DISCUSSION PAPER SERIES

DP16400

Incentives, Self-Selection, and Coordination of Motivated Agents for the Production of Social Goods

Kevin Bauer, Michael Kosfeld and Ferdinand Von Siemens

ORGANIZATIONAL ECONOMICS

CEPR

Incentives, Self-Selection, and Coordination of Motivated Agents for the Production of Social Goods

Kevin Bauer, Michael Kosfeld and Ferdinand Von Siemens

Incentives, Self-Selection, and Coordination of Motivated Agents for the Production of Social Goods

Abstract

We study, theoretically and empirically, the effects of incentives on the self-selection and coordination of motivated agents to produce a social good. Agents join teams where they allocate effort to either generate individual monetary rewards (selfish effort) or contribute to the production of a social good with positive effort complementarities (social effort). Agents differ in their motivation to exert social effort. Our model predicts that lowering incentives for selfish effort in one team increases social good production by selectively attracting and coordinating motivated agents. We test this prediction in a lab experiment allowing us to cleanly separate the selection effect from other effects of low incentives. Results show that social good production more than doubles in the low- incentive team, but only if self-selection is possible. Our analysis highlights the important role of incentives in the matching of motivated agents engaged in social good production.

Kevin Bauer - bauer@safe-frankfurt.de
Leibniz Institute for Financial Research SAFE

Michael Kosfeld - kosfeld@econ.uni-frankfurt.de
Goethe University Frankfurt and CEPR

Ferdinand Von Siemens - vonsiemens@econ.uni-frankfurt.de
Goethe University Frankfurt

Incentives, Self-Selection, and Coordination of Motivated Agents for the Production of Social Goods

Kevin Bauer, Michael Kosfeld, and Ferdinand A. von Siemens[*]

July 20, 2021

Abstract

We study, theoretically and empirically, the effects of incentives on the self-selection and coordination of motivated agents to produce a social good. Agents join teams where they allocate effort to either generate individual monetary rewards (selfish effort) or contribute to the production of a social good with positive effort complementarities (social effort). Agents differ in their motivation to exert social effort. Our model predicts that lowering incentives for selfish effort in one team increases social good production by selectively attracting and coordinating motivated agents. We test this prediction in a lab experiment allowing us to cleanly separate the selection effect from other effects of low incentives. Results show that social good production more than doubles in the low-incentive team, but only if self-selection is possible. Our analysis highlights the important role of incentives in the matching of motivated agents engaged in social good production.

JEL classification: C91, D90, J24, J31, M52

Keywords: incentives, intrinsic motivation, self-selection, public service

1 Introduction

The promotion of social and economic welfare often requires the coordinated activity of motivated individuals. If workers can count on the high motivation and effort of their co-workers and colleagues, overall performance and productivity typically benefit. Itoh (1991) and Kremer (1993) offer classic economic frameworks for the analysis of such production complementarities. Ichniowski et al. (1997), Hamilton et al. (2003), Bandiera et al. (2007, 2009), Iranzo et al. (2008), and Friebel et al. (2017) provide empirical evidence across various organizational contexts. But how do motivated individuals find each other? What role do monetary incentives play in the selection and coordination of individuals with similar motivations to promote particular goals?

In this paper, we develop and test empirically a screening model of team production in a world consisting of motivated and unmotivated agents. Similar to the mission motivation literature triggered by the seminal work of Besley and Ghatak (2005), motivated agents care about team output (what we call a *social good*), while unmotivated agents do not. We extend the existing analysis by assuming that the production of the social good is characterized by complementarities, capturing the need for coordination highlighted above: the higher the effort of other team members, the higher the marginal productivity of each agent, and hence the overall production of the social good. While agents can contribute to the production of the social good, they can also exert effort to generate individual monetary rewards (what we call *selfish effort*). Both effort dimensions are assumed to be substitutes in an agent's cost function capturing a multi-tasking environment (Holmstrom and Milgrom, 1991). Our theoretical results show that monetary incentives for selfish effort play a key role in this context by sorting and coordinating like-minded agents across teams. In particular, motivated agents are attracted by low monetary incentives for selfish effort, for the very reason that unmotivated agents find these incentives unattractive. The latter prefer high incentives (if available), so that motivated agents can be sure to be matched with other motivated agents when choosing a team with low incentives. This mechanism, similarly, enables them to coordinate on a high-effort equilibrium.

The results highlight an important matching mechanism in the recruitment of motivated agents for the production of social goods. A prime example for such goods is public service delivery. Any teacher, social worker, or health worker can only be as productive in the delivery

of her pro-social service as the other team members in her group and department. Recent field studies have analyzed the effects of various forms of incentives in this context, such as wages (Dal Bó et al., 2013; Barfort et al., 2019; Deserranno, 2019) or career perspectives (Ashraf et al., 2020). Results show that incentives entail significant selection effects; however, the direction with respect to pro-social or public service motivation depends on the particular setting and the type of incentive. While some of the findings (e.g., Barfort et al., 2019; Deserranno, 2019) are consistent with the predictions of our model, overall the available field data do not allow for a direct test of the main matching mechanism underlying these predictions.[1] We therefore implement a novel laboratory experiment to test the causal effect of monetary incentives on the production of social goods under self-selection.

A key challenge — both in the lab and in the field — is to separate any selection effect from the effect of incentives *on* the job. In the field this is done by trying to shut down the latter channel once workers have passed the selection stage, e.g., by giving workers in each treatment arm the same incentives after they have been recruited. A potential problem with this procedure is that it may introduce surprise effects, as some workers learn that their initial expectations were wrong. The advantage of the lab is that we can test for the selection effect directly by setting up a control treatment that keeps differences in incentives unchanged but shuts down selection itself. This is achieved by assigning workers randomly to teams, a procedure that is hard to implement and often infeasible in the field. Further, the experimenter can define the production function in the lab and thus has full control over the influence of workers' characteristics on team output. While screening and selection in the field are naturally multidimensional (e.g., with respect to both motivation and ability), the lab setting enables us to target and isolate the relevant heterogeneity we are interested in.

In our experiment, participants allocate effort to the production of a social good. The social good is produced in teams and comes in form of monetary donations to the *Deutsche Krebshilfe*, which is a well-known charitable organization in Germany supporting the fight against cancer. We choose this organization based on the assumption that participants differ in their individual motivation to generate monetary donations to this charity.[2] Effort that is

---

[1]For example, varying incentives in a single job at a particular public service organization might be insufficient to affect applicants' expectations about relevant characteristics of *other* workers in the same organization. Even with an entire recruitment of a new team, characteristics of the organization's incumbent workforce that is not replaced may still play a decisive role in the selection process.

[2]A pilot study prior to the experiment, in which (a different group of) participants could make donations

not allocated to the production of the social good, so-called selfish effort, generates monetary rewards that the participants reap themselves. In the main treatment, participants can choose between two teams that differ in the monetary incentives for selfish effort. Two control treatments vary these conditions. In the first control treatment, incentives differ but we assign participants exogenously and randomly to teams. With this control treatment we isolate the effect of selection. In the second control treatment, participants can choose between teams but incentives for selfish effort in the two teams are now the same, allowing us to identify the effect of different (i.e., low vs. high) monetary incentives. In particular, we isolate the coordination effect of different incentives.

Results from the experiment strongly confirm the theoretical predictions of our model. While production levels of the social good are low in both control treatments, at about 20 percent of maximum social effort on average, they more than double to about 50 percent of maximum social effort in our main treatment. Crucially, this increase happens only in the low-incentive team, while in the high-incentive team production stays at a similarly low level as in either team in the control treatments. Together the results provide clear causal evidence for the effect of low monetary incentives and self-selection on the production of social goods. Additional analyses confirm that the effect is driven by both the separation and coordination of motivated individuals, corroborating the main mechanism of our model. In sum, the empirical analyses elucidate the key role high and low incentives play in the decision making of motivated and unmotivated agents.

Our paper contributes to three different strands in the literature. First, the results relate to a recent set of empirical studies analyzing the selection into public service (Dohmen and Falk, 2010; Serra et al., 2011; Dal Bó et al., 2013; Hanna and Wang, 2017; Barfort et al., 2019; Deserranno, 2019; Friebel et al., 2019; Ashraf et al., 2020). Results in Banuri and Keefer (2016) and Deserranno (2019) are closest to our paper as they also document the selection of pro-socially motivated individuals via low wages. While, e.g., Deserranno (2019) interprets her findings in the light of a signaling model à la Bénabou and Tirole (2003), the results are also consistent with the mechanism studied in this paper highlighting the complementarity in the production of social goods.

Second, our paper adds to the theoretical literature on sorting and selection in the labor market, in particular with respect to intrinsic motivation and work mission (Francois, 2000;

to various charitable organizations confirmed this assumption. Details are available upon request.

Besley and Ghatak, 2005; Brekke and Nyborg, 2008; Delfgaauw and Dur, 2008; Prendergast, 2007; Buurman and Dur, 2012; Auriol and Brilon, 2014; Barigozzi and Burani, 2019). Our model developed in the next section builds on insights from Kosfeld and von Siemens (2009, 2011) analyzing the self-selection of workers with heterogeneous motivations for cooperation, as well as von Siemens and Kosfeld (2014) considering heterogeneity in worker ability. The main contribution of the present paper to this literature stream is its focus on production complementarities and the implication for the selection and coordination of motivated agents via monetary incentives.

Third, our results relate to the lab experimental literature on sorting in public goods experiments (see, e.g., Gächter and Thöni, 2005; Brekke et al., 2011; Aimone et al., 2013). The main difference to our set-up is in the production function of the public good that is considered in these experiments. First, all players in these experiments benefit materially from the public good, which gives rise to classic free-rider incentives. Second, there are no complementarities in production. The latter feature is particularly important for our analysis as it reflects the key incentive for motivated agents to self-select into low-incentive environments: working with like-minded others.

A final comment may be worth making. Our paper shall, of course, not be understood as an argument, let alone a plea for paying motivated workers poorly. Rather, the results identify, theoretically and empirically, a particular matching mechanism via which relatively low monetary incentives for selfish effort are only one possible way to attract motivated workers. Organizations and firms may have also other possibilities to realize similar effects, e.g., the provision of extra time to pursue social projects (Barfort et al., 2019), or training measures focusing on an organization's mission motivated workers care about. In principle, anything that is undesirable for unmotivated workers and keeps them away may work. It is in the interest and in the hands of employers to design incentive schemes and work environments such that they are attractive for motivated workers, taking their utility into account.

The paper is organized as follows. Section 2 presents our theoretical model. The main objective is to generate testable predictions for behavior in the experiment. Section 3 explains details of the experimental design. Section 4 discusses the results. Finally, Section 5 concludes.

2 Theory

2.1 Model

Consider a population of agents with total mass normalized to one. Agents can produce a *social good* — a pro-social output — in teams but differ in their intrinsic motivation to do so. Individual intrinsic motivations are private information. After learning their type (i.e., motivation), agents simultaneously choose between two available teams. They then simultaneously exert *selfish* and *social effort*. Selfish effort generates monetary earnings for an agent, while social effort contributes to the production of the social good. We assume complementarities in the latter production, i.e., an agent's effective contribution to the social good increases in other agents' average social effort in the same team. Teams can differ in the monetary incentives for selfish effort. Let us now fill in the details.

For simplicity, we assume that there are two types of agents $\theta \in \{m, u\}$. Unmotivated types (u) exclusively care about their monetary earnings. Motivated types (m) also experience positive utility from what they contribute to the social good production. Types are i.i.d. and private information. An agent's prior probability to be motivated $\mu_0 \in]0, 1[$ is common knowledge. We assume that there are two teams $k \in \{h, \ell\}$ that may differ in the incentive $b_h \geq b_\ell > 0$ with which selfish effort is monetarily rewarded. While the opportunity to earn monetary income may vary, the social good's production function is the same in both teams.[3]

After learning their individual type, agents decide to join one of the two teams. Let $p(k, \theta)$ denote the type-dependent probability with which an agent of type θ chooses team $k \in \{h, \ell\}$. Once in the team, each agent divides his effort endowment $E > 0$ between social and selfish effort. Let $e(k, \theta)$ denote the type-dependent social effort of an agent of type θ in team k and let $E - e(k, \theta)$ denote the level of selfish effort. Selfish effort generates individual monetary earnings

$$b_k \left(E - e(k, \theta) \right) \tag{1}$$

given incentive $b_k \geq 0$ in the chosen team. Social effort $e(k, \theta)$ adds

$$d = e(k, \theta) \, \bar{e}_k \tag{2}$$

---

[3]We consider incentives to be exogenous in this model. Kosfeld and von Siemens (2009) prove that different monetary incentives co-exist in equilibrium when firms offer contracts in a competitive labor market with free entry.

to the production of the social good, where $\bar{e}_k$ denotes the average social effort of other agents in team k. Equation (2) captures the complementarity in production of the social good: the more social effort the other members in the team exert, the higher is the contribution of an agent's social effort to the social good. Note that $\bar{e}_k$ is the same for all agents in team k, given the total mass of agents in the team is strictly positive. When an agent is alone in the team, there is no average effort of any others. For simplicity, we then set $\bar{e}_k$ equal to $e(k, \theta)$ so that the production of the social good is determined by the agent alone.

All agents maximize expected utility. Unmotivated types only care about their monetary reward, where

$$U(e(k, u), \bar{e}_k, u) := b_k \left(E - e(k, u) \right) \tag{3}$$

describes the utility of an unmotivated agent who joins team k and exerts social effort $e(k, u)$. For simplicity we assume linear utility. Other functional forms may be applicable but do not alter the main insights from our model.[4] Motivated agents in addition care about how much they add to the production of the social good. Let

$$U(e(k, m), \bar{e}_k, m) := b_k \left(E - e(k, m) \right) + \alpha \, e(k, m) \, \bar{e}_k \tag{4}$$

describe the utility of a motivated agent who joins team k and exerts social effort $e(k, m)$, while the other agents in the team exert average social effort $\bar{e}_k$. Parameter α captures the extra utility a motivated agent receives from the effective contribution to the production of the social good. We assume that

$$\alpha \, E > b_h, \tag{5}$$

which implies that motivated agents optimally exert maximum social effort even in the high-incentive team if all other agents in the team exert maximum social effort as well. This assumption makes our analysis relevant and interesting. Otherwise, it would always be optimal for motivated agents to join the high-incentive team and exert only selfish effort, thus leaving no scope for heterogeneity in intrinsic motivation to affect the production of the social good at all.[5] Agents simultaneously self-select into teams and allocate their effort before

---

[4]For example, a non-linear utility specification would yield interior production equilibria with unmotivated and motivated types exerting non-zero and less than maximum social effort, respectively (see Lemma 1 and 2 below). As long as there is the scope and desire for motivated types to exert more social effort than the unmotivated, our main result on separation via low monetary incentives does not change.

[5]To see this, note that the maximum utility a motivated agent can get from exerting social effort is αE^2.

learning about the other agents' selection and effort decisions. The game is static, and we use Bayesian Nash equilibrium as equilibrium concept. A fully specified strategy describes type-dependent probabilities $p(k, \theta)$ with which to choose any team $k \in \{h, \ell\}$ and type-dependent social effort choices $e(k, \theta)$ conditional on the chosen team.[6] Let $\mu_k \in [0, 1]$ denote the probability with which agents believe a member of team $k \in \{h, \ell\}$ to be motivated. We focus on symmetric equilibria in which agents use the same type-dependent strategy. We can thus suppress individual indexation, as already done above. The equilibrium strategy $p^*(k, \theta)$ and $e^*(k, \theta)$ maximizes expected utility for each type given the equilibrium behavior of other agents and the equilibrium beliefs μ_k^* concerning the distribution of types in the two teams. Beliefs must be consistent with equilibrium behavior. We also require type-dependent equilibrium effort choices to maximize expected utility in the team the agent might never join in equilibrium.[7]

2.2 Equilibria

Before considering selection into teams, we first illustrate the two polar production equilibria that can exist *within* teams. In the so-called *unproductive equilibrium*, both motivated and unmotivated agents exert zero social effort. This equilibrium exists for any share μ_k of motivated agents in a given team k.

Lemma 1 (Unproductive Equilibrium). $\forall \mu_k \in [0, 1]$ *there exists a Bayesian Nash equilibrium in which both unmotivated and motivated agents exert zero social effort, i.e.,* $e^*(k, u) = e^*(k, m) = 0$.

Suppose nobody else in the team exerts social effort. In that case, any agent's social effort is completely wasteful, so that not even motivated agents have an incentive to exert

---

Exerting only selfish effort in the high-incentive team generates utility $b_h E$. Therefore, if $\alpha E < b_h$, no agent ever exerts any social effort in equilibrium. We assume strict inequality in (5) to avoid the uninformative knife-edge case in which agents are at most indifferent between exerting some or no social effort.

[6]The utility of the individuals is linear in effort so that an individual gets the same utility if it randomizes over different effort levels and if it only exerts full selfish or full social effort. Therefore, we can ignore equilibria in which individuals randomize over their effort choices without losing any interesting insights.

[7]There is no problem with out-of-equilibrium beliefs in our model. The latter only arise if a team remains empty in equilibrium. If an agent then deviates and joins the empty team, the production of the social good only depends on his own social effort.

social effort. Such an equilibrium exists even if all team members are motivated, and this is common knowledge. The reason is that the complementarity in production generates a coordination problem. However, suppose there are sufficiently many motivated agents in the team. In this case, there also exists a *productive equilibrium*, in which motivated agents exert full social effort, as the next proposition shows.

Lemma 2 (Productive Equilibrium). *If and only if $\mu_k \geq b_k/(\alpha E)$, there exists a Bayesian Nash equilibrium in which unmotivated agents exert zero effort while motivated agents exert full social effort, i.e., $e^*(k, u) = 0$ and $e^*(k, m) = E$.*

Unmotivated agents exert zero social effort in any equilibrium. If all motivated agents exert full social effort, the team's average social effort is equal to $\mu_k E$. The condition on the fraction of motivated agents in the team ensures that motivated agents' net marginal utility to exert social effort is positive.

We now consider team selection. We start by analyzing the case in which monetary incentives to exert selfish effort are strictly larger in one team than in the other, i.e., $b_h > b_l$. We show that there exists a unique Bayesian Nash equilibrium, in which unmotivated agents join the high-incentive team and exert only selfish effort. In contrast, motivated agents join the low-incentive team and exert only social effort. All proofs are in Appendix A.

Proposition 1 (Different Incentives). *Suppose $b_h > b_\ell$ and agents can choose their team. Then there exists a unique Bayesian Nash equilibrium, in which unmotivated agents select the high-incentive team, $p^*(h, u) = 1$, and behave according to the unproductive equilibrium, while motivated agents select the low-incentive team, $p^*(h, m) = 0$, and behave according to the productive equilibrium.*

The fact that agents can choose their team is crucial for Proposition 1. The importance of self-selection is shown by the next proposition that considers the case in which teams have different monetary incentives, but agents are randomly assigned to teams with some fixed probability. Random assignment rules out self-selection, and the fraction of motivated agents is μ_0 in both teams. Lemma 1 and 2 then directly imply the following result.

Proposition 2 (Exogenous Assignment). *Suppose $b_h > b_\ell$ and agents are randomly assigned to the low-incentive team with some common probability $p \in [0, 1]$ and to the high-incentive team with complementary probability $1 - p$. Then the following holds:*

(i) $\forall \mu_0 \in]0, 1[$, *there exists a Bayesian Nash equilibrium in which agents behave according to the unproductive equilibrium in both teams.*

(ii) If $\mu_0 \geq b_\ell/(\alpha E)$, *there exists a Bayesian Nash equilibrium in which agents behave according to the unproductive equilibrium in the high-incentive team, and according to the productive equilibrium in the low-incentive team.*

(iii) If $\mu_0 \geq b_h/(\alpha E)$, *there exists a Bayesian Nash equilibrium in which agents behave according to the productive equilibrium in both teams. Further, there exists a Bayesian Nash equilibrium in which agents behave according to the productive equilibrium in the high-incentive team, and according to the unproductive equilibrium in the low-incentive team.*

Proposition 2 shows that once self-selection is ruled out, there always exists a Bayesian Nash equilibrium in which agents behave according to the unproductive equilibrium in both teams. This result holds even if the ex-ante probability for agents to be motivated is close to one. The reason is that motivated agents need to coordinate on the productive equilibrium. However, with exogenous assignment, it is impossible to update beliefs on other agents' motivation to exert social effort in a given team. If the fraction of motivated agents in the population is low, this equilibrium is the only Bayesian Nash equilibrium under exogenous assignment. If the fraction surpasses a critical threshold, other Bayesian Nash equilibria exist in which agents coordinate on the productive equilibrium in some or even in both teams. There are then multiple equilibria.

We finally analyze the role of different monetary incentives for self-selection into teams and the consequential production of the social good. To do so, we consider the case in which agents can choose their team, but in contrast to Proposition 1 monetary incentives to exert selfish effort are now the same in both teams.

Proposition 3 (Same Incentives). *Suppose* $b_h = b_\ell$ *and agents can choose their team. Then the following holds:*

(i) $\forall \mu_0 \in]0, 1[$, *there exist two Bayesian Nash equilibria in which unmotivated and motivated select into different teams,* $p^*(k, u) = 1$ *and* $p^*(k, m) = 0$ *for some* $k \in \{h, \ell\}$, *and agents behave according to the unproductive equilibrium in the team selected by the unmotivated, and according to the productive equilibrium in the team selected by the motivated agents.*

(ii) $\forall\, \mu_0 \in\,]0, 1[$, *there also exists a continuum of Bayesian Nash equilibria in which unmotivated and motivated agents select into team h with some probability $p^*(h, u) \in [0, 1]$ and $p^*(h, m) \in [0, 1]$, respectively, both teams are never empty, i.e., $p^*(h, u) + p^*(h, m) \in\,]0, 2[$, and agents behave according to the unproductive equilibrium in both teams.*

Part (i) of Proposition 3 confirms that the Bayesian Nash equilibrium with positive contributions to the social good from Proposition 1 also exists if monetary incentives in both teams are the same. The intuition is that motivated agents still have an incentive to separate from unmotivated agents to exert social effort in teams with other motivated agents. Unmotivated agents want to maximize their monetary payoff. If monetary incentives are the same in both teams, they can achieve this by choosing any team and exerting only selfish effort. They have no particular reason to join motivated agents and are happy to leave them alone. Consequently, there exists the equilibrium in which unmotivated and motivated agents separate and in which unmotivated agents exert only selfish effort while motivated agents exert only social effort in their teams.[8]

However, as part (ii) of Proposition 3 shows, if incentives are the same, the equilibrium with positive contributions is no longer unique. Instead, there exists a continuum of other Bayesian Nash equilibria in which no agent exerts any social effort. In the latter equilibria, both teams are never empty, and all agents exert only selfish effort. Therefore, the average social effort of other agents in both teams is zero and cannot be influenced by any agent. It is thus optimal for motivated agents to never exert any social effort. Given these effort choices, all agents are indifferent between both teams because monetary incentives are the same. Any team choices, and therefore also the equilibrium team choices, are optimal. In sum, the proposition shows that the separation and effective production of the social good again poses a coordination problem once monetary incentives are the same in both teams.

Taken together, our theoretical analysis reveals that differences in monetary incentives play a critical *coordinating* role for the effective production of a social good under self-selection. The simple logic is that with different monetary incentives, motivated agents join the team with lower incentives if and only if they expect to be able to produce the social good. Motivated agents know that the low-incentive team is either empty or populated with other motivated agents exerting high social effort. Low monetary incentives thus serve a dual

---

[8]There exist two such equilibria because teams are now identical, and agents can separate in two possible ways.

purpose: on the one hand, they ensure the separation of motivated and unmotivated agents in different teams. On the other hand, they crucially facilitate the coordination of social effort decisions by motivated agents within teams.

The latter mechanism follows from forward induction, and its effectiveness is most strikingly seen once it is switched off. First, if monetary incentives differ but agents are unable to self-select into teams, the zero-social-effort equilibrium exists even if almost everybody in the population is motivated; see Proposition 2 (i). Second, suppose agents can self-select into teams but monetary incentives are the same. Then the zero-social-effort equilibrium exists even if all agents perfectly separate; see Proposition 3 (ii).

We next come to our experiment that we designed to test the theoretical predictions.

3   Experiment

The experimental design closely follows the spirit of our theoretical analysis. By means of three treatments, we identify the causal effect of differences in monetary incentives on the production of social goods allowing for self-selection.

The main structure of the experiment is as follows: After participants have read the written instructions and correctly answered a series of control questions, each participant makes decisions in 20 rounds of the following three-stage game.[9] In the first stage, all participants join one of two possible teams. In the second stage, everybody receives an endowment and decides how much of this endowment he keeps for himself and how much he contributes for generating a charitable donation to the *Deutsche Krebshilfe*. The latter is a well-known German charitable organization that supports the fight against cancer.[10] The allocation decision corresponds to the choice of selfish and social effort in our model above, and the social good is the generated donation. Crucially, the donation a participant generates depends on the share of the endowment the participant contributes and how much the other participants in the same team contribute. In the third stage, all participants receive information on the donation and the private income they have generated and on the average behavior (team

---

[9]Experimental instructions are provided in Appendix C.

[10]See *www.krebshilfe.de* for further information. In the instructions we included a short paragraph summarizing the main mission of the charity, so that participants were fully informed about the purpose of the donations they generate in the experiment. Further, they were given the opportunity to contact us after the experiment to get a receipt documenting that donations were indeed made.

choice and average contribution in both teams) of other participants in the given round.

We now explain in more detail the three stages in each round. We first focus on our *Main Treatment* with self-selection and differing incentives and then explain our two control treatments *Exogenous Assignment* and *Same Incentives*.

3.1   Main Treatment: Different Incentives

In stage 1 of the *Main Treatment*, participants simultaneously and individually decide which of the two teams they join in a given round. We use neutral team labels A and B. Participants cannot communicate with each other during the experiment, such that they cannot coordinate directly.

In stage 2, all participants receive an endowment of $E = 10$ points. They can use this endowment to generate a donation to the *Deutsche Krebshilfe* or to earn a private monetary income. Participants simultaneously enter an integer between 0 and 10, indicating how much of their endowment they contribute to generate a donation. The donation a participant generates is equal to his contribution multiplied by the average contribution of the other participants in the same team. If a participant is alone in a team, the generated donation is the squared contribution.[11] Participants do not know how many other participants are in their team when deciding about their contribution, and they do not know the contributions of any other participants in their team.

While the production function for the social good is the same in both teams, monetary incentives and thus income opportunities differ in the *Main Treatment*. In particular, every point that is not contributed to generate a donation yields private monetary earnings of $b_\ell = 5$ points in Team A and yields private monetary earnings of $b_h = 7$ points in Team B. Thus, payoff functions are equal to $\Pi_A = 5(10 - e)$ in Team A and $\Pi_B = 7(10 - e)$ in Team B with e denoting an participant's contribution.

In stage 3, we remind participants of their own behavior and inform them about their generated donation and monetary income in a given round. In addition, they receive information about the number of other participants and the average contributions in both teams. Participants are thus able to learn also about the team they did not join in this round. This way, we minimize information frictions.

---

[11]This assumption follows our theoretical model. As shown below, the empirical results do not depend on whether we include observations with only a single participant in a team or not. See footnote 14.

3.2 Control Treatment I: Exogenous Assignment

To test whether self-selection causes behavior in our *Main Treatment*, we implement the control treatment *Exogenous Assignment*. This treatment is identical to the *Main Treatment* except that we assign participants to teams randomly. Participants are assigned with equal probability to each team in every round. Random assignment eliminates selection by design, while the difference in incentives between the two teams is maintained.

3.3 Control Treatment II: Same Incentives

While the control treatment *Exogenous Assignment* allows us to identify the effect of selection, our second control treatment *Same Incentives* enables us to analyze to what extent the difference in monetary incentives promotes sorting and coordination. The treatment *Same Incentives* is identical to the *Main Treatment* except that both teams' monetary incentives are now the same. Every point not contributed to the charity generates a private income of 5 points in either team. Thus, $\Pi_A = \Pi_B = 5(10 - e)$.

It is important to note that we equalize incentives at the low level of 5 points. If we had chosen the high-incentives level 7 instead, opportunity costs of social effort in both teams would have been higher compared to the low-incentive Team A in the *Main Treatment*. This would have created a confound if agents contribute more in team A in the *Main Treatment*, which is what our model predicts. With incentives equal to 5, opportunity costs are identical. Furthermore, the treatment thereby allows us to isolate the effect of low incentives per se.

3.4 Procedural Details

The experiment was programmed in z-tree (Fischbacher, 2007) and conducted in the Frankfurt Laboratory for Experimental Economic Research (FLEX) at Goethe University Frankfurt. Subjects were recruited via ORSEE (Greiner, 2015). Altogether we conducted fourteen sessions, five sessions of both the *Main Treatment* and the treatment *Same Incentives*, and four sessions of the treatment *Exogenous Assignment*. In each session, we had between 22 or 24 participants, who were divided into two independent, equally sized matching groups. On the aggregate level, this leaves us with 28 independent matching group observations, which we will use as the basis of our statistical analysis. All participants were students from different fields at Goethe University Frankfurt. Overall, 322 subjects participated in our experiment.

A session lasted about 35 minutes. For each participant, one of the 20 rounds was randomly chosen to be payoff relevant. Every point was worth 20 euro cents. Individuals generated an average donation of 1.45 euros and earned on average 12.99 euros.

3.5 Hypotheses

Before we present the experimental results, let us summarize our hypotheses based on the theoretical analysis in Section 2. In contrast to the situation in Proposition 1 (which we implement in our *Main Treatment*), where a unique Bayesian Nash equilibrium with positive social effort from motivated agents exists, in Proposition 2 and 3 (which we implement in treatment *Exogenous Assignment* and *Same Incentives*, respectively), this equilibrium either fails to exist or the multiplicity of equilibria, in particular those with zero social effort, render coordination on a positive-social-effort equilibrium difficult or impossible.

Thus, we hypothesize that motivated individuals in the *Main Treatment* select into the low-incentive Team A, and that contributions are higher in the low-incentive Team A than in the high-incentive Team B. In contrast, in *Exogenous Assignment*, where we can rule out selection by design, we expect contributions to be low in both teams and similar to those in the high-incentive Team B in the *Main Treatment*. A similar reasoning holds for the treatment *Same Incentives*, where this time the equality of incentives is hypothesized to render the separation and coordination of high contributors difficult. Thus, we expect contributions also to be low in this treatment and similar to those in the treatment *Exogenous Assignment*. The following hypotheses summarize our expectations for participants' behavior in the experiment.

Hypothesis 1 (Main Treatment). *In the Main Treatment, contributions in the low-incentive Team A are higher than in the high-incentive Team B.*

Hypothesis 2 (Exogenous Assignment). *In the treatment Exogenous Assignment, contributions are the same in both teams and on a similarly low level as in the high-incentive Team B in the Main Treatment.*

Hypothesis 3 (Same Incentives). *In the treatment Same Incentives, contributions are the same in both teams and on a similarly low level as in the high-incentive Team B in the Main Treatment.*

4 Results

We discuss our empirical results in two parts. We first study the role of incentives and self-selection for the production of the social good on an aggregate level. This will answer our three hypotheses above. Subsequently, we analyze participants' behavior on an individual level, allowing us to isolate the mechanism through which different incentives affect social effort decisions.

4.1 Team Choices and Social Effort

Figure 1 shows average team choices and social effort decisions within teams for all rounds in all three treatments.
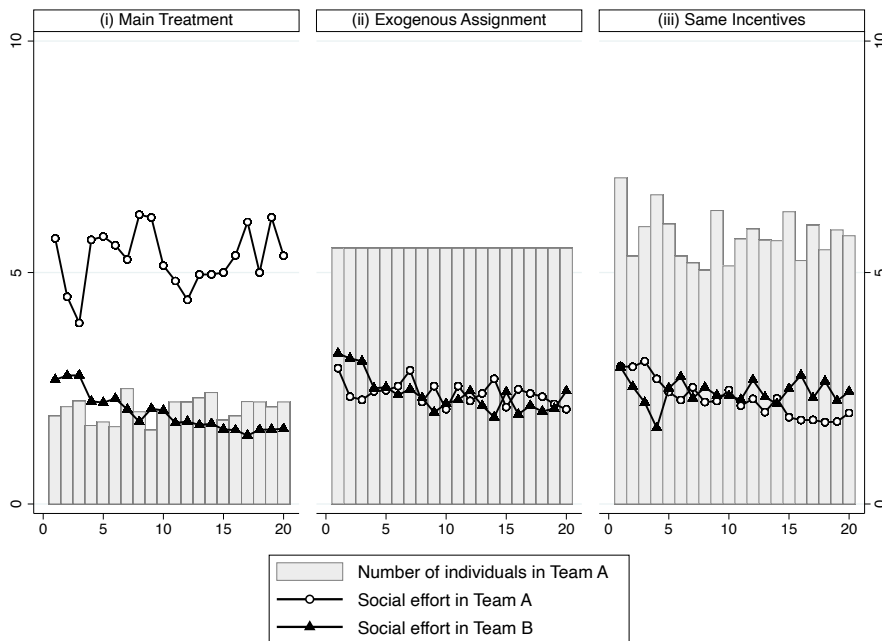

Figure 1: Average Team Choices and Social Effort

Notes: Average team choices and social effort decisions within teams for all rounds and treatments. Bars are the average number of participants in Team A. Lines marked with a white circle are average social effort decisions in Team A, lines marked with a dark triangle are average social effort decisions in Team B. All averages are matching group averages.

Two findings are immediately apparent when comparing behavior in the *Main Treatment*

15

(panel (i)) to the two control treatments (panels (ii) and (iii)). First, significantly fewer participants join Team A (grey bars) in the *Main Treatment*. Second, social effort in Team A in the *Main Treatment* (line marked with white circles) is about twice as high compared to social effort in the other teams and treatments. We now look at these findings in detail.

On average, about two participants (17 percent) choose the low-incentive Team A in the *Main Treatment*. In contrast, a large majority of participants (83 percent) prefer the high-incentive Team B.[12] Aggregate team choices are very stable over the 20 rounds, i.e., there are no time trends and no major fluctuations (Wilcoxon signed-rank test comparing the first and last ten rounds, $p = 0.39$).[13] However, these findings do not imply that the participants in Team A are necessarily the same in each round. We come to this when we analyze individual behavior in the next section. While only a minority of participants choose Team A in the *Main Treatment*, average social effort in this team is equal to 5.19 and thus more than twice as high than social effort in the high-incentive Team B, which is 1.96.[14] The difference is highly statistically significant (see Table 1). We observe no significant change in social effort over rounds in Team A but a significant decline of around 35% in Team B.[15] The larger fluctuations of social effort in Team A stems from the smaller number of participants that self-select into this team.

These observations in the *Main Treatment* are precisely in line with Hypothesis 1. However, we cannot conclude yet that the difference in social effort is caused by self-selection. The reason is that participants might exert more social effort in Team A not because of selection, but because monetary incentives and thereby opportunity costs of social effort (once individuals are in this team) are lower. To identify the selection effect, we need to look at our first control treatment *Exogenous Assignment*, where we rule out selection by design while monetary incentives differ between teams.

Panel (ii) in Figure 1 shows participants' behavior in *Exogenous Assignment*. Average social effort in Team A and B over all rounds is 2.42 and 2.37, respectively, and the small

---

[12]All observations are matching group averages based on 11 to 12 participants in each session.

[13]All non-parametric tests in our analysis are based on comparing the distributions of independent matching group averages. We report p-values from two-sided tests, rounded to two digits. We report p-values smaller than 0.005 as 0.00.

[14]Excluding all rounds, in which there is only a single participant in Team A, gives us an average social effort equal to 5.27 in this team.

[15]Wilcoxon signed-rank test comparing the first and last ten rounds, $p = 0.65$ and $p = 0.01$, respectively.

16

	Team A	Team B	p-value (A vs. B)
Main Treatment	5.19	1.96	0.00
Exogenous Assignment	2.42	2.37	0.40
Same Incentives	2.25	2.32	0.96
p-value (Main vs. EA)	0.00	0.25	
p-value (Main vs. SI)	0.00	0.23	

Table 1: Average Social Effort

Notes: Average social effort over all rounds in Team A and Team B for all treatments. p-values from Wilcoxon signed-rank test for comparison of teams within treatments and Mann-Whitney tests for comparison of treatments for a given team.

differences are not statistically significant (cf. Table 1).[16] These levels are significantly lower than social effort in the low-incentive Team A in the *Main Treatment*, but not significantly different from social effort in the high-incentive Team B in this treatment. Again, see Table 1 for test results. Together, these findings corroborate Hypothesis 2 and show that the differences in social effort between the low- and high-incentive team in our *Main Treatment* are indeed caused by selection and not by the differences in monetary incentives per se once individuals are in a given team.

Finally, we compare behavior in the *Main Treatment* with the treatment *Same Incentives* to identify the role of different levels of incentives for participants' selection and coordination. Recall that different monetary incentives are not a necessary condition for the existence of a separating equilibrium with positive social effort (Proposition 3). Yet, they are necessary to render the equilibrium unique, thereby solving the coordination problem among motivated individuals (Proposition 1). Our second control treatment *Same Incentives* thus enables us to test the importance of different incentives for coordination by allowing participants to self-select into teams while eliminating the difference in monetary incentives between them. Put differently, we deprive participants of the opportunity to coordinate their social efforts within teams via different monetary incentives.

Panel (iii) in Figure 1 illustrates behavior in the treatment *Same Incentives*. Over all

---

[16]We observe no decline in the low-incentive Team A but a significant decline in the high-incentive Team B (Wilcoxon signed-rank test, $p = 0.58$ and $p = 0.04$, respectively).

rounds, about six participants (48 percent) in each matching group self-select into Team A on average in this treatment. This fraction is the same when comparing the first to the last ten rounds (Wilcoxon signed-rank test, $p = 0.88$). Thus, participants spread more or less equally across teams, which is not surprising because teams are now identical. Concerning average social effort, the data show that contributions are very similar and low in both teams: 2.25 in Team A and 2.32 in Team B, and the difference is not statistically significant (see Table 1).[17] Behavior in treatment *Same Incentives* is consistent with Hypothesis 3 and shows that differences in monetary incentives are crucial for the effective coordination on the high-effort equilibrium.

In sum, our results so far show that the production of the social good remains at relatively low levels as long as participants cannot self-select into teams or monetary incentives in the two teams are equally low. Once monetary incentives are high in one of the two teams and participants can self-select, social good production in the low-incentive team increases by more than 130 percent. These findings confirm the predictions of our model which argues that only under differing incentives, high monetary incentives attract unmotivated individuals, creating scope for motivated individuals to self-separate and coordinate on the high-social-effort equilibrium in the low-incentive team. We next present further evidence on the individual level showing that the combination of self-separation and coordination indeed drives the observed effect.

4.2 Self-separation and Coordination of Motivated Agents

This section focuses on two questions. First, can we distinguish between motivated and unmotivated individuals in the experiment? Second, do motivated individuals exert higher social effort when they are in Team A than in Team B? Our theory argues that the answer to both questions is yes, but only in the *Main Treatment*.

With respect to the first question, consider Figure 2 illustrating how often individual participants are in Team A in the three treatments. In the *Main Treatment* (left panel (i)) and in treatment *Same Incentives* (right panel (iii)), observations are based on self-selection decisions. In treatment *Exogenous Assignment* (middle panel (ii)), participants are randomly assigned to the two teams. Panel (i) in Figure 2 shows that 59 of the 118

---

[17]There is a significant decline in Team A but not in Team B (Wilcoxon signed-rank test, $p = 0.01$ and $p = 0.65$, respectively.
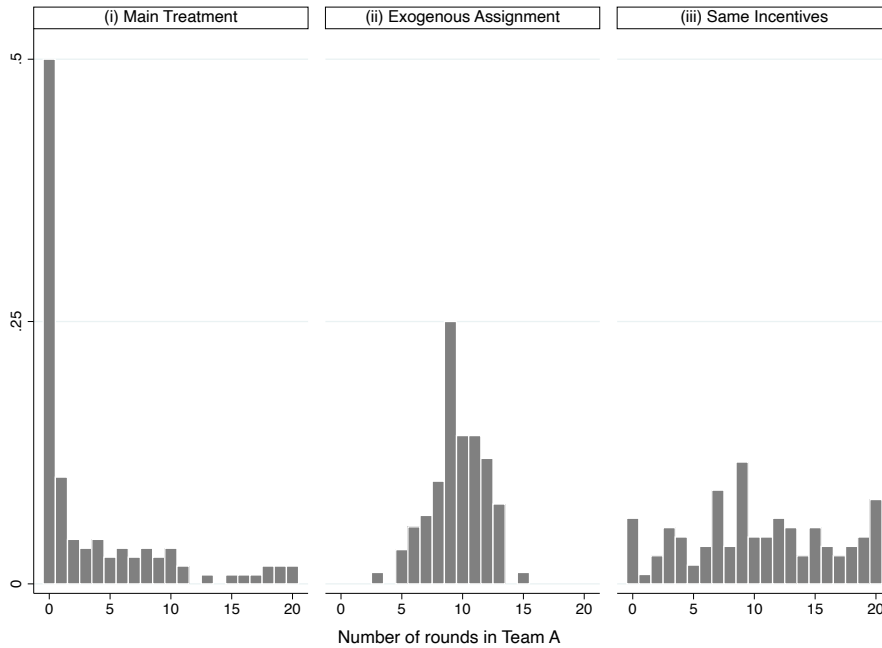
Figure 2: Distribution of Number of Rounds Individuals are in Team A

Notes: Distribution of the number of rounds individuals are in Team A. Bars are the fraction
of individuals for a given number of rounds.

individuals who participate in the *Main Treatment* (i.e., 50 percent) never choose the low-incentive Team A. The remaining participants self-select into Team A at least once, with a considerable heterogeneity in the number of rounds a participant chooses to be in A. For example, around 8 percent of participants self-select into Team A in more than 15 rounds, while about 40 percent choose A in between 1 and 10 rounds. Thus, half of the participants self-select into the high-incentive team for the entire experiment, while almost all of the remaining participants switch back and forth between both teams.[18]

Participants in the *Main Treatment* might select into Team B because they are primarily interested in monetary incentives and care less about the production of the social good. Indeed, Table 2 (upper panel) shows that participants who always choose Team B exert about 60 percent lower social effort on average than participants who select into Team A at least once. The respective average effort levels are 1.39 vs. 3.67, and the difference is

---

[18]Only two participants in our *Main Treatment* never choose the high-incentive Team B.

highly statistically significant. Furthermore, individual average social effort is positively and significantly correlated with the number of rounds an individual chooses Team A (Spearman correlation: 0.64, $p = 0.00$). In the following, we therefore label participants who are always in Team B as *unmotivated* and participants who select into Team A at least once as *motivated*.

	Overall	Team A	Team B	p-value
Main Treatment				
Sometimes in Team A (*motivated*)	3.67	5.19	2.86	0.00
Always in Team B (*unmotivated*)	1.39		1.39	
p-value	0.00		0.01	
Exogenous Assignment				
Sometimes in Team A	2.39	2.42	2.37	0.46
Always in Team B	n.a.	n.a.	n.a.	
p-value				
Same Incentives				
Sometimes in Team A	2.44	2.25	2.37	0.85
Always in Team B	1.76		1.76	
p-value	0.38		0.47	

Table 2: Motivation and Average Social Effort in Teams

Notes: Average social effort over all rounds conditional on sorting behavior and current team choice. p-values from Wilcoxon signed-rank tests.

Whereas team choices in *Main Treatment* allow us to distinguish between motivated and unmotivated individuals, this distinction should not be possible in the two control treatments. In *Exogenous Assignment*, information about the number of rounds a participant is in Team A does not tell us anything about his or her motivation because participants are randomly assigned to teams. Figure 2 (panel (ii)) shows that almost all participants are between 5 and 15 rounds in Team A. Further, the correlation between the number of rounds a participant is in Team A and individual average social effort is insignificant (Spearman correlation: 0.06,

$p = 0.56$).

Next, in *Same Incentives* theory predicts that a separation of motivated and unmotivated individuals is difficult, and also the previous analysis has shown no indication for any separation according to motivation. Panel (iii) in Figure 2 corroborates this with data on the individual level demonstrating that the distribution of the number of rounds a participant is in Team A is more or less flat, quite in contrast to the distribution we observe in the *Main Treatment*. Because incentives are the same, unmotivated individuals are indifferent between teams implying that we should be unable to distinguish between types. Indeed, the correlation between average individual social effort and the number of rounds a participant is in Team A is insignificant (Spearman correlation: -0.09, $p = 0.35$). Further, there is no difference in social effort if we apply the same labeling procedure as above (see Table 2, lower panel): Participants who always remain in Team B over all rounds, exert social effort equal to 1.76 while those, who switch, exert 2.44, and the difference is insignificant.[19] The fact that effort levels of participants who switch tend to be higher, although not significantly, suggests that some individuals might try to coordinate on higher social effort levels, but obviously are not successful.

We can now answer the second question: Do motivated individuals in the *Main Treatment* exert higher social effort in Team A than in Team B? The answer is yes, see Table 2. Motivated individuals, i.e., those who are in Team A at least once, exert social effort equal to 5.19 in Team A but only 2.86 in Team B. The difference is highly statistically significant providing clear evidence for successful coordination on high effort in Team A.

Figure 3 portrays the distributions of social effort decisions in the *Main Treatment*, with individual bars representing the fraction of decisions in which participants of a particular type exert a given social effort level. We find that in about 20 percent of the cases, motivated individuals exert the maximum social effort of 10 when they are in Team A (see middle panel (ii) in Figure 3).[20] In contrast, an even higher share of observations from the *same* group of participants shows zero social effort when in Team B (right panel (iii)). Compare this to the social effort distribution of unmotivated individuals, who always remain in the high-incentive team (left panel (i)). More than 35 percent of these observations show zero

---

[19]The same holds, if we pool participants, who always remain in either A or B, and compare their behavior to those who switch (1.49 vs. 2.49, Wilcoxon signed-rank test $p = 0.28$).

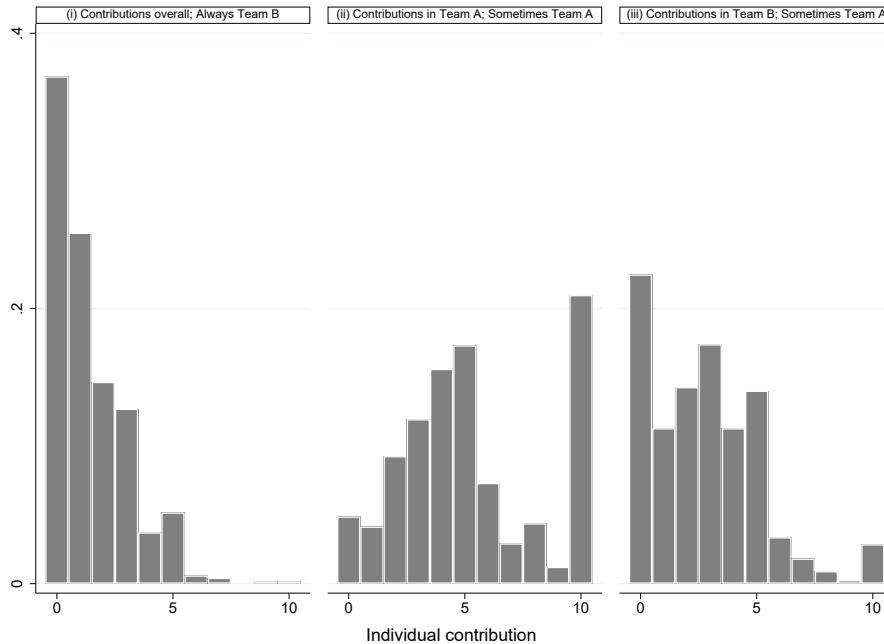[20]Distributions for all treatments are shown in Figure 4 in Appendix B.

Figure 3: Distribution of Social Effort in the Main Treatment

Notes: Distribution of social effort decisions conditional on type and current team in the Main treatment. Bars are the fraction of decisions in which participants of a particular type exert a given effort level.

social effort and we almost never see any effort larger than 5. Interestingly, social effort of motivated individuals in Team B is also significantly higher than social effort of unmotivated individuals (cf. Table 2 for averages and test results). However, the gap is smaller than the difference in the social effort of motivated individuals in the low- and high-incentive team.

Finally, a regression analysis in Table 3 with individual social effort as dependent variable corroborates our results. In the regression we include individual fixed effects and cluster standard errors on the level of matching groups.[21]

Columns (1) analyzes the *Main Treatment*. The dummy variable *Team A* indicates whether a participant is in Team A in the current round. The respective coefficient is estimated based on the behavior of the motivated participants who switch. Results show that

---

[21]Using the wild cluster bootstrap-t procedure proposed by Cameron et al. (2008) to account for the relatively low number of clusters does not change any significance levels. Results are available upon request.

	Main Treatment	Exogenous Assignment	Same Incentives
	(1)	(2)	(3)
Team A	1.46***	-0.06	-0.07
	(0.40)	(0.09)	(0.22)
Constant	2.29***	2.41***	2.38***
	(0.07)	(0.04)	(0.12)
Number of Observations	2360	1840	2240
within R^2	0.07	0.00	0.00
between R^2	0.44	0.01	0.00
overall R^2	0.24	0.00	0.00

Table 3: Individual Fixed Effects Regression: Social Effort

Notes: Individual fixed effects regression with social effort as dependent variable. Column (1) considers data from the *Main Treatment*, column (2) from *Exogenous Assignment*, and column (3) from *Same Incentives*. The dummy variable *In Team A* indicates whether the individual is in Team A in the current round. Standard errors are in parentheses and are clustered at the matching group level. *** indicate statistical significance at the 1% level. The number of observations differ between columns, because we ran four sessions in *Exogenous Assignment* and five sessions in each of the other two treatments.

individuals exert on average 1.46 units more in social effort in Team A than in Team B and the coefficient is significant on the 1 percent level. Thus, motivated individuals coordinate on significantly higher social effort when choosing the low-incentive team. Note that the overall R^2 is high and in particular the variation across individuals is quite well explained. In contrast, there are no effects in the control treatments. Neither in the treatment *Exogenous Assignment* (column (2)) nor in *Same Incentives* (column (3)) do participants contribute significantly more when they are in *Team A*. Further, the R^2 is basically zero in these regressions, both within and between individuals.[22]

In sum, the results show that motivated individuals coordinate on the high-social-effort equilibrium in the low-incentive team — yet, only if they can self-select, and if incentives between teams differ such that unmotivated individuals are attracted by the high-incentive

---

[22]Table 4 in Appendix B replicates the results based on pooling the three treatments and considering interaction effects.

team. This finding confirms that the self-selection into differently incentivized environments is the key mechanism that is responsible for the production of the social good in the experiment.

5 Conclusion

This paper highlights an important role of incentives in the production of what we call social goods. Social goods are goods and services that benefit from positive effort complementarities but at the same time entail significant intrinsic motivational aspects such that not everybody can be expected to be equally motivated to contribute to their production. A prime example is public service delivery as, e.g., in education, healthcare, and social work.

Our theoretical and empirical results show that the provision of low monetary incentives (in the presence of high monetary incentives elsewhere) has a positive effect on social good production by attracting and coordinating intrinsically motivated agents who exert high effort because they correctly expect other intrinsically motivated agents to do the same. The necessary condition for this effect to materialize is that agents can self-select into their labor environment and, importantly, unmotivated agents find low incentives unattractive, i.e., there exist alternative options that provide higher utility to them.

Taken together, the results provide a novel perspective on the interplay of intrinsic motivation, monetary incentives, and labor productivity emphasizing the role of sorting and coordination of heterogeneously motivated agents. While relevant, our results are not limited to the public service sector. More generally, they apply to all organizations that seek individuals with an intrinsic motivation to support their mission. These missions could include a specific focus on customer support, an inclusive work environment, and producing products or services that serve a certain social purpose. Our findings demonstrate that incentive schemes, which are sufficiently unattractive for unmotivated types, can be an important element in shaping organizational cultures that foster the effective pursuit of such missions.

There exist various directions for further research. It would be intriguing, for example, to investigate the main matching mechanism of sorting and coordination in natural labor environments in the field. Public service delivery seems a promising context in this respect as recent studies by Deserranno (2019) and Ashraf et al. (2020) show. Yet, isolating the effects in the field remains a challenging exercise. One possibility could be to manipulate the level of effort complementarity exogenously in different treatments. Alternatively, beliefs

about the effort of other team members could be elicited and varied. In both cases, theory suggests that the positive effect of low incentives on selection and production is stronger the higher the complementarity and the higher beliefs about others' effort. Next, the role of alternative job options (in particular for unmotivated agents) presents an important research area. On the one hand, it seems likely that such options vary significantly between countries and contexts. On the other hand, they are key for the effect of low incentives. If there is no possibility to lure unmotivated agents away, low monetary incentives are expected to have little effect. Finally, monetary incentives are controlled exogenously in our experiment. A next step forward is to allow contracts and incentives to emerge endogenously. Kosfeld and von Siemens (2009) show that different levels of incentives are an equilibrium outcome. However, it is an open question whether such contracts arise (and survive) if decision makers are given the possibility to choose them. In fact, wages may not be the only way to induce the sorting of motivated and unmotivated agents. Alternative instruments might exist that can be tested both in the lab and in the field.

References

Aimone, J. A., Iannaccone, L. R., Makowsky, M. D., and Rubin, J. (2013). Endogenous Group Formation via Unproductive Costs. *Review of Economic Studies*, 80(4):1215–1236.

Ashraf, N., Bandiera, O., Davenport, E., and Lee, S. S. (2020). Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services. *American Economic Review*, 110(5):1355–1394.

Auriol, E. and Brilon, S. (2014). Anti-social behavior in profit and nonprofit organizations. *Journal of Public Economics*, 117(C):149–161.

Bandiera, O., Barankay, I., and Rasul, I. (2007). Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment. *Quarterly Journal of Economics*, 122(1):729–773.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social Connections and Incentives in the Workplace: Evidence From Personnel Data. *Econometrica*, 77(4):1047–1094.

Banuri, S. and Keefer, P. (2016). Pro-social motivation, effort and the call to public service. *European Economic Review*, 83:139–164.

Barfort, S., Harmon, N. A., Hjorth, F., and Olsen, A. L. (2019). Sustaining Honesty in Public Service: The Role of Selection. *American Economic Journal: Economic Policy*, 11(4):96–123.

Barigozzi, F. and Burani, N. (2019). Competition for talent when firms' mission matters. *Games and Economic Behavior*, 116:128–151.

Bénabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520.

Besley, T. J. and Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *The American Economic Review*, 95(3):616–636.

Brekke, K. A., Hauge, K. E., Lind, J. T., and Nyborg, K. (2011). Playing with the good guys. A public good game with endogenous group formation. *Journal of Public Economics*, 95(9-10):1111–1118.

Brekke, K. A. and Nyborg, K. (2008). Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics*, 30(4):509–526.

Buurman, M. and Dur, R. (2012). Incentives and the Sorting of Altruistic Agents into Street-Level Bureaucracies*. *The Scandinavian Journal of Economics*, 114(4):1318–1345.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *Quarterly Journal of Economics*, 128(3):1169–1218.

Delfgaauw, J. and Dur, R. (2008). Incentives and Workers' Motivation in the Public Sector*. *Economic Journal*, 118(525):171–191.

Deserranno, E. (2019). Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda. *American Economic Journal: Applied Economics*, 11(1):277–317.

Dohmen, T. and Falk, A. (2010). You Get What You Pay For: Incentives and Selection in the Education System. *Economic Journal*, 120(546):F256–F271.

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

Francois, P. (2000). 'Public service motivation'as an argument for government provision. *Journal of Public Economics*, 78(3):275–299.

Friebel, G., Heinz, M., Krueger, M., and Zubanov, N. (2017). Team Incentives and Performance: Evidence from a Retail Chain. *American Economic Review*, 107(8):2168–2203.

Friebel, G., Kosfeld, M., and Thielmann, G. (2019). Trust the Police? Self-Selection of Motivated Agents into the German Police Force. *American Economic Journal: Microeconomics*, 11(4):59–78.

Gächter, S. and Thöni, C. (2005). Social learning and voluntary cooperation among like-minded people. *Journal of the European Economic Association*, 3(2-3):303–314.

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

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, 111(3):465–497.

Hanna, R. and Wang, S.-Y. (2017). Dishonesty and Selection into Public Service: Evidence from India. *American Economic Journal: Economic Policy*, 9(3):262–290.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *JL Econ. & Org.*, 7:24.

Ichniowski, C., Shaw, K. L., and Prennushi, G. (1997). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review*, 87(3):291–313.

Iranzo, S., Schivardi, F., and Tosetti, E. (2008). Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data. *Journal of Labor Economics*, 26(2):247–285.

Itoh, H. (1991). Incentives to Help in Multi-Agent Situations. *Econometrica*, 59(3):611–636.

Kosfeld, M. and von Siemens, F. A. (2009). Worker self-selection and the profits from cooperation. *Journal of the European Economic Association*, 7(2-3):573–582.

Kosfeld, M. and von Siemens, F. A. (2011). Competition, cooperation, and corporate culture. *RAND Journal of Economics*, 42(1):23–43.

Kremer, M. (1993). The O-Ring Theory of Economic Development. *Quarterly Journal of Economics*, 108(3):551–575.

Prendergast, C. (2007). The Motivation and Bias of Bureaucrats. *American Economic Review*, 97(1):180–196.

Serra, D., Serneels, P., and Barr, A. (2011). Intrinsic motivations and the non-profit health sector: Evidence from Ethiopia. *Personality and Individual Differences*, 51(3):309–314.

von Siemens, F. A. and Kosfeld, M. (2014). Team production in competitive labor markets with adverse selection. *European Economic Review*, 68(C):181–198.

Appendix A: Proofs

5.1 Proof of Proposition 1

First, consider the unmotivated agents. In any Bayesian equilibrium, they choose zero social effort, $e^*(k, u) = 0$ in both teams $k \in \{\ell, h\}$ and join the high-incentive team, $p^*(h, u) = 1$, because $b_h > b_\ell$ implies $b_h E > b_\ell E$.

Next, consider the motivated agents. First, we show that $p^*(h, m) = 1$ cannot hold in any Bayesian equilibrium: motivated agents join the low-incentive team with strictly positive probability. Note that the utility a motivated agent gets in any team increases in the average social effort in the team. The high-incentive team attracts all unmotivated agents, and these exert only selfish effort. The average social effort in the high-incentive team thus cannot exceed $\mu_0 E$. The maximum equilibrium utility motivated agents can get in the high-incentive team is the maximum of $b_h E$ and $\alpha \mu_0 E^2$. If $p^*(h, m) = 1$, the low-incentive team is empty. A motivated agent can then switch to the empty low-incentive team and produce the social good by herself to get utility αE^2. The latter exceeds her maximum equilibrium utility in the high-incentive teams because of condition (5). Therefore, motivated agents always joining the high-incentive team cannot be a Bayesian equilibrium.

Second, we show that $e^*(\ell, m) = E$ in any Bayesian equilibrium: motivated agents exert full social effort in the low-incentive team. Suppose $e^*(\ell, m) < E$. By their linear utility function, agents are then either indifferent between all social effort levels, or they strictly prefer exerting only selfish effort. Their utility cannot exceed $b_\ell E$. However, $b_\ell E$ is less than $b_h E$ they can get by exerting only selfish effort in the high-incentive team. Motivated agents exert full social effort in the low-incentive team because they could otherwise get a higher utility from joining the high-incentive team.

Thirdly, we show that $p^*(h, m) = 0$ in any Bayesian equilibrium: motivated agents always join the low-incentive team. Because the average social effort in the low-incentive team is E, motivated agents get equilibrium utility αE^2 when joining the low-incentive team. Assumption (5) ensures that this utility is strictly higher than the maximum of $b_h E$ and $\alpha \mu_0 E^2$ that they can get in the high-incentive team. Motivated agents thus always join the low-incentive team. *Q.E.D.*

5.2 Proof of Proposition 2

Because of the exogenous team assignment that is unconditional on the agents' types, the fraction of motivated agents in any team k equals the prior probability for agents to be motivated. Lemma 1 directly implies the existence of a Bayesian equilibrium in which agents behave according to the unproductive equilibrium in both teams. Lemma 2 implies that if $\mu_0 \geq b_\ell/(\alpha E)$, there exists a Bayesian equilibrium in which agents behave according to the productive equilibrium in the low-incentive team and the unproductive equilibrium in the high-incentive team. Similarly, if $\mu_0 \geq b_h/(\alpha E)$, there exists a Bayesian equilibrium in which agents behave according to the productive equilibrium in both teams, or in which they behave according to the productive equilibrium in the high-incentive team and the unproductive equilibrium in the low-incentive team. *Q.E.D.*

5.3 Proof of Proposition 3

First, consider the unmotivated agents. They choose $e^*(k,u) = 0$ for both $k \in \{\ell, h\}$ and exert only selfish effort in any Bayesian equilibrium. They get the same utility $b_h E = b_\ell E$ in both teams if $b_h = b_\ell$. Any team choice is optimal for them, also separation from the motivated agents.

Second, consider the motivated agents in the Bayesian equilibrium characterized in the first part of the proposition. A motivated agent gets equilibrium utility αE^2 by joining team k and then exerting full social effort $e^*(k,u) = E$. Exerting full social effort in team k is optimal by (5) because the average social effort is E in the team. If the agent switches to the other team, it is optimal for her to exert only selfish effort because the other team's agents exert only selfish effort. Switching to the other team generates a deviation payoff of at most $b_k E$, which is smaller than the equilibrium utility by (5). It is optimal for the motivated agents to separate themselves from the unmotivated agents and then exert full social effort.

Thirdly, consider the motivated agents in the Bayesian equilibrium characterized in the second part of the proposition. The condition $p^*(h,u) + p^*(h,m) \in]0,2[$ implies that none of the two teams is empty. In both teams average social effort is equal to zero. It is then optimal for a motivated agent to exert only selfish effort in either team, $e^*(h,m) + e^*(\ell,m) = 0$. The equilibrium utility is $b_h E = b_\ell E$ in both teams for the motivated agents. Because they are indifferent between the two teams, any team choice is optimal. *Q.E.D.*

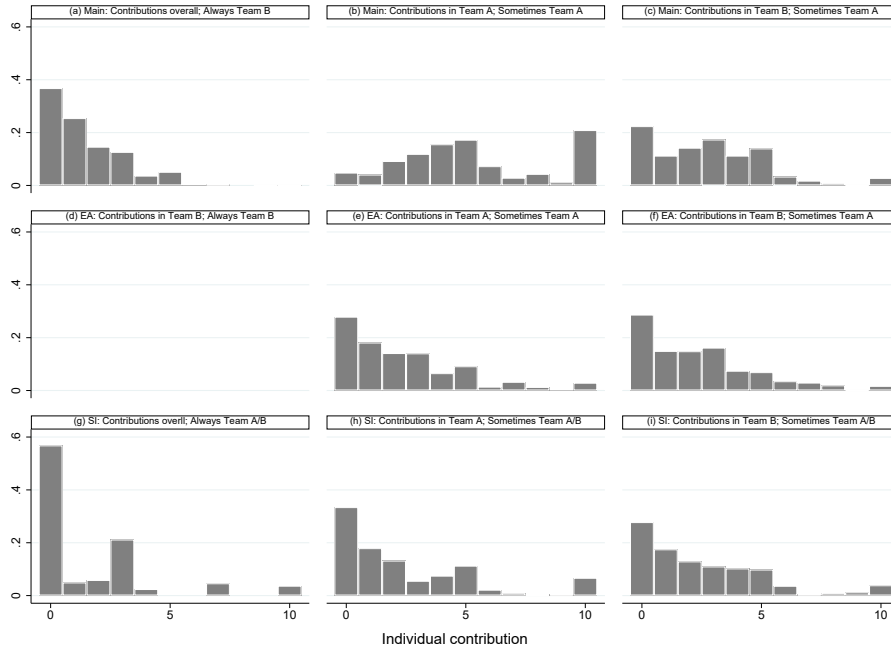# Appendix B: Additional Figures and Tables

Figure 4: Distribution of Social Effort in all Treatments

Notes: Distribution of social effort decisions conditional on type, current team, and treatment. Bars are the fraction of decisions in which participants of a particular type exert a given effort level.

	Benchmark	
	Exogenous Assignment	Same Incentives
	(1)	(2)
Team A	0.03	-0.14
	(0.09)	(0.64)
Main Treatment	-0.40	-0.45
	(0.36)	(0.37)
Main Treatment × Team A	3.28***	3.45***
	(0.42)	(0.76)
Same Incentives	0.05	
	(0.46)	
Same Incentives × Team A	-0.16	
	(0.64)	
Exogenous Assignment		-0.05
		(0.46)
Exogenous Assignment × Team A		0.17
		(0.64)
Constant	2.37***	2.42***
	(0.32)	(0.33)
Number of Observations	6440	6440
adjusted R^2	0.09	0.088

Table 4: Individual Regression: Social Effort

Notes: Linear regression with social effort as dependent variable. Benchmark treatment is *Exogenous Assignment* in column (1) and *Same Incentives* in column (2). The dummy variable *In Team A* indicates whether the individual is in Team A in the current round. Standard errors are in parentheses and are clustered at the matching group level. *** indicate statistical significance at the 1% level.

Appendix C: Instructions

Please find below the English translation of the experimental instructions. The original instructions are in German and available upon request. The parts that were specific for the different treatments are in brackets and in different color. The black parts are the same in all treatments. The below remarks and comments were of course not included in the instructions. Subjects also filled out a short questionnaire before and at the end of the experiment. The questionnaires are available upon request.

General Information

Thank you very much for your participation in today's experiment. In this experiment, you can earn money and at the same time generate donations for the German Cancer Aid (Deutsche Krebshilfe). More information about the German Cancer Aid can be found below.

During the experiment, all income and donations will be given in points. Each point has a value of 20 cents. The points will be converted into Euro at the end of the experiment and then paid to you in cash or donated by us to the German Cancer Aid.

We guarantee that at no time another experiment participant will be informed of your personal decisions. During the experiment, we ask you not to communicate with other participants. Please also switch off your mobile phone. In case you have a question, please raise your hand. We will then come to your table and answer your question personally.

Information on the German Cancer Aid

Every year, around 500,000 new cases of cancer are diagnosed in Germany. Helping those affected and their relatives is the most important concern of the German Cancer Aid. According to the motto "Help. Research. Inform", the non-profit organization supports projects to improve prevention, early detection, diagnosis, therapy, medical aftercare and psychosocial care, including cancer self-help. The German Cancer Aid is the most important private donor in the field of cancer research in Germany. Source: http://www.krebshilfe.de

Specific information

[Exogenous Assignment] The experiment consists of a total of 20 rounds. In each round, you will be randomly assigned either Team A or Team B. Within a team you can generate donations for the German Cancer Aid and at the same time income for yourself. The two teams differ in how your personal income is calculated.

After you have been randomly assigned to one of the two teams, you have 10 points available in each round. Your decision will be to determine how many of these points you will provide for the production of your personal donation to the German Cancer Aid. The remaining points will automatically be used to generate your personal income.

[Main Treatment and Same Incentives] The experiment consists of a total of 20 rounds. In each round, you have to join either Team A or Team B. Within the selected team you can generate donations for the German Cancer Aid and at the same time income for yourself. The two teams differ in the way they calculate their personal income.

After you have decided on one of the two teams, you have 10 points available in each round. Your decision will be to determine how many of these points you will provide for your personal donation to the German Cancer Aid. The remaining points will automatically be used for your personal income.

How is your personal donation to the German Cancer Aid calculated?

The amount of your personal donation in both teams depends on how many points you provide and how many points the other members of your chosen team provide on average for the donation. Your points are multiplied by the average contribution of the other team members. The resulting amount is your personal donation to the German Cancer Aid in this round. If there is no other participant in your team, the points you provide will be multiplied by itself to calculate your donation. These rules apply to both teams.

How is your personal income calculated?

[Main Treatment and Exogenous Assignment] Any point you do not provide for the production of your donation will automatically be used for generating your personal income. The amount of your personal income depends on the choice of the team. In team A, each point is multiplied by a factor of 5. In team B, each point is multiplied by a factor of 7. The resulting amount is your personal income in this round.

The following diagram summarizes the calculation of your personal donation and income in the teams:

Team A

Your personal donation to the German Cancer Aid

= (number of points provided by you) · (average points provided
by the other members of Team A)

Or if you're the only member of Team A:

= (number of points provided by you)^2

Your personal income

= 5 · [10 - (number of points you have provided)]

Team B

Your personal donation to the German Cancer Aid

= (number of points provided by you) · (average points provided
by the other members of Team B)

Or if you're the only member of Team B:

= (number of points provided by you)^2

Your personal income

= 7 · [10 - (number of points you have provided)]

The following examples are supposed to illustrate the calculation.

Example 1: Assume that you have been randomly assigned to Team A. You yourself provide 4 points for the donation and keep 10 - 4 = 6 points for yourself. The other members of Team A provide on average 5 points. Thus, your personal donation for the German Cancer Aid amounts to 4 · 5 = 20 points. Your personal income amounts to 5 · 6 = 30 points.

Example 2: Assume that you were randomly assigned to Team B, as were two other participants. You provide 8 points for the donation and keep 10 - 8 = 2 points for yourself. The other members of Team B provide an average of 7 points. Thus, your personal donation for the German Cancer Aid amounts to 8 · 7 = 56 points. Your personal income is 7 · 2 = 14 points.

Example 3: Assume that you are the only person randomly assigned to Team B. You provide 3 points for the donation and keep 10 - 3 = 7 for yourself. Then your personal donation to the German Cancer Aid is 3 · 3 = 9 points. Your personal income is 7 · 7 = 49 points.

[Comment Same Incentives] The instructions are essentially the same for our Same Incentives treatment as for our Main treatment, with the only difference that the number of points that participant receive for every point not donated is 5 points in both Team A and Team B.

Personal income and donation

Altogether the experiment goes over 20 rounds. In each round, you have 10 points at your disposal. At the end of the experiment, one round is selected at random. Your decision in this specific round is then relevant for your personal income and donation to the German Cancer Aid. Since each of the 20 rounds can be relevant with equal probability, you should make your decisions carefully in each round.

At the end of the experiment, you will receive your personal income in cash.

Your personal donation will be carried out by us for you. We guarantee that the amount you generate will be donated to the German Cancer Aid. If you would like to find out more, please contact us by e-mail (Blinded for review).