_A^x(\sigma_h,d_f))) + b(s_A^x(\sigma_h,d_k)) - b(s_A^x(\sigma_h,d_f)))$$
$$+ p(\sigma_l,\theta_h|\rho_h,x)(\mu(U_A(\theta_h,s_A^x(\sigma_l,d_k)) - U_A(\theta_h,s_A^x(\sigma_l,d_f))) + b(s_A^x(\sigma_l,d_k)) - b(s_A^x(\sigma_l,d_f)))$$
$$+ p(\sigma_h,\theta_h|\rho_h,x)(\mu(U_A(\theta_h,s_A^x(\sigma_h,d_k)) - U_A(\theta_h,s_A^x(\sigma_h,d_f))) + b(s_A^x(\sigma_h,d_k)) - b(s_A^x(\sigma_h,d_f)))f(x)dx.$$

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 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 default setter conveys her information to the agents. Hence, $p(\theta_h|\sigma,\rho_h,x)$ is weakly larger than $p(\theta_h|\sigma,d_k,x)$, which represents the assessment of agent x after he observes default d_k. As the default setter puts more probability weight on situations where a higher z is profitable for the agent, her expectation about the advantage of $s_A^x(\sigma,d_k)$ over $s_A^x(\sigma,d_f)$ is weakly higher than the agent's own expectation, which is given by

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

This in turn is strictly larger than zero as long as $s_A^x(\sigma,d_k) \neq s_A^x(\sigma,d_f)$ 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 default setter receives a high signal. Hence, every default played in response to a high signal exhibits the same information. All other defaults reveal that the default setter received a low signal.

Proof of Proposition 1

The proof consists of two steps. First, we show that there is at most one positive transmission rate c^{pd} that is associated with an informative Pareto-efficient equilibrium. Second, we show that this transmission rate increases in μ.

For the first step suppose there are two informative mixed equilibria with $0 < c_1 < c_2 < 1$. Agents always prefer an equilibrium corresponding to c_2 since more information yields a better decision and, thus, a higher expected payoff. The same is true for the default setter if she receives a high signal because preferences are aligned in this case. If the default setter receives a low signal, she is indifferent between both equilibria. To see this, suppose the default setter chooses a message from the set D_l. Since this action reveals her signal (see Lemma 1), the inherent information and the expected payoff associated with a low default are equal in both equilibria. Furthermore, the default setter is also indifferent between a low and a high default. Hence, the expected payoff after a high default must also be identical in both equilibria. Overall, if there exist two mixed equilibria, the default setter prefers the one with the higher information transmission rate. Moreover, she prefers an equilibrium with $c = 1$ to any equilibrium with partial information revelation if the former exists. If she receives a low signal the utility in a full revelation equilibrium is identical to the utility in any mixed equilibrium. After a high signal, however, she benefits from the revelation of her signal. Hence, only the highest information rate that is implementable in equilibrium can be associated with a Pareto-efficient equilibrium.

In the second step, we show that c^{pd} is weakly increasing in μ. For this purpose first assume that there exists an informative Pareto-efficient equilibrium with $c^{pd} < 1$. Due to Lemma 1, $\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, d_l))|\rho_h, x]f(x)dx$ does not vary in the transmission rate c^{pd}. In the assumed equilibrium, the default setter is indifferent between a high and a low default given she received a low signal

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

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

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

$$= -\frac{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, d_h))|\rho_l, x]f(x)dx - \int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, d_l))|\rho_l, x]f(x)dx\right)/\partial\mu}{\partial\left(\int_{\frac{1}{2}}^{1} E_{\sigma,\theta}[U_P(\theta, s_A^x(\sigma, d_h))|\rho_l, x]f(x)dx\right)/\partial c^{pd}}.$$

The numerator of this fraction is smaller than zero, if

$$\begin{aligned}\int_{\frac{1}{2}}^{1} & p(\sigma_l, \theta_l|\rho_l, x)(U_A(\theta_l, s_A^x(\sigma_l, d_h)) - U_A(\theta_l, s_A^x(\sigma_l, d_l)))) \\ & + p(\sigma_h, \theta_l|\rho_l, x)(U_A(\theta_l, s_A^x(\sigma_h, d_h)) - U_A(\theta_l, s_A^x(\sigma_h, d_l))) \\ & + p(\sigma_l, \theta_h|\rho_l, x)(U_A(\theta_h, s_A^x(\sigma_l, d_h)) - U_A(\theta_h, s_A^x(\sigma_l, d_l))) \\ & + p(\sigma_h, \theta_h|\rho_l, x)(U_A(\theta_h, s_A^x(\sigma_h, d_h)) - U_A(\theta_h, s_A^x(\sigma_h, d_l)))f(x)dx \leq 0.\end{aligned}$$

This follows if the inequality holds for all x. Since defaults convey less information than the signal of the default setter and agents maximize their utility the same argument as in Lemma 1 applies. We conclude that the numerator is strictly smaller than zero. Thus,

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

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 the postulated equilibrium with transmission rate c^{pd} is Pareto-efficient. Hence, the information transmission rate is increasing for informative Pareto-efficient equilibria with partial information transmission. Therefore, it is also weakly increasing for any informative Pareto-efficient equilibrium. Since agents are Bayesian updaters, this leads to weakly stronger responses, 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

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 updaters, they weigh their private signal stronger if it conveys more information. For any given equilibrium, agents are thus more prone to follow the default if they have less private information.

A.2 Supplementary Figures

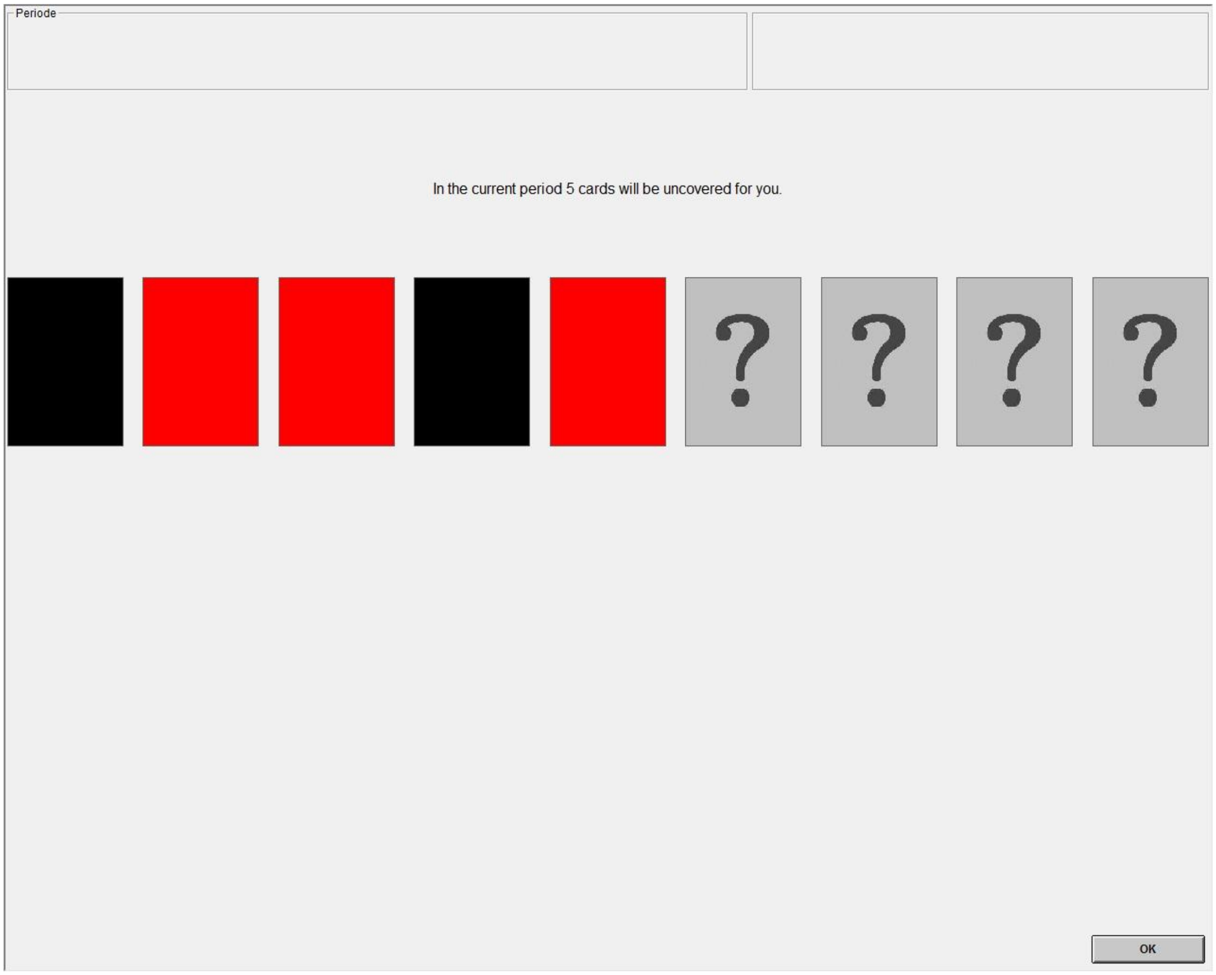


Figure A.1: *Agent's information screen. Example in which two black and three red cards are revealed.*

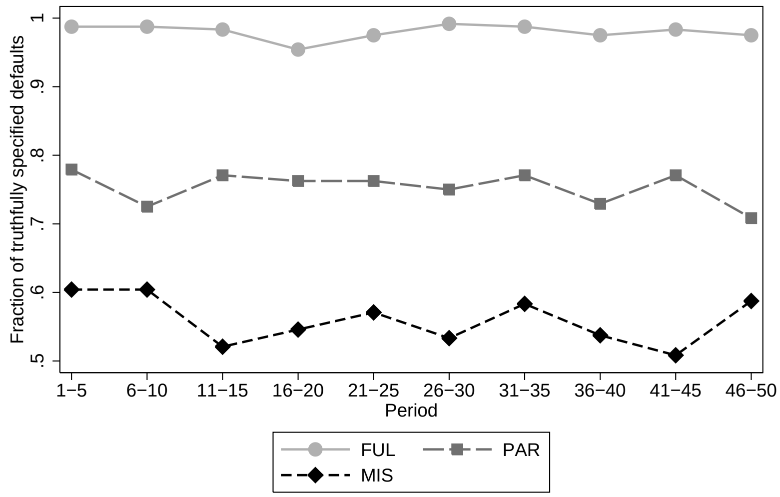


Figure A.2: *Frequency with which default setters set default according to their signal. Average values for 5-period intervals in FUL, PAR, and MIS.*

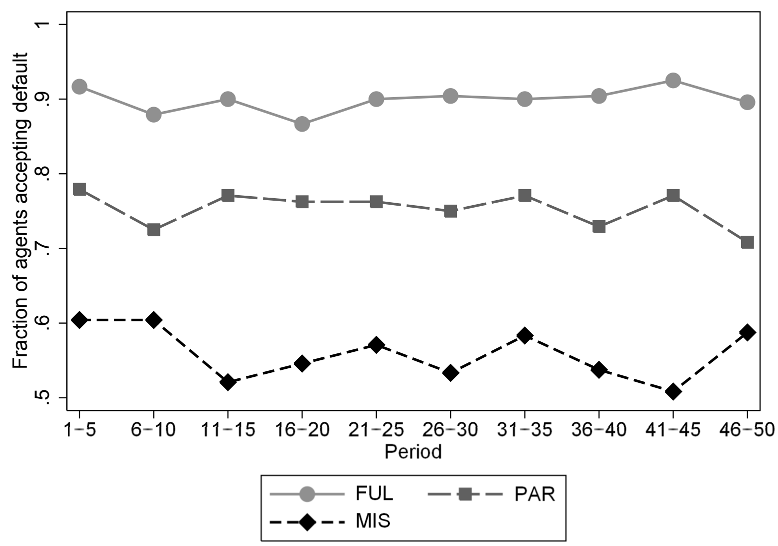


Figure A.3: *Frequency with which agents follow default option. Average values for 5-period intervals in FUL, PAR, and MIS.*

A.3 Instructions of the Experiment

In what follows, we present a translation of the instructions for the PAR treatment. The instructions for the other treatments had a similar structure and differed only in the description of player A's incentives.

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 received documents 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 to you on your screen. Your role remains the same during the entire experiment.

In each period of the experiment groups will be formed either consisting of one player 1A and one player 2 **or** one player 1B and one player 2. 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 chooses 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.

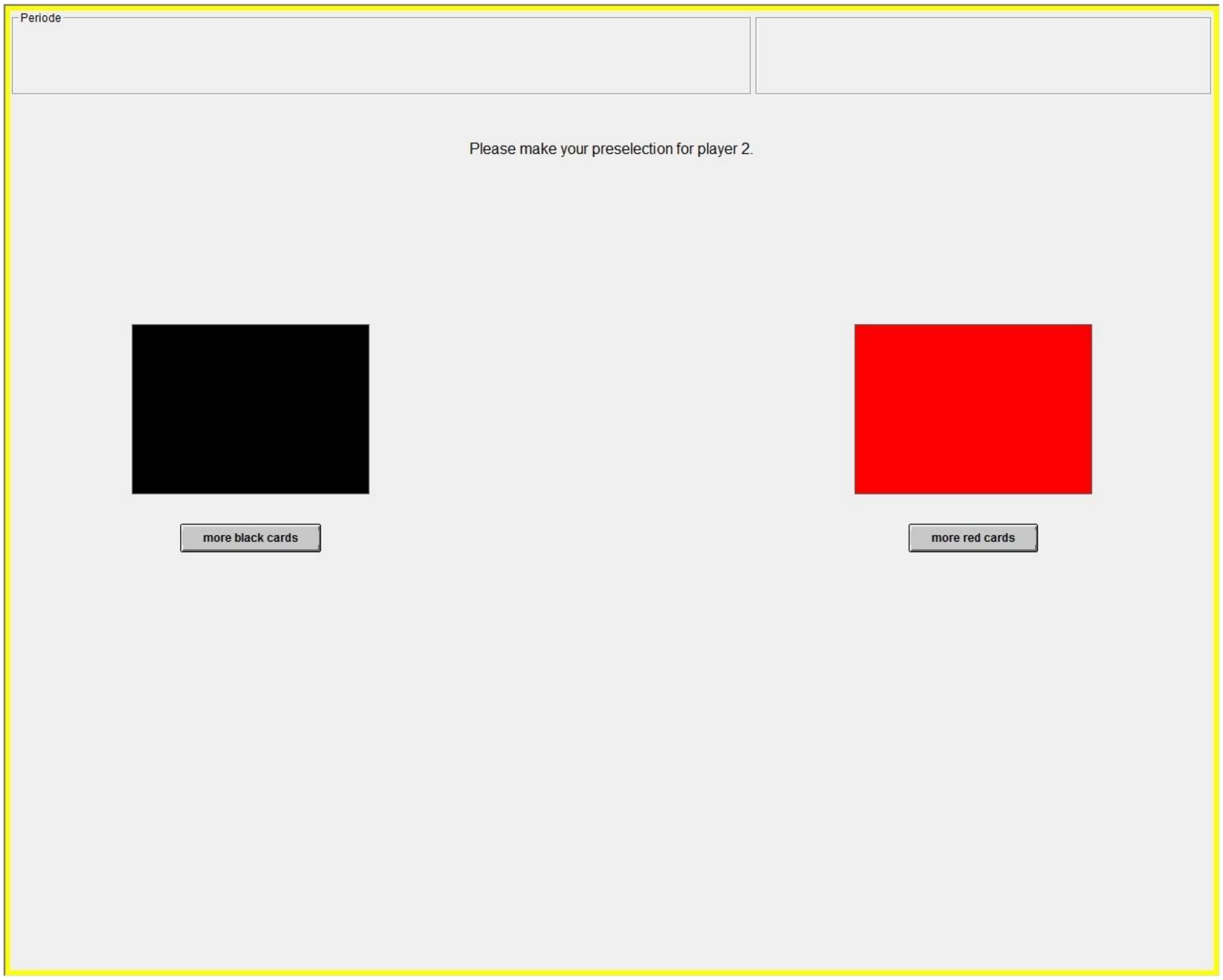
Player 1 (that is, player 1A or 1B) receives his information about the amount of red and black cards in the current set of cards via a signal on his screen. There are two possible **signals which can be displayed on the screen**: either the signal for player 1 is "more red cards" or the signal is "more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

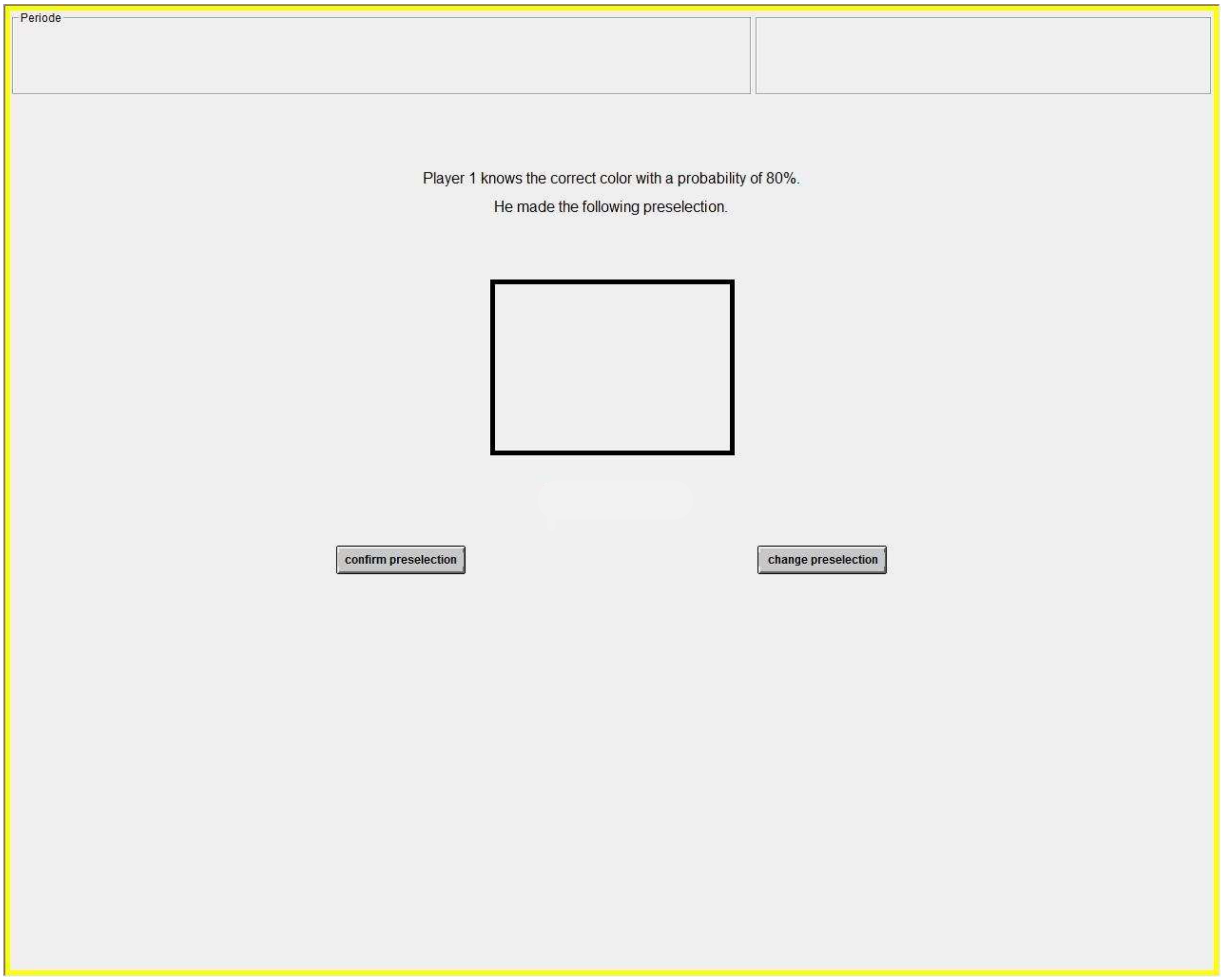
Player 2 receives his information about the number of red and black cards **by the uncovering of a subset of the set of 9 cards.** In each period either **two** or **five** cards will be uncovered for player 2. The probability that two cards will be uncovered in a given period is 50%. The probability that five cards will be uncovered in a given 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 given 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". The following screen is displayed to her, on which she can make her decision. She chooses the preselection by clicking the respective button.



Afterwards, player 2 is informed about the preselection which the other group member chose for him. The following screen is shown to player 2, on which the preselection of player 1 is displayed (instead of gray rectangular 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 given period are calculated as follows.

Player 1A:

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

Player 1B:

- If 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**.
- If 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:

- If 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**.
- If 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.