STOCKHOLM SCHOOL OF ECONOMICS

Department of Economics

5350 Master's thesis in economics

Academic year 2017–2018

Information Design for the Provision of Public Goods under Interdependent preferences

Julia Baumann (41034)

Abstract

This master's thesis uses an information design approach to improve contributions in a public goods game where players have interdependent preferences. The idea is that an information designer commits to a decision rule giving action recommendations such that players will contribute more often to a public good. The players are conditional cooperators in the sense that they want to contribute together with their own type and prefer not to collaborate with a different type. There is uncertainty about the type they are collaborating with. In a model that is based on Bergemann and Morris (2016a), we first characterize the set of attainable equilibria and then identify the best possible outcome for the information designer. We find that a designer can improve expected utility in a voluntary contribution mechanism by giving action recommendations that satisfy the players' incentive compatibility constraints. Furthermore, the information design problem is studied with other mechanisms, which gives the result that the designer's utility is highest in a mechanism with an entry fee when the public good is excludable.

Keywords: Information design, Asymmetric information, Public goods provision, Voluntary contribution mechanism, Social preferences
JEL-Codes: D82, D83, H41.

Supervisor: Albin Erlanson and Alia Gizatulina (University of St. Gallen)

Date submitted: May 14, 2018

Date examined: May 30, 2018

Discussant: Kajsa Schrewelius

Examiner: Maria Perrotta Berlin

Contents

1 Introduction and motivation **1**

1.1 Overview . . . 1

1.2 Contributions to public goods . . . 2

1.2.1 Mechanisms in public goods provision . . . 3

1.2.2 Incentives in public goods provision . . . 4

2 The information design approach: Theory and literature **8**

2.1 The information design problem . . . 8

2.1.1 Introduction and an example . . . 8

2.1.2 A formal representation . . . 11

2.2 Information design literature . . . 14

3 Model **22**

3.1 Information design with a VCM . . . 23

3.1.1 Payoffs under a VCM . . . 23

3.1.2 Signal distribution . . . 24

3.1.3 Decision rule . . . 25

3.1.4 Obedience constraints . . . 25

3.1.5 Analysis of optimal contribution probabilities . . . 27

3.2 Excludable good . . . 33

3.2.1 Obedience constraints . . . 34

3.2.2 Analysis . . . 35

3.3 Only the high type can produce the good on his own . . . 36

3.3.1 Obedience constraints . . . 37

3.3.2 Analysis . . . 38

3.4 Comparison of the designer's payoffs . . . 39

3.5 Example . . . 41

3.6 Discussion . . . 51

4 Conclusion **56**

References **57**

Appendix **62**

List of Tables

1 Example: The firm's payoff in the investment game 9

2 Example: The signal structure $\sigma(a|\theta)$ for the investment game 10

3 Coordination game 13

4 Payoffs in VCM 24

5 Signal distribution 25

6 Decision rule 26

7 Optimal decision rule 31

8 Payoffs in entry fee mechanism 33

9 Payoffs with asymmetric abilities to produce the good 36

10 Parameter values for example 41

11 Example: Payoffs in VCM 42

12 Example: Best attainable decision rule with information design in VCM ($q = \frac{1}{3}$) 44

13 Example: Payoffs in entry fee mechanism 45

14 Example: Best attainable decision rule with information design in entry fee mechanism ($q = \frac{1}{3}$) 47

15 Example: Payoffs when only the high type produces good on his own . . . 49

16 Example: Best attainable decision rule with information design with asymmetric abilities to produce the good ($q = \frac{1}{3}$) 51

List of Figures

1 An overview of information design 16

2 Set of BCE in VCM; $q = \frac{1}{3}$ 45

3 BCE in VCM with varying values of q 46

4 BCE in VCM with varying value of q; 2-D 47

5 Set of BCE in entry fee mechanism; $q = \frac{1}{3}$ 48

6 BCE in entry fee mechanism with varying values of q 49

7 BCE in entry fee mechanism with varying values of q; 2-D 50

8 Set of BCE with only $\overline{\theta}$ producing good; $q = \frac{1}{3}$ 52

9 BCE with only $\overline{\theta}$ producing good with varying values of q 53

10 BCE with only $\overline{\theta}$ producing good with varying values of q; 2-D 54

1 Introduction and motivation

1.1 Overview

In a survey among the Linux community, Hertel et al. (2003) study the motivations and determinants for collaboration on the Linux system. The authors of the study find that programmers contribute more effort to the project when they believe that they have the same goals as their team members. This leads them to conclude that the belief in having the same valuation for a project within a work group is a crucial factor for programmers' motivation. In general, collaboration on the production of public goods is an essential part of the modern economy. Especially in the digital space, people working together on a project without direct compensation and without knowing each other has become an important feature. In investigating what determines contribution levels it has been found numerous times that people contribute conditionally on who they are working with.

In this thesis we ask: Can this kind of conditional cooperation be used to increase contribution levels in public goods production under different mechanisms? Our approach employs an information design model that works by changing players' beliefs. Information design is a relatively nascent field that explains how and when a principal can persuade agents to take a certain action by giving them action recommendations. As opposed to mechanism design, which takes the information structure of a game as given, it endogenizes the information structure while keeping the game constant. Importantly, the "information designer" can commit to the signal structure that she will provide the players with before the types of players and the state of the world have realized (Bergemann and Morris, 2018). In information design, there is a general trade-off for the designer between recommending the players to take a desirable action and satisfying their incentive compatibility constraints.

We compare an information designer's ability to influence contribution probabilities across two different one-shot voluntary contribution mechanisms (VCMs) and an entry fee mechanism where free-riding is not possible. We treat a two-player game in all cases. The preference structure is based around the idea of conditional cooperation, where players prefer to collaborate with certain types over others. This can, for example, be interpreted as different valuations for the goals of a group, as it has been shown to be a factor in determining contribution levels to the Linux system by Hertel et al. (2003).

So far, applications for information design have been found, for instance, in lobbying and advertising (Kamenica and Gentzkow, 2011). This thesis explores the potential for information design in public goods provision. We find that an information designer can increase the probabilities for contribution in a two-player game with heterogeneous types, where payoffs depend on the collaborator's type, which is uncertain. Our main game specification is a type of VCM where it is possible to free-ride on the other player's contribution. We compare this to information design in a game where the good is excludable such that only contributors can gain from the good. Lastly, we analyze a VCM where only the high type is able to produce the good on his own.

Many applications rely on these public goods mechanisms and this thesis contributes to the literature by suggesting a starting point for improving the amount of contributions via signals to the players. This is relevant for digital applications as discussed in the case of the Linux system, but can also be transferred to in-person team collaborating efforts where it is a priori unclear whether members share the same valuations for a goal.

This thesis is structured as follows: The remainder of Chapter 1 motivates our information design model for public goods provision by introducing public goods mechanisms and giving some background on the empirical public goods literature. Chapter 2 is dedicated to presenting the theory of information design and reviewing the literature on this topic. Finally, Chapter 3 presents, analyzes, and discusses our model. The thesis concludes with some directions for future research and a brief summary.

1.2 Contributions to public goods

The "zero contribution thesis" predicts that a rational actor with self-regarding preferences will not contribute to a public good unless the group of players is rather small or there are external devices to incentivize contributions (Olson, 1965; Ostrom, 2000). The incentive to free-ride for each individual will cause the project to be unsuccessful on the whole. On the one hand, this prediction largely fails in reality: We observe charitable giving, contributions to political campaigns, collaboration on open source software and countless other instances where people decide to put in effort and/or money to produce what can be regarded as a public good. On the other hand, free-riding and thus the inefficient provision of public goods can also be widely observed from experimental laboratory studies to teamwork settings.

This free-rider problem stems from non-excludability, which is one of the two characteristics that define a pure public good (Cowen, 1992, p.3). The second feature is that it is non-rivalrous in consumption. A good that is non-rivalrous but excludable is called a club good.[1] This distinction will become important when picking a mechanism which produces the good.

1.2.1 Mechanisms in public goods provision

This section briefly introduces different mechanisms that can be used in the provision of public goods. A crucial aspect in this context is whether the good is a pure public good in the sense that it is both non-rival and non-excludable. Since different options are possible when a good is excludable, the choice of mechanism will largely be based on these characteristics. Note that when a good is fully non-rival, the act of exclusion creates inefficiencies: Once the good has been provided, it is costless to admit more agents to the use of the good, who could then derive utility from it. Gailmard and Palfrey (2005) call these *exclusionary inefficiencies*.

A substantial number of mechanisms meant to elicit efficient public goods provision has been derived. Largely, they are based on transfer functions that incentivize truthful reporting of one's type. This includes the famous Vickrey-Clarke-Groves (VCG) mechanism that is based on the principle that all players pay their own externality.[2] Another example is the Ledyard-Groves mechanism which sets up transfers such that payments are made according to the deviation from the average promised payment of all players (Groves and Ledyard, 1977). Alternatively, the provision point mechanism requires that the public good is binary, i.e. it is either provided or not provided. It collects contributions from players but returns them if the total amount of contributions is below the threshold needed to provide the good. Clearly, this can only be done with monetary contributions, not, for example, time.

A classic VCM is based on the utility function:

$$e_i - x_i + m \sum_{j=1}^{N} x_j$$

[1]We do not attempt to analyze rivalrous goods in this thesis. These would either be private goods (excludable) or common-pool resources (non-excludable).

[2]See for example Jehle and Reny (2001) for a textbook treatment

where e_i is the endowment, x_i the contribution and m represents the marginal return to capital, i.e. the benefit from the public good (see for example Fischbacher et al., 2001). This setup can often be observed in experimental studies, some of which we mention in the next section. The unique Nash equilibrium for $m < 1$ is $x_i = 0$ for all i, which is where the zero contribution thesis originates from. In the main part of our model, we use an implicit version of a VCM but add payoffs from social preferences that can make it desirable for players to contribute to the good from a theoretical perspective.

When the good is non-rival but excludable, a club good, it is possible to use an entry fee mechanism. A simple example for this would be a park that charges a fee to be able to enter it (Diamantaras et al., 2009, p.235). Since we attempt to compare information design under different mechanisms, we choose a version of the entry fee mechanism for our model that only allows players to benefit from the good if they choose to contribute (with a binary action space, i.e. contribute or not contribute).

While we compare a voluntary contribution game with an entry fee mechanism, we do not change the amount of the transfer in any game. We are going to assume a fixed cost of contribution in all games. One reason why we do not directly manipulate transfers is because it has been shown that intrinsic motivations may potentially be crowded out, for instance when agents are paid for their efforts. These types of state-dependent or endogenous preferences are the subject of Reeson and Tisdell (2008) (empirical) and Bowles and Hwang (2008) (theoretical), where the policy relevance of motivational crowding out is emphasized. One reason why we think it is beneficial to examine information design in the public goods context is that mechanism design can have unintended effects on the belief structure of players by changing transfers. Information design directly tackles the belief structure and can thus help to achieve a better understanding of decision making in a cooperational context.

1.2.2 Incentives in public goods provision

Our model on public goods provision is motivated by the results from empirical literature in this section. On the one hand, we present papers that investigate the relationship between contributions to public goods and the information structure of the game. These show that signals that increase contributions are not only something that comes from our theoretical model, but are also observed in reality. On the other hand, we provide some motivation for the preference structure of our model using research that states that agents want to contribute depending on who they are cooperating with.

Several studies have examined the effect of varying the players' information structure on contributions to public goods. One part of the literature investigates how social comparison interacts with contributions by providing the players with social information. Frey and Meier (2004) conduct a field experiment using voluntary donations by students. They find that when students receive information indicating that a large amount of other students has made a donation, they are more likely to contribute themselves. A different field experiment conducted by Chen et al. (2010) examines contributions to an online movie rating platform. Users are provided with information about how their contributions, i.e. the number of ratings, range compared to others. This leads to a large increase in contributions from users who are initially below the median and a decrease in ratings submitted by users who are told that they are above the median. The authors interpret this as evidence that personalized social information can increase contributions to public goods. This is in line with an earlier laboratory study by Sell and Wilson (1991) who find that individualized information has the largest positive impact on contributions compared to an aggregate information or no information environment.

An earlier work is a series of laboratory experiments conducted by McCabe et al. (1996). They vary matching protocols, payoffs, and payoff information and find evidence in favor of cooperation under complete information, against game theoretic predictions. Also in a laboratory setting Shang and Croson (2009), investigate whether the contributions of others are complements or substitutes for a player's own contribution. Their findings indicate that players give more when they know that others make high contributions, implying that players' contributions are complements.[3] Thus, it has repeatedly been shown that the players' information structure is a factor in determining contributions, both in field and laboratory experiments.

Further, we present the findings that motivate the preference structure in our model. One factor that has repeatedly been established to be important in determining contributions to public goods is conditional cooperation. Fischbacher et al. (2001) were the first ones to find that the largest fraction of their subjects were conditional cooperators who adjusted the amount contributed to a public good upwards if others gave more.

[3]This partially contradicts Chen et al. (2010) since they find that others' contribution are substitutes for players above the median.

Other literature shows that behavior is not only conditioned on the amount of contributions made but also on the identity of the other players. This group identity condition implies that players make a distinction between in-group and out-group members.[4] A meta-analysis of psychology literature by Balliet et al. (2014) finds that while the evidence on this effect is not entirely consistent, overall there is a moderately sized positive effect on collaboration when cooperating with in-group members compared to out-group members. The effect is found to be especially large when the player's own payoff is affected by the other players' actions.

Economic papers that conduct experiments on this issue include for instance Keser and Van Winden (2000), who observe contributions in a repeated public goods game when the group members remain the same (partner condition) and when the composition changes in every round (strangers condition). Their result is that the stranger treatment shows the usual declining pattern over time, while the partner treatment exhibits high contributions throughout the experiment. Eckel and Grossman (2005) also examine a repeated public goods game which is framed as a team production problem. Thus, they are specifically interested in whether a team identity can prevent shirking and free-riding behavior. They find that random assignment of team identities does not increase cooperation, while factors like common goals do increase team production. In Chen and Li (2009), subjects execute different tasks to enhance team identity. They observe the effect of in-group membership on different social preferences and find that participants matched with an in-group member show higher charity concerns and less envy. Furthermore, they are more likely to reward good behavior toward other team members and are less likely to punish them for misbehavior. Guala and Filippin (2016) examine whether established favorable behavior towards in-group members is the manifestation of actual preferences or merely a heuristic that boundedly rational agents use to simplify decision-making. Their finding is that while there may be a group identity effect in a simple dictator game with two allocation options, this effect disappears when the environment becomes more complex by adding more options.[5]

Akerlof and Kranton (2010 and 2000) have been cultivating the idea that identity matters in group production and that in-group versus out-group identification can greatly

[4] In some experiments they are called "partners" and "strangers".

[5] While these are mostly experiments in repeated games, our model will treat a one-shot game. We thus do not tackle the development of group identity, but assume that there already is a preference for working with a certain type.

change agents' incentives. Akerlof and Kranton (2010) list examples from different companies where it has been shown that a shared identity in the form of norms and goals can increase effort levels.

Thus, we can see that agents behave depending on who they collaborate with. In reality, it will not always be clear who their counterpart is and how much players identify with each others' goals and conceptions of the project. In this case, the players' beliefs are a crucial factor in the efficient provision of public goods since they will determine whether and/or how much they will contribute. This implies that with conditional cooperators, there is support for the possibility to design optimal cooperation beyond providing monetary incentives, as in classical mechanism design. By changing the players' beliefs, the outcome can be changed, as we will show in the next chapter, where we present the theory of information design.

2 The information design approach: Theory and literature

The theory of information design analyzes how a game of incomplete information can be influenced by changing the beliefs of the players. It is an emerging field that is related to mechanism design. However, one can make a distinction between the two: Mechanism design is about optimally setting the rules of a game while taking its information structure as given. Information design, on the other hand, endogenously determines the optimal information structure, but takes the rules of the game as given (Bergemann and Morris, 2018).

In a game of incomplete information, players are uncertain about the state of the world, which determines their payoffs. The information designer can provide them with strategic signals to influence their actions. A crucial feature of the information design problem which distinguishes it from a classical game of incomplete information with communication is that the principal has an informational advantage over the players. Depending on the information designer's objective, she will use this advantage optimally to maximize her expected payoff. This is done by giving action recommendations conditional on the state of the world. This approach is used to analyze the model on public goods provision in Chapter 3.

To clarify the notation, assumptions, and procedure, this section provides a general description of information design. The concepts introduced here will be used throughout the thesis. Sections 2.1 and 2.2 describe the problem and deal with the theoretical background of the problem, while Section 2.3 gives an overview of the information design literature to this date.

2.1 The information design problem

2.1.1 Introduction and an example

Information design refers to a game in which an information designer ("sender") has the option to provide player(s) ("receiver(s)") with information in addition to the one the receivers may or may not already have. This is possible since the sender has an informational advantage over the receivers. She wishes to achieve a certain objective and her payoffs are dependent on the actions that the players take. Therefore, the sender will want to influence these actions, that is, to *persuade* the receivers to take their actions in her favor. If the interests of the sender and the receiver do not align, it is usually optimal for the information designer to not fully disclose all information, but rather to obfuscate some information.

To illustrate this principle, we use a simple example from Bergemann and Morris (2018). Suppose there is a government and a firm. Furthermore, there are two states of the world: good and bad. The government wants to maximize the probability that the firm invests, irrespective of the state. The firm, on the other hand, only finds it profitable to invest when the state is good and will otherwise prefer to do nothing. Thus, there is a conflict of interest between the two actors. The government knows whether the state is good or bad and can send the firm signals about the state to increase the probability of investment. We can see that there is a constraint: For the government, it would be optimal to always tell the firm that the state is good such that the firm follows this recommendation and always invests. However, the firm would then know that the government's advice is not valuable and would not follow the investment recommendation. Thus, the sender is constrained by the fact that the firm has to want to follow the government's recommendation, i.e. the firm should not lose from being obedient. It is intuitive to realize that in the good state, the actors' interests are aligned: Both will want the investment to take place. Thus, the government will fully disclose the correct state in this case. The situation is different in the bad state: The government will try to obfuscate some of the information that it has, so that the firm will still invest. Table 1 gives the firm's payoffs in each state where the payoff in the good state is $0 < x < 1$. Table 2 shows the signal structure on which the government has to decide depending on the state, $\sigma(a|\theta)$. p_B is the probability with which the government recommends the firm to invest despite being in the bad state.

	bad state B	good state G
invest	-1	x
not invest	0	0

Table 1: Example: The firm's payoff in the investment game

Under the assumption that the firm's prior belief is that each state is equally likely, the firm would not invest without signals from the government as its expected payoff from this is negative. When the government sends a signal to invest, this translates into the

	bad state B	good state G
invest	p_B	1
not invest	$1 - p_B$	0

Table 2: Example: The signal structure $\sigma(a|\theta)$ for the investment game

following incentive constraint for the firm:

$$\frac{1}{2}p_B(-1) + \frac{1}{2}x \geq 0 \quad \Leftrightarrow \quad p_B \leq x$$

Since the government wants to set the probability of investment as high as possible, the constraint will be satisfied with equality, and thus $p_B = x$. Thus, the government will give an investment recommendation with probability x in the bad state, which the firm will follow so that the probability of investment is now positive. This concludes the example.

Bergemann and Morris (2018) use a linear programming approach to find the optimal information structure. In the first step, they identify the feasible set of outcomes by setting up obedience constraints and then, in the second step pick the most-preferred outcome by the sender among the feasible set. Apart from this two-step procedure, one can also employ the concavification approach. This uses a geometric analysis to concavify the payoff function as shown by Kamenica and Gentzkow (2011).[6] In short, the designer increases her expected payoff by sending signals to the receiver such that her payoff function becomes concave. This procedure is called *Bayesian persuasion* in the case when there is only one sender and one receiver.[7] However, extending the concavification approach to a game of multiple players is not straightforward, so that the linear programming approach is generally less complex in these cases. The solution that we reach via the latter procedure will still correspond to concavifying the designer's objective function (Bergemann and Morris, 2018, Section 4.3).

Having outlined the general idea of information design and how to solve a problem of this type in principle, we proceed to describe its formal representation.

[6] With seminal work done previously by Aumann et al. (1995).

[7] We follow Bergemann and Morris (2018) in their definition of Bayesian persuasion.

2.1.2 A formal representation

Following Bergemann and Morris (2018), we present the general information design problem and introduce the relevant notation for the remainder of the thesis.

Basic game G

There is a finite number of players, $1, 2, ..., I$, who can take a finite set of possible actions A_i. Furthermore, there are Θ payoff states of the world with θ_i as one individual realization from the state space. The players share the same full support prior $\Psi \quad \Delta(\Theta)$ about the probability that each state will occur.

Thus, we can write the utility function for each player i as:

$$u_i : A \times \Theta \to \mathbb{R}$$

where $A = A_1 \times ... \times A_I$ and $a = (a_1, ..., a_I)$ is a typical element of A.

The information designer has an objective function: $v : A \times \Theta \to \mathbb{R}$.

Information structure S

As common for incomplete information games, we define the information structure S in addition to the basic game G. The belief structure consists of the set of types $t_i \quad T_i$ and the signal distribution that is a mapping from the payoff state space to the players' type space:

$$\pi : \Theta \to \Delta(T)$$

where $T = T_1 \times ... \times T_I$.

The basic game G and the information structure $S = \left((T_i)_{i=1}^I, \pi\right)$ together make up the incomplete information game (G,S).

Timing

A crucial feature of the information design problem is that the designer picks and commits to a rule that specifies how to provide the players with additional information before the state and the types are realized. This commitment to the decision rule σ is the first step, and only afterwards the state θ is realized and the players learn their types t_i.

Knowing their types, the players receive additional information from the designer and pick their action based on all the information available to them now. The payoffs are then realized.

Revelation principle

Under certain assumptions, we can evoke the revelation principle in analogy to mechanism design: This implies that the signal space will be equal to the action space. The idea behind this is that any message will induce an action in equilibrium, so we label messages according to the actions they give rise to. Thus, the decision rule chosen by the designer will give "action recommendations". This principle implicitly makes use of the assumption that the designer is able to choose the equilibrium that is played. Furthermore, all information structures must be feasible, the cost of using information is zero, there is a single information designer and the setting is static (Bergemann and Morris, 2018, Section 1). That the designer can pick the equilibrium may not be realistic and we will comment on this issue later on. For now, as is common in the literature, we maintain this claim and thus make use of the revelation principle.

Decision rule

Based on the revelation principle discussed above, the designer selects the decision rule which can condition on the true state and the type vector:

$$\sigma : T \times \Theta \to \Delta(A)$$

This rule gives the players information about the state of the world, thus influencing the action profile a.

Obedience constraints

The decision rule is constructed to maximize the designer's objective function. However, it is restricted by incentive constraints, or obedience constraints, as Bergemann and Morris (2018) coin it: The players have to want to follow the action recommendations that the designer gives in all cases. In other words, their expected utility from following the recommendation must be higher than taking a different action. Formally, this translates into the following expression from an ex ante perspective (Bergemann and Morris, 2018, Definition 1):

Decision rule σ is *obedient* for (G,S) if, for each $i = 1, ..., I$, t_i T_i, and a_i A_i, we have:

$$\Sigma_{a_{-i},t_{-i},\theta}\Psi(\theta)\pi((t_i, t_{-i})|\theta)\sigma((a_i, a_{-i})|(t_i, t_{-i}), \theta)u_i((a_i, a_{-i}), \theta)$$

$$\geq$$

$$\Sigma_{a_{-i},t_{-i},\theta}\Psi(\theta)\pi((t_i, t_{-i})|\theta)\sigma((a_i, a_{-i})|(t_i, t_{-i}), \theta)u_i((a_i, a_{-i}), \theta)$$

for all a_i A_i.

This can also be written in terms of the interim beliefs with Bayes' rule. However, the resulting inequality then has the same denominator on both sides such that it cancels out, again yielding the expression above. We make use of this fact in our model, which will be written from an interim perspective, i.e. the players know their own type, but not the other player's type.

Solution concept: Bayes Correlated Equilibrium

Bergemann and Morris (2016a) define the solution concept of Bayes correlated equilibrium: It is a version of the correlated equilibrium introduced by Aumann (1974) for incomplete information environments.

Let us first explain the concept of a correlated equilibrium in a complete information setting. We illustrate it with an example of the classic coordination game.

		Player 2	
		A	B
Player 1	A	3, 1	0, 0
	B	0, 0	1, 3

Table 3: Coordination game

The game has three Nash equilibria: (A, A), (B, B) and a mixed equilibrium where A is played with probability $p = \frac{1}{4}$. The expected payoff from the last strategy is $\frac{1}{4} \times 3 + \frac{3}{4} \times 1 = \frac{3}{2}$. Now suppose that there is a correlation device that helps players coordinate on one equilibrium. For example, one might think of the flip of a fair coin: When Head comes up, both players shall play strategy A. With Tail, both play B. Both players now have an expected payoff of $\frac{1}{2} \times 3 + \frac{1}{2} \times 1 = 2$, which is higher than under the mixed strategy. Unilateral deviation from adhering to the correlation device can only lead to a lower

expected payoff. Thus, this is a correlated equilibrium in a complete information setting. Bergemann and Morris (2016a, p. 493) define a *Bayes* correlated equilibrium as follows: *A decision rule* σ *is a Bayes correlated equilibrium (BCE) of (G,S) if it is obedient for (G,S).*
Again, G is the basic game and S the information structure and obedience is defined as in the previous subsection. This means that for a decision rule to be a BCE, it must induce action recommendations from which the players will not want to deviate unilaterally.

2.2 Information design literature

Overview

Having described the basic concept of information design and its components, we now proceed to review some of the literature on the topic. Bergemann and Morris (2018) offer an extensive overview of the current information design literature. In this section we only attempt to summarize some seminal concepts and papers that have the most relevance for our model on public goods provision.
There are multiple different scenarios where information design is relevant. We have already mentioned the most basic scenario developed by Kamenica and Gentzkow (2011) where there is only one sender and one receiver without prior information. There are obvious extensions to this both in terms of the basic structure of the game and the information structure. As Figure 1 shows, one can go from a single-player situation to a multi-player game and from a binary action space to a richer action space.
Of course, there are many options to vary the information structure, as Figure 1 also presents. In particular, the players can have either no prior information about the realization of the state of the world apart from the designer's signal; or their type t_i can induce some additional signal that lets them update their prior about the state. Moreover, the information designer can have different degrees of information about players' types (Bergemann and Morris, 2016b, p. 588):

- *Omniscient*: The sender knows the receivers types and can therefore condition directly on them.
- *With elicitation/private persuasion*: The sender does not know the receivers types but can elicit them by asking to report them. This introduces additional truth-

telling constraints and strictly decreases the set of attainable outcomes compared to the omniscient case.

- *Without elicitation/public persuasion*: Again, the sender does not know the receivers types and in this case cannot elicit them either. Kolotilin et al. (2017) show that under certain conditions, private and public persuasion are equivalent.

In our model, we stick to the assumption of an omniscient designer who has thus perfect information about the social type of the players. This simplifies the analysis significantly, but of course limits the scope of applying the model to reality. We elaborate on this in the discussion.

Furthermore, there may be restrictions on the signals that the designer is able to send. Depending on what degree of correlation between the players' actions the designer wants to induce, she may either want to send private signals or public signals. In some circumstances, private signals may not be possible, thus restricting the set of feasible outcomes when the designer has to stick to signals that are common to all players.

Lastly, the information design problem can be examined in a dynamic setting rather than a static one, which is especially important for long-run processes where experimentation matters (see Hörner and Skrzypacz, 2016).

Having established the variations that are possible in principle, we will examine the literature more closely. Note that we divide papers into rough categories, even though some of the literature may fit into multiple of these baskets. However, we attempt to associate each paper with the category that it focuses on.

As we have already established, the current research on information design builds on the seminal paper by Kamenica and Gentzkow (2011) that found a geometric solution, concavification, to solve the information design problem with a single sender and receiver without private information. Other important conceptual contributions were made in a series of papers by Bergemann and Morris (2018, 2016b, 2016a). They find an approach to categorize new and already existing work on information design that not only includes Bayesian Persuasion, but also the more traditional literature on communication in games (Bergemann and Morris, 2018).[8]

Bergemann and Morris (2016a) define "Bayes correlated equilibria", the set of outcomes

[8] In which the sender is not able to commit to the signal structure before types are realized.

Figure 1: An overview of information design

that can arise in Bayes Nash equilibrium if players observe the given information structure but potentially also additional signals. In this process, they find a definition for what it means to have "more information": Additional information on the side of the players imposes more incentive constraints that decrease the set of possible outcomes. Bergemann and Morris find conditions under which this is the case. Their analysis extends the framework to a many-player structure.

Taneva (2015) follows Bergemann and Morris' methodology and construct the set of BCE under the assumption that agents have no information beyond their prior beliefs. In this manner, she finds the set of all Bayes Nash equilibria without explicitly using information structures.

Alonso and Câmara (2016a) characterize the information problem when sender and receiver have different priors. They find that with heterogeneous priors, full information disclosure can be suboptimal even when preferences are aligned. Gentzkow and Kamenica (2014) and Hedlund (2015) extend the original Bayesian Persuasion framework to a situation where sending signals is costly. This affects the sender's payoff and thus also the optimal information structure. The former paper characterizes cost functions that are compatible with the concavification approach.

More conceptual contributions were made by a number of papers following Kamenica and Gentzkow (2011): Gentzkow and Kamenica (2016) extend the work, characterizing all feasible distributions of the posterior mean. Kolotilin (2015) imposes more structure on the information design problem and derive necessary and sufficient conditions for the sender's and receiver's welfare to be monotonic in information. This paper finds that the sender's welfare increases with the precision of the *sender's* information and decreases with the precision of public information. Under a different set of assumptions and with a linear programming approach, Kolotilin (2016) shows that the expected utilities are not monotonic in the precision of the *receiver's* private information. Kolotilin et al. (2017) establish the equivalence of persuasion mechanisms and experiments when the receiver is privately informed and the sender cannot directly observe the receiver's type. This means that the sender is indifferent between asking the receiver to report his type and then sending a private signal (persuasion mechanism) and giving a public signal which is observed by all receivers equally (experiment).

Gentzkow and Kamenica (2017a) and Gentzkow and Kamenica (2017b) examine the effects of having multiple senders (and one receiver) who thus are in a competition with each other.

Multi-agent games

We now move on to review the literature on many-player frameworks in information design. Our main reference papers Bergemann and Morris (2016b) and Bergemann and Morris (2018), have their focus on BCEs and information structures in games. In particular, they characterize what it means for players to have "more information": "[...] [A]n information structure giving rise to a smaller set of Bayes correlated equilibria in all games corresponds to a *more* informed information structure" (Bergemann and Morris, 2016b, p. 502). They find that a richer information structure on the side of the players actually leads to a lesser amount of attainable outcomes for the designer. This is because more information leads to more constraints on behavior.

As mentioned in Section 2.1.2, Mathevet et al. (2017) also examine games and information design. In a setting with more than one player, there are now not only beliefs about the state, but also higher-order beliefs. That is, beliefs that players have about each others' beliefs (about each others' beliefs and so on). They show that every optimal information structure can be decomposed into a public and a private component in the form of signals to the agents. An important feature of their analysis that distinguishes it from, for instance Bergemann and Morris (2018), is that they do not invoke the revelation principle. They stress that while the revelation principle applies very generally in single-player settings, this may often not be the case in multi-player games. Their approach is also robust to different solution concepts, as for example level-k reasoning.

Arieli and Babichenko (2016) also extend the single player situation to multiple receivers. They investigate the information design problem under different types of payoff functions. Compared to Bergemann and Morris (2016b), their setup is more concrete, so that Arieli and Babichenko are able to characterize the optimal revenue and information disclosure explicitly.

While we present some primarily conceptual papers in this section, more research on many-player information design problem is reviewed under "Applications" since they have a more applied focus.

Dynamic information design

While it is beyond the scope of this thesis, there is also an emerging literature on dynamic information design that gives implications about optimal information disclosure under learning and experimentation. We still mention some of these papers here, since a dynamic setup can easily be justified in public goods provision where a good is produced repeatedly or has to be sustained over time. Dynamic information design is especially relevant for long-run environments where agents may act differently in later periods depending on what happened in earlier periods. A survey containing much of this literature can be found in Hörner and Skrzypacz (2016). Sequential information design in games (i.e. multi-player settings) is examined by Doval and Ely (2016). Incidentally, they use a public goods game as an example to show that outcomes can be improved with sequential information design compared to a static action recommendation from the designer. Their set-up is a classic public goods game where free-riding is possible and there are no social preferences. The uncertainty stems from the value of the public good as opposed to our model, where players are uncertain about their collaborators' types. The idea is that the designer not only determines the release of the information on the state, but also fixes the timing of play. It is found in the example that as the number of stages goes to infinity, the number of attainable contributions goes to an upper bound value.

Other papers study situations where the designer observes a stochastic process or has the option to experiment (privately). Ely (2017) states that because of the evolving state, the agents' beliefs will evolve even without additional information from the designer. The designer has the possibility to shape the path of belief evolution but finds that there is a trade-off between current persuasion and persuasion in the future. An interesting application is found by Ely et al. (2015): Drawing from the idea that people derive utility from suspense and surprise in entertainment, they model the demand for information over time. They solve for the optimal information revelation that maximizes expected suspense or surprise for Bayesian subjects.

Che and Hörner (2018) note the increasing importance of recommender systems in the digital space and derive optimal information design for a situation where users learn collaboratively about a product. Here, the designer's trade-off is between full disclosure and over-recommending a product whose quality has not been explored yet. On the one hand there is a risk to derive disutility from this new product (thus decreasing the users trust in the platform), on the other hand the product's quality needs to be explored for the community. The optimal mechanism finds a balance between these effects.

The optimal information disclosure policy in a setting where the principal wants to maximize social welfare within a recommender system is characterized in Kremer et al. (2014). Agents' incentives to explore and generate new information are affected by the information structure. They assume agents to be self-interested and myopic and find that a threshold policy maximizes welfare, while neither no disclosure nor full disclosure is optimal.

This section provides evidence on dynamic information design. One can see that this research is primarily very recent and there are a lot of powerful applications to be discovered. We emphasize the potential for public goods provision with social preferences, but leave this extension for future research.

Applications

Having reviewed the primarily conceptual side of the literature, we now present papers that apply the general conceptual results to a more specific environment. Coordination games and their application to a potential regime change are examined in Goldstein and

Huang (2016). Goldstein and Huang present a model where the designer is a policymaker who would like to keep the current regime, while the agents potentially prefer to abandon the regime depending on fundamentals. In this setup, it is optimal for the policymaker to sometimes abandon the regime themselves to decrease the probability that agents will attack. Ostrovsky and Schwarz (2010) analyze the phenomenon that universities intentionally hold back information on their students' achievements towards potential employers. This obfuscation can, according to their findings, prevent unraveling in the market, i.e. students getting hired in increasingly earlier semesters.

Furthermore, a few papers examine information design in the context of voting, which is another relevant application. One can think of a politician (or multiple candidates) trying to convey signals to voters that make it the most likely to achieve the politician's preferred outcome. The most notable study on this subject is Alonso and Câmara (2016b).

Several studies focus on auctions and information design: Bergemann et al. (2016), Bergemann and Hörner (2018), Bergemann and Wambach (2015) all treat information design in auctions, while Zhang and Zhou (2016) study information design in contests. Perhaps unexpectedly, there are not only applications of information design in microeconomics, but even on a macroeconomic level: In a simple stylized economy, Bergemann et al. (2015) examine aggregate volatility and the effect of agents' information structure on volatility. They consider a setting with symmetric agents with linear best responses. The result is that maximum aggregate volatility is obtained in an information structure where agents are uncertain about the size of the shock to the economy. This leads to overreaction and thus maximum volatility.

Providing optimal information to consumers from the seller's perspective is another application that has warranted some attention. For instance, Li and Shi (2017) show that disclosing different signals to different buyer types dominates a full disclosure policy in a setting where the buyer has incomplete knowledge about his own type.

Little work has been done on combining the mechanism and the information design problem such that the allocation mechanism and the belief structure can be altered simultaneously. Bergemann and Pesendorfer (2007) investigate the optimal choice of information structure and auction format. They find a trade-off between the minimization of information rent and the maximization of allocational efficiency, which is balanced by the optimal information structure. As opposed to Bergemann and Pesendorfer (2007) who assume that the bidders initially have no private information, Daskalakis et al. (2016) analyze the optimal auction and information design when both the seller and the bidders have private information about the object's value.

A strand of the literature on team production examines information sharing in the team production process. A trade-off has been noted between information sharing and adaptation, as for instance Blanes i Vidal and Möller (2016) have described it: Sharing additional information may have detrimental effects on the motivation of the players, but may on the other hand increase the likelihood of adaptation. Thus, some information can intentionally be held back and groups fail to aggregate information efficiently. The mechanism developed in Blanes i Vidal and Möller (2016) determines how much information is to be shared in the optimum, when there is a noncontractible agreement on a project between the players.

Celik and Peters (2016) examine collaboration agreements with reciprocal relationships between players. This means that each player can make commitments based on the other players' commitments while they form a contract among themselves. They find that partial revelation of information at the contracting stage increases the set of supportable outcomes compared to a centralized mechanism.

This section shows that there is a wide range of potential applications where information design can play a role in designing optimal policies. We add to this literature in the next chapter by presenting a model based on a public goods game.

3 Model

This analysis treats the optimal design of information in public goods games. Our main focus is on public goods provision in games that rely on voluntary contributions. Since we cannot or do not want to change this mechanism in many instances, classical mechanism design theory does not generally seem fit for improving contribution levels. There is still scope for information design, however, when the information structure can be optimally adjusted. For instance, this can be the case in digital applications: A platform designer would usually have a lot of power about what users know about each other and could use this flexibility to optimally adjust players' beliefs. However, the notion of an information designer need not necessarily be literal, as we discuss in Section 3.6.

We choose to compare the information design problem under different mechanisms. This allows us to compare combined mechanism design and information design problems when the designer has the option to influence both the payoffs and the belief structure. The following sections first treat the information design problem in the VCM in detail and then proceed to consider an entry fee mechanism where free-riding is not possible and a mechanism where the ability to produce the good is asymmetrically distributed between the players. Lastly, we compare the designer's utility across these cases. The analysis is based on the procedure developed by Bergemann and Morris (2018 and 2016a). The main difference is that in our model, the players' types directly determine the state of the world. In the original model, the state is not directly related to the players. This change implies four different possible states and thus a higher-dimensional problem. Furthermore, the payoff structure that we impose varies from the model by Bergemann and Morris.

This model has possible applications in collaboration processes. For instance, Hertel et al. (2003) find in their survey on participation motives to the Linux kernel that team activities are determined by team goals. They observe that some projects are conducted with a team on subsystems. Within these subsystems, developers are more willing to spend time when they have higher subjective evaluation of the subsystem's goals. That is, contribution activity is determined by how highly individual participants value the teams' goals. Of course, there is some uncertainty involved in this subjective evaluation, since the programmers may not know these objectives perfectly a priori. Thus, their decision to contribute will be shaped by their beliefs about the team's goals. If they value these expected goals highly, they will contribute more.

However, this models' applications are not limited to the digital space. One may also think of a situation in a more general team production process. For example, in a group project where individual members can potentially have different valuations of the outcome, it may be uncertain before the project is executed whether the goals are shared or not. This becomes clear only during the project and team members gain more utility from having worked with someone with the same valuation.

3.1 Information design with a VCM

3.1.1 Payoffs under a VCM

The first scenario consists of a simplified VCM with a nonexcludable good. Players can either contribute ("C") or not contribute ("NC"). The good is produced whenever there is at least one contribution. There are two players of type $\theta = \{\underline{\theta}, \overline{\theta}\}$ which implies that they have either low or high valuation for the good. Both players incur a cost c when they decide to contribute. The players have a preference for collaborating with their own type, but not with the type that has a different valuation. This assumption represents the feature that people like to be part of a like-minded community. Thus, they receive a premium z when contributing together with the same type but a penalty y when they contribute together with a different type. We call the different situations the social ($\underline{\theta\theta}$,$\overline{\theta\theta}$) and the asocial state ($\underline{\theta}\overline{\theta}$,$\overline{\theta}\underline{\theta}$). Furthermore, when only player contributes on their own, they also incur a penalty (x) since they see the other player free-riding. They achieve the payoffs depicted in Table 4.[9]

The information designer is omniscient in the sense that she knows the players' type with certainty. Furthermore, it is possible to send private signals to each player. The parameters are such that:

- $0 < \underline{\theta} < \overline{\theta}$.
- $(x, y, z, c) > 0$.
- $z > c$: This implies that free-riding is not a best response in the social state.
- From the points above it follows that $\underline{\theta} - c + z > \underline{\theta} > 0$ (and thus $\overline{\theta} - c + z > \overline{\theta} > 0$).
- $\overline{\theta} - c - y < 0$ (and thus $\underline{\theta} - c - y < 0$).

[9]Nash equilibria are marked in color throughout the thesis.

$\underline{\theta}, \underline{\theta}$	C	NC
C	$\underline{\theta} - c + z, \underline{\theta} - c + z$	$\underline{\theta} - c - x, \underline{\theta}$
NC	$\underline{\theta}, \underline{\theta} - c - x$	0,0

$\underline{\theta}, \overline{\theta}$	C	NC
C	$\underline{\theta} - c - y, \overline{\theta} - c - y$	$\underline{\theta} - c - x, \overline{\theta}$
NC	$\underline{\theta}, \overline{\theta} - c - x$	0,0

$\overline{\theta}, \underline{\theta}$	C	NC
C	$\overline{\theta} - c - y, \underline{\theta} - c - y$	$\overline{\theta} - c - x, \underline{\theta}$
NC	$\overline{\theta}, \underline{\theta} - c - x$	0,0

$\overline{\theta}, \overline{\theta}$	C	NC
C	$\overline{\theta} - c + z, \overline{\theta} - c + z$	$\overline{\theta} - c - x, \overline{\theta}$
NC	$\overline{\theta}, \overline{\theta} - c - x$	0,0

Table 4: Payoffs in VCM

- $\underline{\theta} - c - x < 0$ and $\overline{\theta} - c - x > 0$: In the asocial state, it is a best response for the high type to contribute when the other player does not contribute. For the low type, it is a dominant strategy not to contribute against a high type in the asocial state.

This implies that both players prefer (C, C) in the social states. However, (C, C) is the worst outcome in the asocial state. The high type would still like to contribute since he gets a positive payoff (despite the penalty x), but only when the low type does not contribute.

3.1.2 Signal distribution

In addition to knowing their own types, the players get a signal t_i with $i \in \{s, a\}$ which gives them information about being either in the social or asocial state. According to the signal distribution below, the players will always observe the same signal and will know with certainty that they are in the asocial state when they observe the signal t_a. This distribution follows Bergemann and Morris (2016a).

The common prior probability for the social and asocial state is Ψ and $1 - \Psi$, respectively. $(\underline{\theta}, \underline{\theta})$ and $(\overline{\theta}, \overline{\theta})$ are equally likely, as are the states $(\underline{\theta}, \overline{\theta})$ and $(\overline{\theta}, \underline{\theta})$. For example, the probability for $(\underline{\theta}, \underline{\theta})$ is then $\frac{1}{2}\Psi$.

This signal distribution can, of course, be varied. For example, it is possible that players observe asymmetric signals or that they never know for sure what state they are in after observing the signals. For now, we keep this signal distribution constant across the different public goods games.

$\pi(\cdot \vert \underline{\theta},\underline{\theta})$	$(t_s, \underline{\theta})$	$(t_a, \underline{\theta})$
$(t_s, \underline{\theta})$	1	0
$(t_a, \underline{\theta})$	0	0

$\pi(\cdot \vert \underline{\theta},\overline{\theta})$	$(t_s, \overline{\theta})$	$(t_a, \overline{\theta})$
$(t_s, \underline{\theta})$	$1 - q$	0
$(t_a, \underline{\theta})$	0	q

$\pi(\cdot \vert \overline{\theta},\underline{\theta})$	$(t_s, \underline{\theta})$	$(t_a, \underline{\theta})$
$(t_s, \overline{\theta})$	$1 - q$	0
$(t_a, \overline{\theta})$	0	q

$\pi(\cdot \vert \overline{\theta},\overline{\theta})$	$(t_s, \overline{\theta})$	$(t_a, \overline{\theta})$
$(t_s, \overline{\theta})$	1	0
$(t_a, \overline{\theta})$	0	0

Table 5: Signal distribution

3.1.3 Decision rule

The decision rule determines the contribution probabilities in each state and depends on which signal has been observed by the players. Here, we restrict attention to the case when t_s has been observed since only then the players are still unsure about the actual state. This leads the designer to setting up four different decision rules, one for each state. Clearly, though, the two asocial states are completely symmetric, so that the parameters will be the same in $(\underline{\theta}, \overline{\theta})$ and $(\overline{\theta}, \underline{\theta})$. In the case of the two social states, we can also find symmetry: While the payoffs differ in $(\underline{\theta}, \underline{\theta})$ and $(\overline{\theta}, \overline{\theta})$, the desirable equilibrium is (C, C) in both cases, and so we can say that the social states are equivalent in terms of the decision rule. Thus, there are only five parameters to solve for (see Table 6).
γ_i is the probability that both players contribute in the high state, while α_i is the probability that an individual player contributes. The probability that only one player contributes is then $\alpha_i - \gamma_i$. There is only one value for individual contributions (α_s) in the social state since the two players are symmetric. In the asocial state, there is a probability that the high type contributes (α_1) and a probability that the low type contributes (α_2).[10]

3.1.4 Obedience constraints

The designer's recommendations are constrained by the requirement that the players will want to follow her action recommendation. That is, their expected payoff from obeying the designer must be at least as large as the expected payoff from taking the opposite

[10]The notation for the decision rule again follows Bergemann and Morris (2018).

$\sigma(\underline{\theta},\underline{\theta},t_s)$	C	NC
C	y_s	$\alpha_s - y_s$
NC	$\alpha_s - y_s$	$1 + y_s - 2\alpha_s$

$\sigma(\underline{\theta},\overline{\theta},t_s)$	C	NC
C	y_a	$\alpha_2 - y_a$
NC	$\alpha_1 - y_a$	$1 + y_a - \alpha_1 - \alpha_2$

$\sigma(\overline{\theta},\underline{\theta},t_s)$	C	NC
C	y_a	$\alpha_1 - y_a$
NC	$\alpha_2 - y_a$	$1 - y_a - \alpha_1 - \alpha_2$

$\sigma(\overline{\theta},\overline{\theta},t_s)$	C	NC
C	y_s	$\alpha_s - y_s$
NC	$\alpha_s - y_s$	$1 + y_s - 2\alpha_s$

Table 6: Decision rule

action. These incentive constraints are written out below for each type and each action recommendation. The LHS depicts the expected payoff from contributing while the RHS gives the expected payoff from not contributing.

For type $\overline{\theta}$:

- When the action recommendation is C

$$\begin{aligned}&\Psi\left(y_s(\overline{\theta} - c + z) + (\alpha_s - y_s)(\overline{\theta} - c - x)\right) + \\ &+ (1-\Psi)(1-q)\left(y_a(\overline{\theta} - c - y) + (\alpha_1 - y_a)(\overline{\theta} - c - x)\right) \\ &\geq \Psi(y_s\overline{\theta}) + (1-\Psi)(1-q)(y_a\overline{\theta})\end{aligned} \tag{1}$$

- When the action recommendation is NC

$$\begin{aligned}&\Psi\left((\alpha_s - y_s)(\overline{\theta} - c + z) + (1 + y_s - 2\alpha_s)(\overline{\theta} - c - x)\right) + \\ &+ (1-\Psi)(1-q)\left((\alpha_2 - y_a)(\overline{\theta} - c - y) + (1 + y_a - \alpha_1 - \alpha_2)(\overline{\theta} - c - x)\right) \\ &\leq \Psi((\alpha_s - y_s)\overline{\theta}) + (1-\Psi)(1-q)((\alpha_2 - y_a)\overline{\theta})\end{aligned} \tag{2}$$

Symmetrically, the obedience constraints for type $\underline{\theta}$ are:

- When the action recommendation is C

$$\begin{aligned}&\Psi\left(y_s(\underline{\theta} - c + z) + (\alpha_s - y_s)(\underline{\theta} - c - x)\right) + \\ &+ (1-\Psi)(1-q)\left(y_a(\underline{\theta} - c - y) + (\alpha_2 - y_a)(\underline{\theta} - c - x)\right) \\ &\geq \Psi(y_s\underline{\theta}) + (1-\Psi)(1-q)(y_a\underline{\theta})\end{aligned} \tag{3}$$

- When the action recommendation is NC

$$\begin{aligned}
&\Psi\left((\alpha_s - y_s)(\underline{\theta} - c + z) + (1 + y_s - 2\alpha_s)(\underline{\theta} - c - x)\right) + \\
&+ (1 - \Psi)(1 - q)\left((\alpha_1 - y_a)(\underline{\theta} - c - y) + (1 + y_a - \alpha_1 - \alpha_2)(\underline{\theta} - c - x)\right) \\
&\qquad\qquad \leq \Psi((\alpha_s - y_s)\underline{\theta}) + (1 - \Psi)(1 - q)((\alpha_1 - y_a)\underline{\theta})
\end{aligned} \tag{4}$$

When these constraints are satisfied, the players will want to follow the information designer's action recommendation. The probabilities that do fulfill these constraints are classified as Bayes Correlated Equilibria (BCE) (Bergemann and Morris, 2016a). This is therefore the set of outcomes that can be attained. In the next section, the designer picks the most preferred outcome according to her utility function.

3.1.5 Analysis of optimal contribution probabilities

In principle, the information designer can have different objective functions that determine how she will optimally pick the decision rule giving the contribution probabilities. In this section, we look at two different objectives: Firstly, we analyze the optimal solution for a designer that wants to maximize the number of contributions by the players, no matter what state they are in. In the second part, the designer would like to obtain a socially optimal outcome by maximizing the sum of players' utilities.

Maximize the number of contributions A designer who wants to maximize the number of contributions has the following utility function:

$$V = \mathbb{1}_{\{a_1 = C\}} + \mathbb{1}_{\{a_2 = C\}}$$

The designer's expected utility in this case is:

$$\begin{aligned}
E[V] = \Psi &\left(\frac{1}{2}(2y_s + 2(\alpha_s - y_s)) + \frac{1}{2}(2y_s + 2(\alpha_s - y_s))\right) + \\
&(1 - \Psi)(\frac{1}{2}\left(q \cdot 1 + (1 - q)(2y_a + (\alpha_1 - y_a) + (\alpha_2 - y_a)\right) + \\
&\frac{1}{2}\left(q \cdot 1 + (1 - q)(2y_a + (\alpha_1 - y_a) + (\alpha_2 - y_a)\right)) \\
&= 2\Psi\alpha_s + (1 - \Psi)\left(q + (1 - q)(\alpha_1 + \alpha_2)\right)
\end{aligned} \tag{5}$$

With signal t_a in the asocial state, there will be exactly one contribution: The high type will contribute, while the low type will not. Signal t_s is observed with probability $1 - q$ in the two asocial states.

Thus, the designer maximizes her utility by maximizing a_i, the probabilities that the individual players contribute. Note that while her priority is actually to maximize γ_i (the probability for joint contribution), this is done by maximizing α_i since γ_i must be smaller or equal to the individual contribution probabilities (see Table 6).

We first analyze the probabilities for the social state.

Proposition 1. *In the state where the designer and the players prefer the same equilibrium outcome, it is optimal not to obfuscate information. The designer thus sets* $\gamma_s = \alpha_s = 1$.

Proof. The designer wants to induce (C, C) and therefore set γ_s as high as possible. We observe that the players' contribution constraints are slackened by simultaneously increasing γ_s and α_s. This is because a positive payoff is attached to contributing together in the social state, which is larger than the payoff from free-riding. We can see this by rewriting Equation (3):

$$\Psi \left(\gamma_s(-c + z) + (\alpha_s - \gamma_s)(\underline{\theta} - c - x)\right) + \\ + (1 - \Psi)(1 - q) \left(\gamma_a(-c - y) + (\alpha_2 - \gamma_a)(\underline{\theta} - c - x)\right) \geq 0$$

Independent of the decision rules in the asocial state, increasing γ_s and minimizing the difference between γ_s and α_s slackens the constraint since $c < z$. This holds analogously for the high type. Therefore, the designer can pick her most preferred decision rule, that is always recommending to contribute to both of them in the social state by setting $\gamma_s = \alpha_s = 1$. □

With the observation that the interests of the designer and the players are aligned in the social state, we use the result from Proposition 1. With some rearrangements, the obedience constraints from the former section then become:

$$\Psi(-c+z)+(1-\Psi)(1-q)\left(y_a(-c-y)+(\alpha_1-y_a)(\bar{\theta}-c-x)\right)\geq 0 \tag{6}$$

$$(1-\Psi)(1-q)\left((\alpha_2-y_a)(-c-y)+(1+y_a-\alpha_1-\alpha_2)(\bar{\theta}-c-x)\right)\leq 0 \tag{7}$$

$$\Psi(-c+z)+(1-\Psi)(1-q)\left(y_a(-c-y)+(\alpha_2-y_a)(\underline{\theta}-c-x)\right)\geq 0 \tag{8}$$

$$(1-\Psi)(1-q)\left((\alpha_2-y_a)(-c-y)+(1+y_a-\alpha_1-\alpha_2)(\underline{\theta}-c-x)\right)\leq 0 \tag{9}$$

Note that Equation (9) is satisfied for all probabilities, since $\underline{\theta}-c-x$ is negative. Furthermore, when Equation (8) is satisfied, then (6) automatically holds, which we show in the Appendix.

The first-best solution from the designer's perspective would be to also set $y_a=\alpha_1=\alpha_2=1$. However, this is not always attainable, since the players can get negative utility from contributing in the asocial state. The obedience constraints are therefore potentially not satisfied. The following proposition states the condition such that the first-best outcome is feasible:

Proposition 2. *A designer maximizing the number of contributions can implement the first-best decision rule $y_s=y_a=1$ if $q\geq 1-\frac{\Psi(c-z)}{(1-\Psi)(-c-y)}$, that is if the signal is accurate enough.*

Proof. For making the contribution probabilities in the asocial state as large as possible, the relevant binding constraint is the obedience constraint for the low type when told to contribute. We know that the high type's contribution constraint will then be automatically satisfied. Setting all parameters equal to one and solving for q yields:

$$\Psi(-c+z)+(1-\Psi)(1-q)(-c-y)\geq 0 \tag{10}$$

$$q\geq 1-\frac{\Psi(c-z)}{(1-\Psi)(-c-y)}$$

□

Note that the threshold value is independent of the players' types. This is because their valuations enter both RHS and the LHS of the obedience constraints so that they cancel out when solving for the cut-off values.

We now analyze the case where the first-best outcome is not attainable: If q is below the threshold, the contribution constraint for the low type is binding and the designer has to choose $y_a < 1$. To maximize the probability that both will contribute, she sets $y_a = \alpha_2$ and solves the low type's contribution constraint (Equation (8)) as an equality:

$$\Psi(-c+z) + (1-\Psi)(1-q)\left(\alpha_2(-c-y)\right) = 0$$

$$\alpha_2 = y_a = \frac{\Psi(c-z)}{(1-\Psi)(1-q)(-c-y)} \tag{11}$$

Again, note that the low type's valuation does not enter this expression. Regarding α_1, the designer will now both want and have to set it equal to 1:

- After maximizing the common contribution probability, the designer sets the individual contribution probabilities as large as possible.
- The no contribution constraint (7) for the high type is only satisfied for the value $\alpha_1 = 1$ when $\alpha_2 = y_a$. To see this examine the constraint and plug in the results just obtained:

$$(1-\Psi)(1-q)\left((\alpha_2 - y_a)(-c-y) + (1+y_a-\alpha_1-\alpha_2)(\bar{\theta}-c-x)\right) \leq 0$$
$$(1+\alpha_2-\alpha_1-\alpha_2)(\bar{\theta}-c-x) \leq 0$$
$$(1-\alpha_1)(\bar{\theta}-c-x) \leq 0$$

  The last equation is only satisfied when $1-\alpha_1 = 0$, since $\bar{\theta}-c-x > 0$.

This concludes the search for the optimal contribution probabilities and we can see the decision rule in Table 7. Under these probabilities, the expected utility of the designer is:

$$E[V] = 2\Psi + (1-\Psi)\left(q + (1-q)(1+\alpha_2)\right) = 2\Psi + (1-\Psi)\left(1 + \frac{\Psi(c-z)}{(1-\Psi)(-c-y)}\right)$$

$\sigma(\underline{\theta},\underline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

$\sigma(\underline{\theta},\overline{\theta},t_s)$	C	NC
C	α_2	0
NC	$1-\alpha_2$	0

$\sigma(\overline{\theta},\underline{\theta},t_s)$	C	NC
C	α_2	$1-\alpha_2$
NC	0	0

$\sigma(\overline{\theta},\overline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

Table 7: Optimal decision rule

Maximize the sum of utilities Now suppose that the designer's utility function is determined by the sum of the players' utilities, i.e. the designer cares about social efficiency. This translates into the objective function:

$$\begin{aligned}
E[V] =&\Psi[\frac{1}{2}(2y_s(\underline{\theta}-c+z)+(\alpha_s-y_s)(2\underline{\theta}-c-x))+\\
&+\frac{1}{2}\left(2y_s(\overline{\theta}-c+z)+(\alpha_s-y_s)(2\overline{\theta}-c-x)\right)]\\
&+(1-\Psi)[q(\underline{\theta}+\overline{\theta}-c-x)+(1-q)(y_a(\underline{\theta}+\overline{\theta}-2c-2z)+\\
&+(\alpha_1-y_a)(\underline{\theta}+\overline{\theta}-c-x)+(\alpha_2-y_a)(\underline{\theta}+\overline{\theta}-c-x))]
\end{aligned} \tag{12}$$

Again, the players never contribute after observing t_a, which implies that there is no utility for the designer in this case. We can see from this that the designer would like to maximize $y_s, \alpha_s, \alpha_1, \alpha_2$, while minimizing y_a. To see this intuitively, let us have a look at the payoff matrices: In the social states, maximizing the utilities means making both players contribute. In the asocial states, (C, C) is the worst possible outcome. The designer prefers (NC, C) or (C, NC). She is indifferent between which type contributes, since the sum of the utilities is the same regardless of who contributes.

Note that the obedience constraints remain unchanged. In other words, the set of attainable outcomes is still the same, but the designer picks a different optimal solution. Again, the interests of the designer align in the social state, where they agree that (C, C) is the preferred equilibrium. This implies that setting $y_s = \alpha_s = 1$ is optimal in this state, as we have observed in Proposition 1.

Secondly, the designer sets $y_a = 0$ to minimize the probability of the socially inefficient equilibrium in the asocial state that both players contribute. Thus, she wants to achieve $\alpha_1 + \alpha_2 = 1$ such that there is always exactly one player that contributes in the asocial state.

Proposition 3. *A designer maximizing the sum of player's utilities can always achieve the first-best payoff.*

Proof. Since the obedience constraints are unchanged, we can use the cut-off value for q from Proposition 2 and distinguish between cases:

- $q < 1 - \frac{\Psi(c-z)}{(1-\Psi)(-c-y)}$:
  - Set $\alpha_2 = 0$ and $\alpha_1 = 1$: Since $\overline{\theta} - c - x > 0$, α_1 can be made arbitrarily large and the contribution constraint for the high type will still be satisfied.
  - $0 < \alpha_2 < \frac{\Psi(-c+z)}{(1-\Psi)(1-q)(\underline{\theta}-c-x)}$: The designer can set the contribution probability of the low type to any number between the two bounds that we established above. Naturally, α_1 is set to $1 - \alpha_2$ in all cases.
  - Set $\alpha_2 = \frac{\Psi(-c+z)}{(1-\Psi)(1-q)(\underline{\theta}-c-x)}$ and $\alpha_1 = 1 - \alpha_2$: From the designer's utility function, it is equivalent whether the high type or the low type contributes, even though the low type incurs a negative payoff from contributing on their own. If the designer decides to put a positive probability on the contribution of the low type, the upper bound for this probability is given by the contribution constraint of the low type.[11]

  All these options lead to the same expected utility for the designer since the sum of the individual contribution probabilities is always 1.

- $q \geq 1 - \frac{-\Psi(-c+z)}{(1-\Psi)(-c-y)}$:
  - None of the contribution constraints is binding as long as $\alpha_1 + \alpha_2 = 1$, which also maximizes the designer's utility.

The no contribution constraints are easily satisfied because the probability for no one to invest is zero: $(1 + \gamma_a - \alpha_1 - \alpha_2) = (1 + 0 - (1 - \alpha_2) - \alpha_2) = 0$.
In either case, the designer sets the joint contribution probability γ_a to zero and the sum of individual contribution probabilities to 1. □

Thus, the designer can achieve the maximal payoff attainable regardless of the signal precision and by shuffling around the individual contribution probabilities depending on

[11]This puts the same requirement on the precision of q as in the previous section.

the situation. While the precision of q potentially limits α_2, α_1 can be set very freely since the high type can gain positive utility even when contributing on his own.

This section shows that a designer who (in some states) works against the interests of the players will find it more difficult to maximize her utility, since the incentive constraints are binding more often. Even though the designer may obfuscate information such that players get negative utility under some outcomes, overall they cannot be worse off than without information design, which is ensured by the obedience constraints.

3.2 Excludable good

We have found an optimal information design solution in the case where all players contribute voluntarily and where free-riding on someone's contribution is possible. In this section, we answer the question of how this information design solution changes if a different mechanism can be used. The game that is now considered is not about a pure public good, but a club good. This means that the good is excludable such that free-riding is not possible since only the players who choose to contribute are able to benefit from the good's value. Depending on the application, it may be a valid option to exclude participants based on low activity, so that this scenario will be relevant.

$\underline{\theta}, \underline{\theta}$	C	NC
C	$\underline{\theta} - c + z, \underline{\theta} - c + z$	$\underline{\theta} - c - x, 0$
NC	$0, \underline{\theta} - c - x$	$0,0$

$\underline{\theta}, \bar{\theta}$	C	NC
C	$\underline{\theta} - c - y, \bar{\theta} - c - y$	$\underline{\theta} - c - x, 0$
NC	$0, \bar{\theta} - c - x$	$0,0$

$\bar{\theta}, \underline{\theta}$	C	NC
C	$\bar{\theta} - c - y, \underline{\theta} - c - y$	$\bar{\theta} - c - x, 0$
NC	$0, \underline{\theta} - c - x$	$0,0$

$\bar{\theta}, \bar{\theta}$	C	NC
C	$\bar{\theta} - c + z, \bar{\theta} - c + z$	$\bar{\theta} - c - x, 0$
NC	$0, \bar{\theta} - c - x$	$0,0$

Table 8: Payoffs in entry fee mechanism

As before, we make the following assumptions about the parameters:

- $0 < \underline{\theta} < \bar{\theta}$.

- $(x, y, z, c) > 0$.
- $z > c$: This implies that free-riding is not a best response in the social state.
- From the points above it follows that $\underline{\theta} - c + z > \underline{\theta} > 0$ (and thus $\overline{\theta} - c + z > \overline{\theta} > 0$).
- $\underline{\theta} - c - x < 0$ and $\overline{\theta} - c - x > 0$: In the asocial state, it is a best response for the high type to contribute when the other player does not contribute. For the low type, it is a dominant strategy not to contribute against a high type in the asocial state.
- $\overline{\theta} - c - y < 0$ (and thus $\underline{\theta} - c - y < 0$).

We use the same signal distribution as in the previous section (see Table 5) and the parameters for the decision rule also follow the same reasoning as under the VCM (Table 6).

3.2.1 Obedience constraints

Again, we start by identifying the set of BCE. The obedience constraints for this payoff structure follow the same logic as in the previous section. We write the payoff from contributing on the LHS for the cases when the type is high or low and when the designer recommends to contribute or not to contribute. Note that not contributing always implies a payoff of 0 with certainty so that the RHS of the constraints is always zero. Thus, the payoff from contributing must be positive for both types when told to contribute and negative when they are told not to contribute.
For type $\overline{\theta}$:

- When the action recommendation is C

$$\begin{aligned}\Psi\left(\gamma_s(\overline{\theta} - c + z) + (\alpha_s - \gamma_s)(\overline{\theta} - c - x)\right) + \\ + (1 - \Psi)(1 - q)\left(\gamma_a(\overline{\theta} - c - y) + (\alpha_1 - \gamma_a)(\overline{\theta} - c - x)\right) \geq 0 \quad (13)\end{aligned}$$

- When the action recommendation is NC

$$\begin{aligned}\Psi\left((\alpha_s - \gamma_s)(\overline{\theta} - c + z) + (1 + \gamma_s - 2\alpha_s)(\overline{\theta} - c - x)\right) + \\ + (1 - \Psi)(1 - q)\left((\alpha_2 - \gamma_a)(\overline{\theta} - c - y) + (1 + \gamma_a - \alpha_1 - \alpha_2)(\overline{\theta} - c - x)\right) \leq 0 \quad (14)\end{aligned}$$

The obedience constraints for type $\underline{\theta}$ are:

- When the action recommendation is C

$$\Psi\left(\gamma_s(\underline{\theta}-c+z)+(\alpha_s-\gamma_s)(\underline{\theta}-c-x)\right)+ \\ +(1-\Psi)(1-q)\left(\gamma_a(\underline{\theta}-c-y)+(\alpha_2-\gamma_a)(\underline{\theta}-c-x)\right)\geq 0 \quad (15)$$

- When the action recommendation is NC

$$\Psi\left((\alpha_s-\gamma_s)(\underline{\theta}-c+z)+(1+\gamma_s-2\alpha_s)(\underline{\theta}-c-x)\right)+ \\ +(1-\Psi)(1-q)\left((\alpha_1-\gamma_a)(\underline{\theta}-c-y)+(1+\gamma_a-\alpha_1-\alpha_2)(\underline{\theta}-c-x)\right)\leq 0 \quad (16)$$

3.2.2 Analysis

We go through the same procedure as in the previous section, with the designer maximizing the number of contributions.

Again, interests are aligned in the social state, which implies that there is no restriction on maximizing γ_s. The designer chooses $\gamma_s=\alpha_s=1$:

$$\Psi(\overline{\theta}-c+z)+(1-\Psi)(1-q)\left(\gamma_a((\overline{\theta}-c-y)+(\alpha_1-\gamma_a)(\overline{\theta}-c-x)\right)\geq 0 \quad (17)$$

$$(1-\Psi)(1-q)\left((\alpha_2-\gamma_a)(\overline{\theta}-c-y)+(1+\gamma_a-\alpha_1-\alpha_2)(\overline{\theta}-c-x)\right)\leq 0 \quad (18)$$

$$\Psi(\underline{\theta}-c+z)+(1-\Psi)(1-q)\left(\gamma_a(\underline{\theta}-c-y)+(\alpha_2-\gamma_a)(\underline{\theta}-c-x)\right)\geq 0 \quad (19)$$

$$(1-\Psi)(1-q)\left((\alpha_2-\gamma_a)(\underline{\theta}-c-y)+(1+\gamma_a-\alpha_1-\alpha_2)(\underline{\theta}-c-x)\right)\leq 0 \quad (20)$$

The derivation of the optimal decision rule in the asocial state is very similar to the previous section. Again, we note that (20) is always satisfied and that (17) holds when (8) is satisfied.

The threshold value of q for the first-best solution is derived exactly as before by setting all parameters equal to 1 and solving the contribution constraint of the low type for q, which gives the value $1-\frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(\underline{\theta}-c-y)}$.

Under the condition that q is below this threshold value, the designer chooses the second-best solution, which implies that the contribution constraint of the low type is solved with equality:

$$\alpha_2 = \gamma_a = \frac{-\Psi(\underline{\theta} - c + z)}{(1 - \Psi)(1 - q)(\underline{\theta} - c - y)} \tag{21}$$

while $\alpha_1 = 1$ with the same reasoning as in the previous case. The decision rule then looks like the one in Table 7, but with a different value for α_2.

A designer that cares about the sum of utilities instead of maximizing contribution probabilities can simply employ a full disclosure policy. This will be such that $\gamma_s = \alpha_s = 1$ and $\gamma_a = \alpha_2 = 0$, $\alpha_1 = 1$. The designer can therefore always achieve (C, C) in the social state and (NC, C), (C, NC) in the states $(\underline{\theta}, \bar{\theta})$,$(\bar{\theta}, \underline{\theta})$, respectively. Since the payers have the same preferred outcomes, the obedience constraints are satisfied with these probabilities and all parties profit from this action recommendation.

3.3 Only the high type can produce the good on his own

Thirdly, we now consider a situation where the high type has the ability to produce the good on his own, while the low type can only produce it with the help of the other player (high or low type). Depending on the application, this may either be the case because the designer determines the mechanism to be this way or because of the nature of the good, which requires at least one player with a high type (e.g. high expertise). The payoff matrix is depicted in Table 9.

$\underline{\theta}, \underline{\theta}$	C	NC
C	$\underline{\theta} - c + z$,$\underline{\theta} - c + z$	$-c$,0
NC	0, $-c$	0,0

$\underline{\theta}, \bar{\theta}$	C	NC
C	$\underline{\theta} - c - y$,$\bar{\theta} - c - y$	$-c$,0
NC	$\underline{\theta}$,$\bar{\theta} - c - x$	0,0

$\bar{\theta}, \underline{\theta}$	C	NC
C	$\bar{\theta} - c - y$,$\underline{\theta} - c - y$	$\bar{\theta} - c - x$,$\underline{\theta}$
NC	0, $-c$	0,0

$\bar{\theta}, \bar{\theta}$	C	NC
C	$\bar{\theta} - c + z$,$\bar{\theta} - c + z$	$\bar{\theta} - c - x$,$\bar{\theta}$
NC	$\bar{\theta}$,$\bar{\theta} - c - x$	0,0

Table 9: Payoffs with asymmetric abilities to produce the good

The payoffs have the same characteristics as in the previous cases:

- $0 < \underline{\theta} < \overline{\theta}$.
- $(x, y, z, c) > 0$.
- $z > c$: This implies that free-riding is not a best response in the social state.
- It follows that $\underline{\theta} - c + z > \underline{\theta} > 0$ (and thus $\overline{\theta} - c + z > \overline{\theta} > 0$).
- $\overline{\theta} - c - y < 0$ (and thus $\underline{\theta} - c - y < 0$).
- $\overline{\theta} - c - x > 0$: In the asocial state, it is a best response for the high type to contribute when the other player does not contribute.

The low type simply incurs the cost c when he contributes on his own since the good is not produced (and there is thus no free-riding, i.e. no penalty x). We again keep the same signal distribution as in the previous cases to make the mechanisms comparable.

3.3.1 Obedience constraints

In principle, the obedience constraints are set up as before. However, types now face asymmetric payoffs that are not only due to their differing valuations:
For type $\overline{\theta}$:

- When the action recommendation is C

$$\begin{aligned}\Psi\left(y_s(\overline{\theta} - c + z) + (\alpha_s - y_s)(\overline{\theta} - c - x)\right) + \\ + (1 - \Psi)(1 - q)\left(y_a(\overline{\theta} - c - y) + (\alpha_1 - y_a)(\overline{\theta} - c - x)\right) \geq \Psi(y_s\overline{\theta})\end{aligned} \quad (22)$$

- When the action recommendation is NC

$$\begin{aligned}&\Psi\left((\alpha_s - y_s)(\overline{\theta} - c + z) + (1 + y_s - 2\alpha_s)(\overline{\theta} - c - x)\right) + \\ &+ (1 - \Psi)(1 - q)\left((\alpha_2 - y_a)(\overline{\theta} - c - y) + (1 + y_a - \alpha_1 - \alpha_2)(\overline{\theta} - c - x)\right) \\ &\leq \Psi((\alpha_s - y_s)\overline{\theta})\end{aligned} \quad (23)$$

The obedience constraints for type $\underline{\theta}$ are:

- When the action recommendation is C

$$\Psi\left(\gamma_s(\underline{\theta} - c + z) + (\alpha_s - \gamma_s)(-c)\right) + \\ + (1 - \Psi)(1 - q)\left(\gamma_a(\underline{\theta} - c - y) + (\alpha_2 - \gamma_a)(-c)\right) \geq (1 - \Psi)(1 - q)(\gamma_a\underline{\theta}) \quad (24)$$

- When the action recommendation is NC

$$\Psi\left((\alpha_s - \gamma_s)(\underline{\theta} - c + z) + (1 + \gamma_s - 2\alpha_s)(-c)\right) + \\ + (1 - \Psi)(1 - q)\left((\alpha_1 - \gamma_a)(\underline{\theta} - c - y) + (1 + \gamma_a - \alpha_1 - \alpha_2)(-c)\right) \\ \leq (1 - \Psi)(1 - q)((\alpha_1 - \gamma_a)\underline{\theta}) \quad (25)$$

These constraints give us the set of BCE.

3.3.2 Analysis

The analysis of the optimal decision rule is along the lines of the previous games. Again, the optimal policy in the social state is always recommending to contribute to both players because designer and player principles are aligned in this case (see Proposition 1).
The designer that maximizes the number of contributions will again find that the contribution constraint for the low type is binding if q is sufficiently imprecise. The threshold value for q to be able to attain the first-best outcome in this case is $1 - \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(-c-y)}$. If $q < 1 - \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(-c-y)}$, the designer solves for the second-best outcome by solving the low type's contribution constraint as an equality. This yields:

$$\alpha_2 = \gamma_a = \frac{-\Psi(\underline{\theta} - c + z)}{(1 - \Psi)(1 - q)(-c - y)} \quad (26)$$

Once again, $\alpha_1 = 1$ is optimal and satisfies the constraints.
If the designer cared about social efficiency, she could simply set $\gamma_s = \alpha_s = 1$, $\gamma_a = \alpha_2 = 0$ and $\alpha_1 = 1$ because the players prefer the same outcomes as the designer so that the obedience constraints are easily satisfied.

3.4 Comparison of the designer's payoffs

We now proceed to compare the designer's expected payoffs under the VCM, the mechanism with an entry fee, and in the case where only the high valuation player can produce the good. Recall that the utility function is as follows when the designer gets utility from the number of contributions:

$$E[V] = 2\Psi\alpha_s + (1-\Psi)(1-q)(\alpha_1+\alpha_2)$$

Since $\alpha_s = \alpha_1 = 1$ in all cases, we need to compare the contribution probabilities of the low type in the asocial state, α_2.

Proposition 4. *The expected payoff for the designer maximizing contributions is highest in a mechanism with an entry fee when a first-best solution is not feasible. This is followed by the mechanism where only the high type can produce the good on his own, while the expected utility from the VCM is lowest.*

Proof. We first compare the VCM and the entry fee mechanism:

$$\begin{aligned}
\alpha_{2,VCM} &\ ?\ \alpha_{2,EF} \\
\frac{-\Psi(-c+z)}{(1-\Psi)(1-q)(-c-y)} &\ ?\ \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(1-q)(\underline{\theta}-c-y)} \\
\frac{-c+z}{-c-y} &\ ?\ \frac{\underline{\theta}-c+z}{\underline{\theta}-c-y} \\
(-c+z)(\underline{\theta}-c-y) &\ ?\ (-c-y)(\underline{\theta}-c+z) \\
\underline{\theta}z &\ ?\ -\underline{\theta}y \\
z &> -y
\end{aligned}$$

Since both z and y are positive numbers, we conclude that $\alpha_{2,VCM} > \alpha_{2,EF}$ and therefore that the utility from the entry fee mechanism is larger than from the VCM.[12] Now compare the entry fee mechanism to the situation where only high valuation players can produce the good on their own:

[12]For $\underline{\theta} > 0$. If $\underline{\theta} = 0$, then the mechanisms are equivalent.

$$\alpha_{2,HT} \ ? \ \alpha_{2,EF}$$

$$\frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(1-q)(-c-y)} \ ? \ \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(1-q)(\underline{\theta}-c-y)}$$

$$-c-y < \underline{\theta}-c-y$$

We can see that as long as $\underline{\theta} > 0$, the contribution probability is higher in the entry fee mechanism.

□

Furthermore, note that the threshold values for the signal precision q is closely related to a_2. The higher a_2, the lower the threshold will be. With a lower precision threshold, for a given q it will be more likely that the designer can implement the first-best decision rule and therefore obtain a higher payoff.

Proposition 5. *The likelihood that the first-best decision rule (from the perspective of a contribution maximizing designer) can be implemented is largest in the entry fee mechanism, followed by the mechanism where only the high type produces the good on his own and lastly the VCM.*

Thus, not only does the VCM yield the lowest expected payoff under the second-best decision rule, it is also most likely to make the second-best decision rule necessary in the first place since it has a higher threshold value than the other mechanisms examined here.

Proof. The threshold values are:

$$\begin{aligned}
q_{VCM} &\geq 1 - \frac{-\Psi(-c+z)}{(1-\Psi)(-c-y)} \\
q_{EF} &\geq 1 - \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(\underline{\theta}-c-y)} \\
q_{HT} &\geq 1 - \frac{-\Psi(\underline{\theta}-c+z)}{(1-\Psi)(-c-y)}
\end{aligned}$$

We observe that the second term on the RHS side corresponds to the contribution probabilities of the low type, α_2. Drawing on the result from the previous proof, we can state: $q_{VCM} > q_{HT} > q_{EF}$ □

This section has shown that from an information designer's perspective, the VCM makes it the most difficult to achieve the first best solution compared to the two other mechanisms. This is the case because the VCM makes it easiest to free-ride: The expected payoff from not contributing is larger, which constrains the LHS of the obedience constraints more. In the entry fee mechanism the RHS is always 0, and in the asymmetric mechanism, only games against a high type make free-riding possible.

Thus, a designer that would like to manipulate not only the information structure, but also the game, should consider the entry-fee mechanism when the good is excludable. Alternatively, she could also "raise" the bar for contributions and only provide the good when either high valuation types provide it or enough players of the low valuation type. Transferring this back to our open source example would mean that programmers with lower valuation (and potentially lower ability) can not provide code that is accepted on their own.

However, if the mechanism cannot or is not supposed to be changed, we show with an example in the next section that the designer can still improve the outcome only using action recommendations.

3.5 Example

We now proceed to calculate the set of BCE with specific parameter values and depict them as figures. We choose the parameter values as in Table 10. Note that the precision of the signal q is flexible for now.

Parameter	Value
$\underline{\theta}$	1
$\overline{\theta}$	2
c	1
z	2
y	2
x	0.5
Ψ	$\frac{1}{3}$

Table 10: Parameter values for example

$\underline{\theta},\underline{\theta}$	C	NC
C	2, 2	−0.5,1
NC	1, −0.5	0,0

$\underline{\theta},\overline{\theta}$	C	NC
C	−2, −1	−0.5,2
NC	1,0.5	0,0

$\overline{\theta},\underline{\theta}$	C	NC
C	−1, −2	0.5,1
NC	2, −0.5	0,0

$\overline{\theta},\overline{\theta}$	C	NC
C	3, 3	0.5,2
NC	2,0.5	0,0

Table 11: Example: Payoffs in VCM

VCM

We start with the VCM game described under Section 4.1, which now has the payoffs depicted in Table 11. The first step is to calculate the Bayesian Nash Equilibrium (BNE) that would arise without any correlation device, i.e. without the designer. This is done to be able to calculate the benefit that arises from being able to employ information design compared to a case where this is not done.

Firstly, note that in the BNE, there are only two contribution probabilities: $\overline{\alpha}$ and $\underline{\alpha}$. This is because the players cannot distinguish between the social and the asocial state. Furthermore, the probability for joint contribution (γ) is the product of the two players' contribution probabilities since they contribute independently. We can reuse the obedience constraints to find the equilibrium by relabeling the probabilities (and plugging in the parameters). For the low type:

$$\frac{1}{3}(2\underline{\alpha}\underline{\alpha} - \frac{1}{2}(\underline{\alpha} - \underline{\alpha}\underline{\alpha})) + \frac{2}{3}(1-q)(-2\underline{\alpha}\overline{\alpha} - \frac{1}{2}(\underline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq \frac{1}{3}1\underline{\alpha}\underline{\alpha} + \frac{2}{3}(1-q)1\underline{\alpha}\overline{\alpha} \tag{27}$$

It is important to note that the probabilities depend on the state of the types, i.e. whether the low type is playing against a high type ($\overline{\alpha}$) or a low type ($\underline{\alpha}$). The constraint for the high type is now:

$$\frac{1}{3}(3\overline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \overline{\alpha}\overline{\alpha})) + \frac{2}{3}(1-q)(-1\underline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq \frac{1}{3}2\overline{\alpha}\overline{\alpha} + \frac{2}{3}(1-q)2\underline{\alpha}\overline{\alpha} \tag{28}$$

There is a pure strategy equilibrium in which the players always contribute, $\underline{\alpha} = \overline{\alpha} = 1$, when q is above a threshold value:
Low type:

$$\frac{1}{3}(2) + \frac{2}{3}(1-q)(-2) \geq \frac{1}{3}(1) + \frac{2}{3}(1-q)(1)$$
$$q \geq \frac{5}{6} \tag{29}$$

For the high type, this translates into the same cut-off value:

$$\frac{1}{3}(3) + \frac{2}{3}(1-q)(-1) \geq \frac{1}{3}(2) + \frac{2}{3}(1-q)(2)$$
$$q \geq \frac{5}{6} \tag{30}$$

Both types have the same threshold because the individual valuations cancel out from the equation since they are on the right and the left hand side. Thus, if $q \geq \frac{5}{6}$, there is a pure strategy BNE where all types always contribute.
Furthermore, there is an equilibrium with $\overline{\alpha} = 1$ and $\underline{\alpha} = 0$ for all values of q and another one with $\overline{\alpha} = 0$ and $\underline{\alpha} = 1$ when $q < \frac{11}{12}$. For the values $\frac{11}{14} < q < \frac{5}{6}$, there is also a mixed strategy BNE.[13]
There can be no equilibrium $\underline{\alpha} = \overline{\alpha} = 0$: If the low type sets $\underline{\alpha} = 0$, then the high type would always want to contribute, since he then can never obtain a negative payoff, even in the asocial state. This is the case even for the low type, who would still get a positive expected payoff from always contributing against a high type who never contributes. In other words, the players no contribution constraints would not be satisfied.

Going back to the information design problem, we find the following:
The threshold for being able to implement the first-best solution is $q = \frac{5}{6}$, which is the same cut-off value as in the BNE. The second-best solution is given by:

$$\alpha_2 = \gamma_a = \frac{1}{6(1-q)} \tag{31}$$

Plugging in $q = \frac{1}{3}$ gives $\alpha_2 = \gamma_a = \frac{1}{4}$, which is summarized in Table 12. For the designer, this implies the payoffs:

[13] We discuss the BNE for all mechanisms in more detail in the Appendix.

- Under the BNE:

$$
\begin{aligned}
E[V] &= \Psi\left(\frac{1}{2}(2\cdot 0)+\frac{1}{2}(2\cdot 1)\right)+\\
&+(1-\Psi)\left(\frac{1}{2}\left(q\cdot 1+(1-q)(1+0)\right)+\frac{1}{2}\left(q\cdot 1+(1-q)(1+0)\right)\right)\\
&= \frac{1}{3}+\frac{2}{3}=1 \qquad (32)
\end{aligned}
$$

- With information design:

$$
E[V] = 2\cdot\frac{1}{3}\cdot 1+(1-\frac{1}{3})(q+(1-q)(1+\frac{1}{4})) = \frac{3}{2}-\frac{1}{6}q \qquad (33)
$$

  This gives $E[V]=\frac{13}{9}$ for $q=\frac{1}{3}$, which is larger than the expected payoff in the BNE.

Thus, the information design has clearly improved the designer's payoff.

The set of BCE is depicted in Figures 2 to 4. We hold $\alpha_s=\gamma_s$ constant at 1 and draw the points that satisfy the obedience constraints. α_1 is displayed on the x-axis, α_2 on the y-axis and γ_a on the z-axis. Figure 2 gives the set of BCE for $q=\frac{1}{2}$ and marks the first-best action profile as well as the preferred second-best BCE for a contribution maximizing designer (which is at $\alpha_2=\gamma_a=\frac{1}{4}$). Figure 3 compares the set of BCE for different values of q, where it becomes clear that the highest attainable contribution probabilities increase with q, thus also increasing the designer's utility. That is, the more accurate the signal, the higher the designer's utility. At the threshold value $q=\frac{5}{6}$, the first-best solution can be achieved, as calculated earlier. In Figure 4, we visualize the possible combinations of

$\sigma(\underline{\theta},\underline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

$\sigma(\underline{\theta},\overline{\theta},t_s)$	C	NC
C	$\frac{1}{4}$	0
NC	$\frac{3}{4}$	0

$\sigma(\overline{\theta},\underline{\theta},t_s)$	C	NC
C	$\frac{1}{4}$	$\frac{3}{4}$
NC	0	0

$\sigma(\overline{\theta},\overline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

Table 12: Example: Best attainable decision rule with information design in VCM ($q=\frac{1}{3}$)

α_1 and α_2 for $\gamma_a = 0$ and the different values of q from Figure 3. (still holding the social contribution probabilities constant).

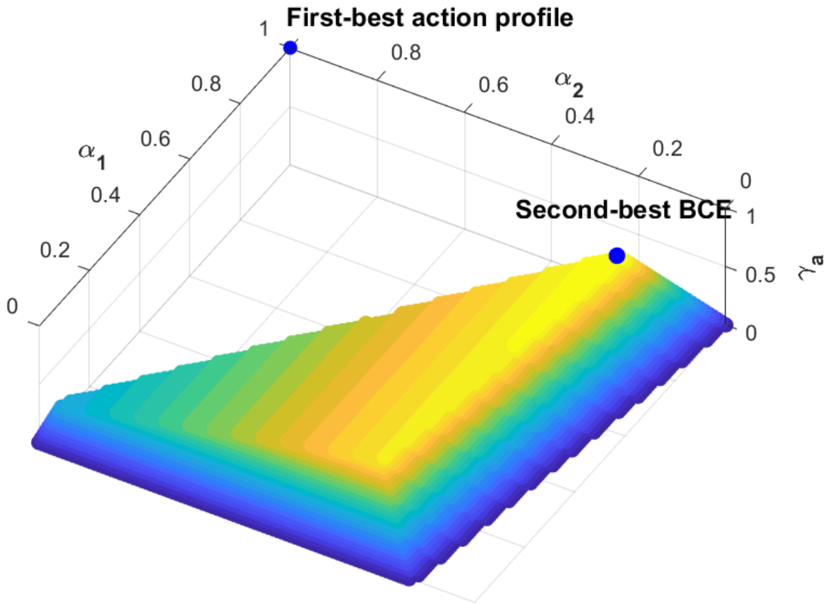

Figure 2: Set of BCE in VCM; $q = \frac{1}{3}$

Excludable good

We now apply the same principle to the mechanism with an entry fee.

$\underline{\theta}, \underline{\theta}$	C	NC
C	2, 2	−0.5,0
NC	0, −0.5	0,0

$\underline{\theta}, \overline{\theta}$	C	NC
C	−2, −1	−0.5,0
NC	0,0.5	0,0

$\overline{\theta}, \underline{\theta}$	C	NC
C	−1, −2	0.5,0
NC	0, −0.5	0,0

$\overline{\theta}, \overline{\theta}$	C	NC
C	3, 3	0.5,0
NC	0,0.5	0,0

Table 13: Example: Payoffs in entry fee mechanism

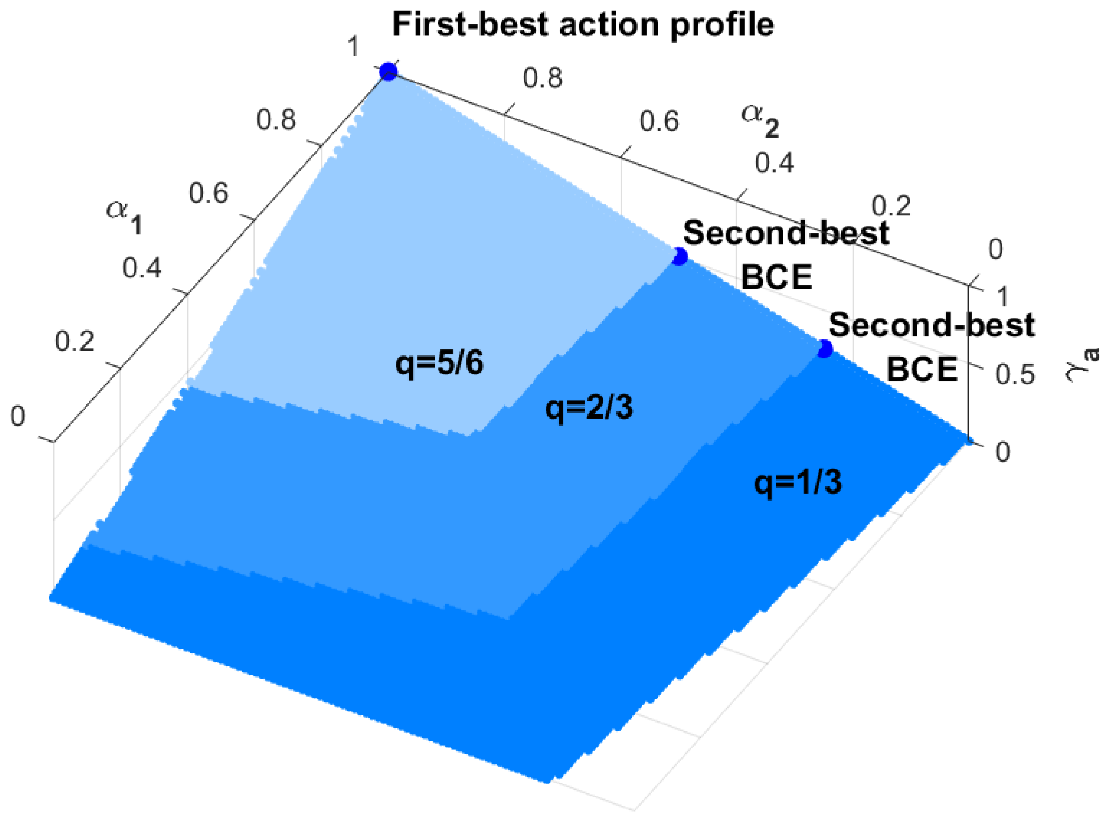

Figure 3: BCE in VCM with varying values of q

The modified obedience constraints to find the BNE are now:

Low type:

$$\frac{1}{3}(2\underline{\alpha}\underline{\alpha} - \frac{1}{2}(\underline{\alpha} - \underline{\alpha}\underline{\alpha})) + \frac{2}{3}(1-q)(-2\underline{\alpha}\overline{\alpha} - \frac{1}{2}(\underline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq 0 \tag{34}$$

High type:

$$\frac{1}{3}(3\overline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \overline{\alpha}\overline{\alpha})) + \frac{2}{3}(1-q)(-1\underline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq 0 \tag{35}$$

The threshold for both types to always contribute is $q = \frac{1}{2}$ in this case. There is also an equilibrium with $\underline{\alpha} = 1$ and $\overline{\alpha} = 0$ for $q < \frac{3}{4}$ and an equilibrium with $\overline{\alpha} = 1$ $\underline{\alpha} = 0$ for all values of q. Note that the threshold value for always contributing is much lower than in the VCM, since there is no free-riding benefit to be obtained. Now we compare this

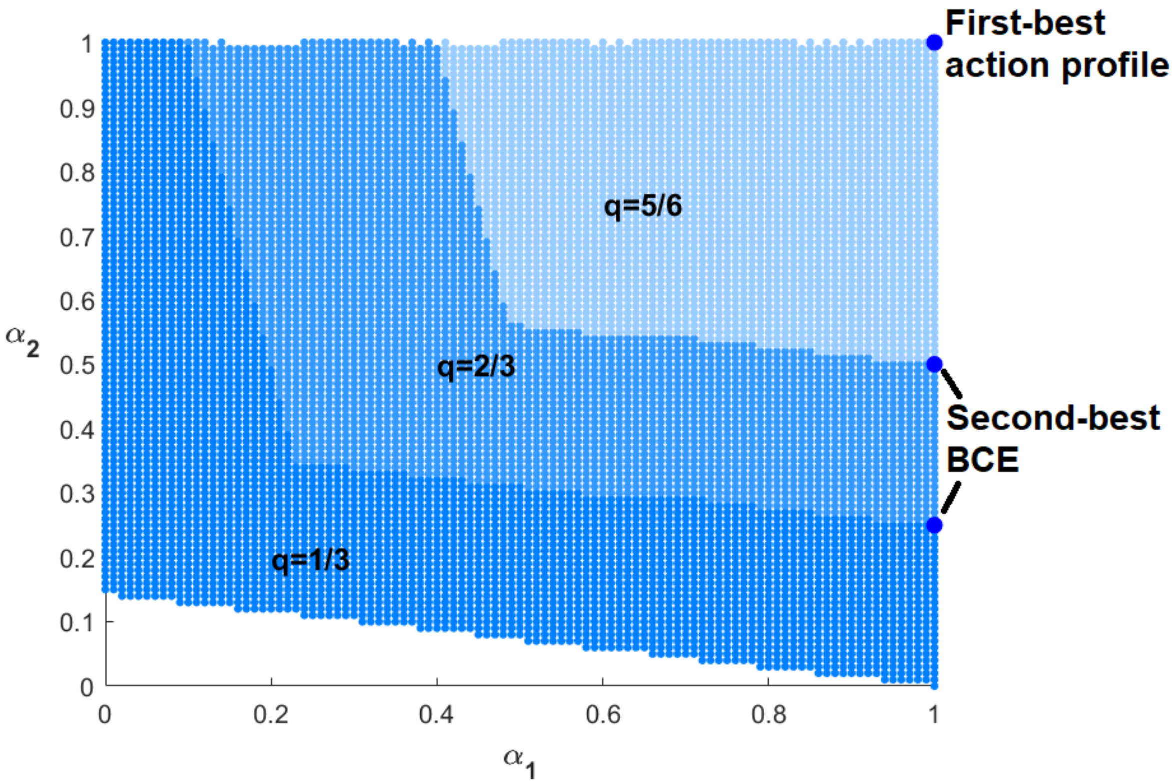

Figure 4: BCE in VCM with varying value of q; 2-D

result to the optimal BCE:

$$\alpha_2 = y_a = \frac{-\Psi(\underline{\theta} - c + z)}{(1-\Psi)(1-q)(\underline{\theta} - c - y)} = \frac{-2 \cdot \frac{1}{3}}{-2 \cdot \frac{2}{3}(1-q)} = \frac{1}{2(1-q)} \tag{36}$$

This implies $\alpha_2 = y_a = \frac{3}{4}$ for $q = \frac{1}{3}$.

We now proceed to analyze the set of BCE under different precision values q. Firstly, Figure 5 shows the set of BCE with $q = \frac{1}{3}$. It also displays the first-best solution, which

$\sigma(\underline{\theta},\underline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

$\sigma(\underline{\theta},\overline{\theta},t_s)$	C	NC
C	$\frac{3}{4}$	0
NC	$\frac{1}{4}$	0

$\sigma(\overline{\theta},\underline{\theta},t_s)$	C	NC
C	$\frac{3}{4}$	$\frac{1}{4}$
NC	0	0

$\sigma(\overline{\theta},\overline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

Table 14: Example: Best attainable decision rule with information design in entry fee mechanism ($q = \frac{1}{3}$)

is not attainable in this case, and the second-best BCE. The latter is at $\alpha_2 = \gamma_a = \frac{3}{4}$ in this case. Comparing this set of BCE to Figure 2, we observe that the best attainable contribution probabilities are much larger than in the VCM at the same precision q. This is because without the possibility to free-ride, contribution is relatively more attractive and therefore the designer is less constrained in her choice.

In Figures 6 and 7, we compare the set of BCE for different values of q. We can again see that with increased signal precision, the best attainable BCE gets closer to the optimal solution. As observed earlier, this first-best solution is attainable for $q \geq \frac{1}{2}$, which is a much lower threshold than under the VCM.

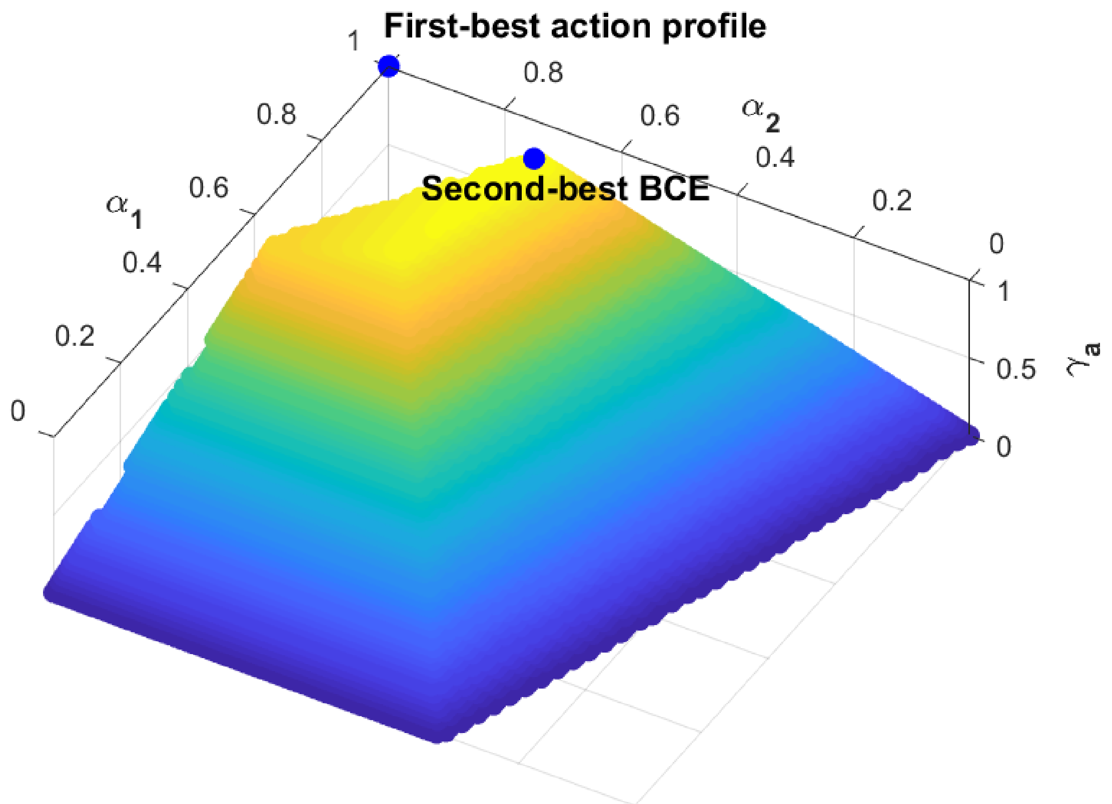

Figure 5: Set of BCE in entry fee mechanism; $q = \frac{1}{3}$

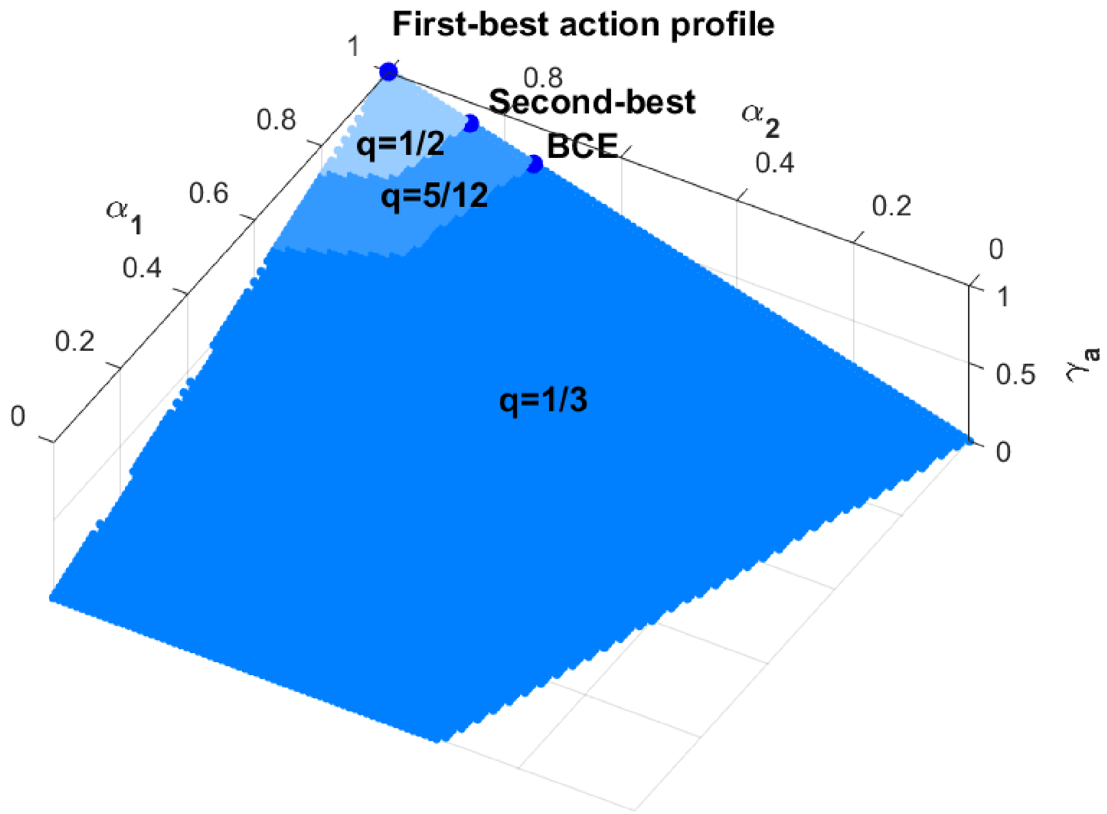

Figure 6: BCE in entry fee mechanism with varying values of q

Only the high type can produce the good on his own

Lastly, we solve for the equilibria in the game where only the high type can produce the good on his own with the numbers from Table 10. Once again, we start out by finding

$\underline{\theta},\underline{\theta}$	C	NC
C	2, 2	−1,0
NC	0,1	0,0

$\underline{\theta},\overline{\theta}$	C	NC
C	−2, −1	−1,0
NC	1,0.5	0,0

$\overline{\theta},\underline{\theta}$	C	NC
C	−1, −2	0.5,1
NC	0, −1	0,0

$\overline{\theta},\overline{\theta}$	C	NC
C	3, 3	0.5,2
NC	2,0.5	0,0

Table 15: Example: Payoffs when only the high type produces good on his own

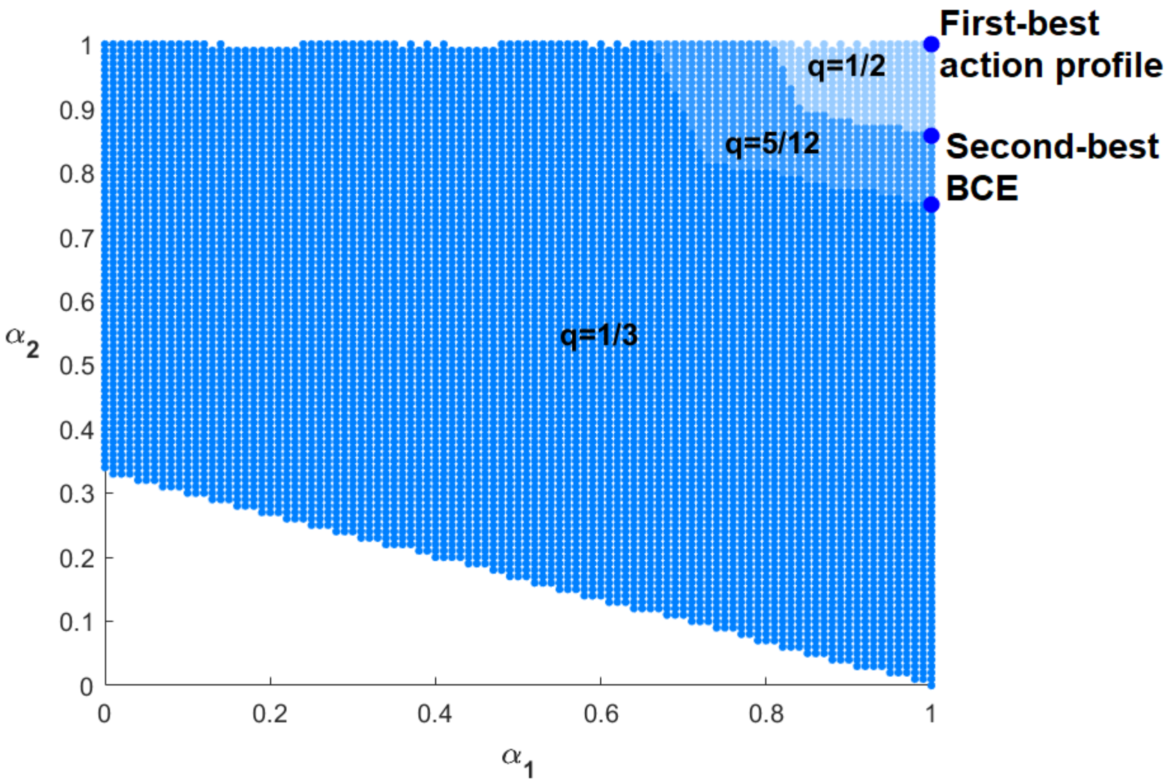

Figure 7: BCE in entry fee mechanism with varying values of q; 2-D

the standard BNE. The modified obedience constraints are now:

$$\frac{1}{3}(2\underline{\alpha}\underline{\alpha} - 1(\underline{\alpha} - \underline{\alpha}\underline{\alpha})) + \frac{2}{3}(1-q)(-2\underline{\alpha}\overline{\alpha} - 1(\underline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq \frac{2}{3}(1-q)1\underline{\alpha}\overline{\alpha} \tag{37}$$

$$\frac{1}{3}(3\overline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \overline{\alpha}\overline{\alpha})) + \frac{2}{3}(1-q)(-1\underline{\alpha}\overline{\alpha} + \frac{1}{2}(\overline{\alpha} - \underline{\alpha}\overline{\alpha})) \geq \frac{1}{3}2\overline{\alpha}\overline{\alpha} \tag{38}$$

With these constraints, the threshold for always contributing is $q > \frac{2}{3}$. As in the first mechanism, there can be a mixed equilibrium when q is in a certain range. Here it is $\frac{1}{2} < q < \frac{15-\sqrt{33}}{16}$. Again, the equilibrium where the high type contributes and the low type does not exist for all values of q. For values of q below $\frac{3}{4}$, there is also the equilibrium where the low type contributes and the high type does not.

In the information design problem, the second-best optimal solution (for $q < \frac{2}{3}$) can be computed as follows:

$$\alpha_2 = y_a = \frac{-\Psi(\underline{\theta} - c + z)}{(1-\Psi)(1-q)(-c-y)} = \frac{-\frac{1}{3}(2)}{(1-\frac{1}{3})(1-q)(-3)} = \frac{1}{3(1-q)} \tag{39}$$

$\sigma(\underline{\theta},\underline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

$\sigma(\underline{\theta},\overline{\theta},t_s)$	C	NC
C	$\frac{1}{2}$	0
NC	$\frac{1}{2}$	0

$\sigma(\overline{\theta},\underline{\theta},t_s)$	C	NC
C	$\frac{1}{2}$	$\frac{1}{2}$
NC	0	0

$\sigma(\overline{\theta},\overline{\theta},t_s)$	C	NC
C	1	0
NC	0	0

Table 16: Example: Best attainable decision rule with information design with asymmetric abilities to produce the good ($q = \frac{1}{3}$)

In Figure 8, we compute the set of BCE for $q = \frac{1}{3}$ with the second-best solution $\alpha_2 = y_a = \frac{1}{2}$. In terms of the designer's utility, this is in between the regular VCM and the entry fee mechanism. Figures 9 and 10 compare the correlated equilibria and their optimal solutions for varying values of q, showing a similar picture to before. In fact in Figure 10, the lower edge of the set for $q = \frac{1}{3}$ is equal to the one under the entry fee mechanism (in Figure 7). This is because when $y_a = 0$, the contribution constraint of the low type in the last game is equivalent to the one in the entry fee mechanism.

This concludes the analysis of the numerical example. Again, we saw that the best attainable outcome gets closer to the first-best solution with an increase in q, i.e. the precision of the signal. Furthermore, the designer can achieve higher utilities for a given value of q in the entry fee mechanism compared to the VCM as her choice of the signal distribution is less constrained without free-riding.

3.6 Discussion

The model in this thesis has shown us that an information designer can improve the outcome of a public goods game by implementing a decision rule that satisfies the participants' incentive compatibility constraints. In the games, we have assumed that players want to contribute to the public good together with their own type, while they get negative utility from working with a different type.

Our results imply that a platform designer would want to make action recommendations to two agents who want to collaborate on a project together. When the designer can observe their types, she can always tell types who want to contribute together to contribute.

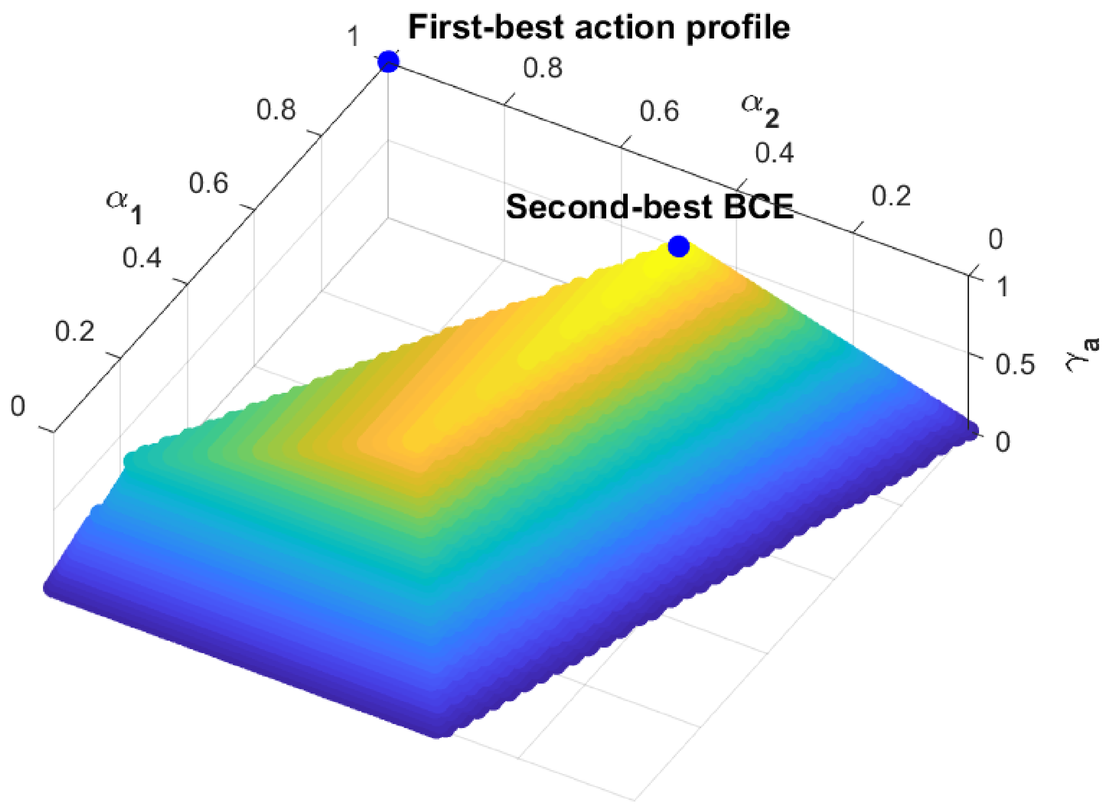

Figure 8: Set of BCE with only $\overline{\theta}$ producing good; $q = \frac{1}{3}$

When they have opposing types, which could mean that they have different goals, then the designer who would like to maximize contributions still gives both players an action recommendation with some probability and recommends only the high type to contribute, who will then work on his own, in all other cases. A designer trying to maximize social efficiency could simply act as a correlation device by always telling only the high type to contribute in the asocial state, which maximizes the sum of utilities in each state. Let us go back to the Linux situation from the introduction, where people work in subteams (Hertel et al., 2003) and are unsure about their team members' goals. A central platform designer could now make action recommendations to the programmers based on their types to maximize her utility. As discussed earlier, this would imply full transparency when players have the same valuation for the team's goals and some potential obfuscation when they do not have the same goals (depending on the designer's objective).

Here, it is an exogenously imposed assumption that the players prefer to work with their own type and derive disutility from working with a different type. One can imagine a

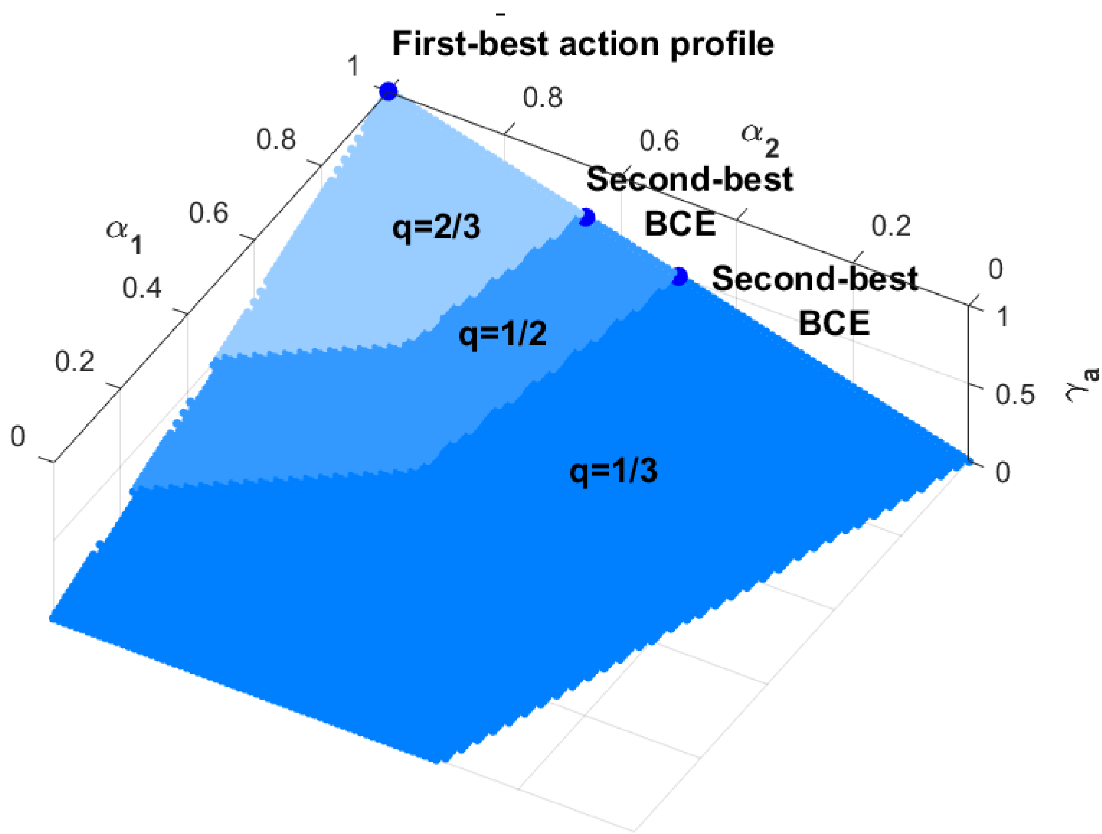

Figure 9: BCE with only $\overline{\theta}$ producing good with varying values of q

situation where conditional cooperators prefer to work with one type rather than another and they are unsure about their counterpart's type. This effect has been supported by previous research as laid out in Chapter 1. However, an interesting extension to this model would be to endogenize this preference for working with one specific type. This may for example arise from a game with multiple states where players adjust their reaction to the outcome from the previous stage.

This dynamic approach would also be insightful from another perspective: Often, public goods are provided over a long period of time, so that collaboration on these projects should also be maintained over more than one period. Extending the analysis to a perspective where an information designer sets up an optimal decision rule for different types and states over multiple periods could therefore give relevant results for the design of public goods provision. This would imply designing the optimal revelation of information over time. Combined with endogenous types, this could also determine which players an agent prefers to work with in the first place.

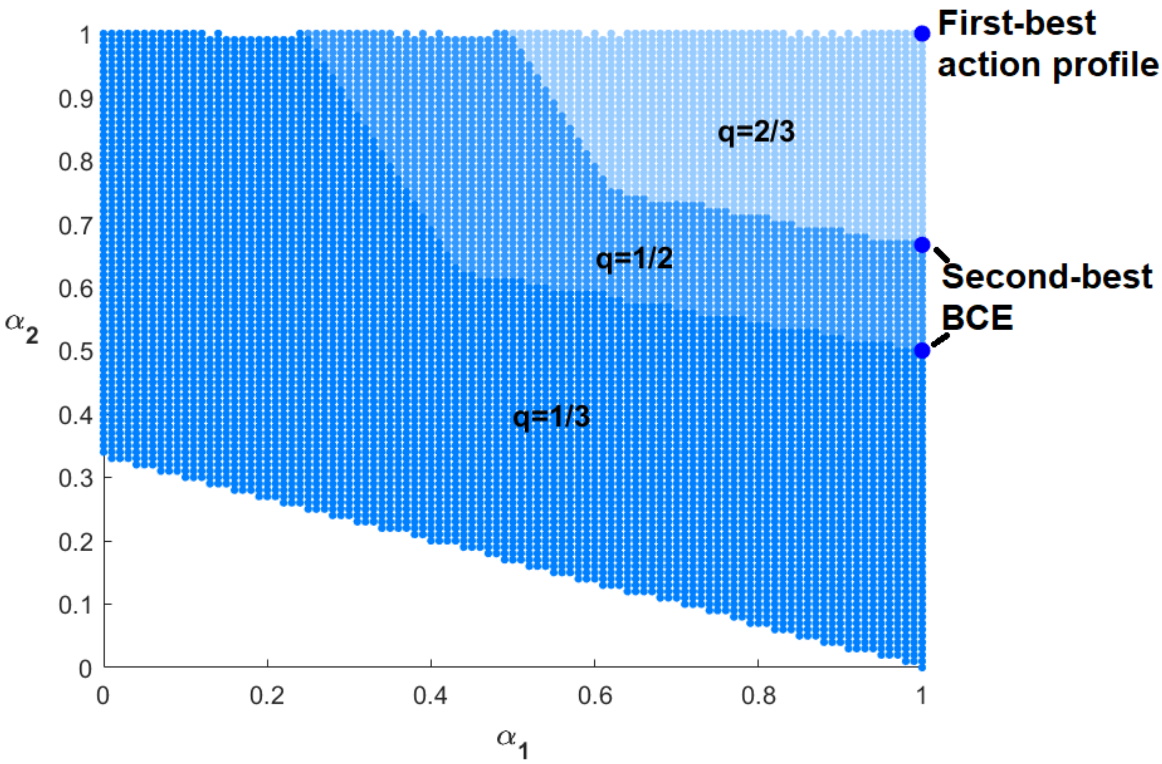

Figure 10: BCE with only $\overline{\theta}$ producing good with varying values of q; 2-D

A potentially problematic point of this analysis is that it is assumed that the designer is omniscient in the sense that she knows the players' types with certainty. Clearly, this may not always be the case in relevant applications, since players' types could be imperfectly observable for the designer. In this case, information design with elicitation would be better suited to analyze the situation: In addition to the incentive compatibility constraints, there would be truth-telling constraints such that the players would truthfully reveal their type. These additional constraints would decrease the set of feasible BCE.

While we have invoked a version of the revelation principle following Bergemann and Morris (2018, 2016a, 2016b), as Mathevet et al. (2017) note, it is not evident that the precondition required to do so is always fulfilled: The designer has to be able to pick the most-preferred equilibrium in order for the revelation principle to apply. This is usually not the case in the many-player case. Our results therefore give an upper benchmark that the designer can achieve in the best case.

As Bergemann and Morris (2018) note, information design, like mechanism design, can have a literal or a metaphorical interpretation. This means that even in cases where there is no actual designer sending signals, the set of BCE "is precisely the set of outcomes that can arise with extra information for a given basic game and prior information structure" (Bergemann and Morris, 2018, Section 5). Thus, the results are still informative about possible outcomes when players receive information in addition to the prior information structure, even if there is no literal information designer. Bergemann and Morris (2016a) treat metaphorical information design in more detail by finding an expansion for the information structure in the form of a larger type space. The expansion lets the players have the same beliefs that they would have, had they observed actual action recommendations. Therefore, this gives rise to a set of BNE that corresponds to the set of BCE under explicit information design.

Overall, the model has a few restrictive assumption that are, however, quite common in the current literature on information design. To be able to make more general statements about information design in public goods provision, it would be beneficial to generalize the relatively specific situation that we have analyzed in this situation. One option would be to vary the players' signal structure which we have assumed to be of a specific form here. The distribution could be varied in many ways, for example to accommodate for asymmetric signals. This would however not change the analytic procedure in principle. Furthermore, we have seen in the introduction, that cooperation can be a complex phenomenon that is determined by multiple different factors like reciprocity, altruism, fairness considerations, group size, communication between players and dynamic aspects. For future research, it is therefore important to try to identify and incorporate these insights into the information design problem. A deeper analysis would therefore include a more complex utility function. Thus, this model is one starting point that may give recommendations for a certain situation.

4 Conclusion

This thesis has analyzed ways to increase contributions in public goods games compared to a situation without information design. An information designer is able to increase expected utility for herself by influencing the players' beliefs through giving action recommendations. The players may benefit from this as well, since they are better or at least not worse off than without information design, otherwise their obedience constraints would not be satisfied and they would not want to follow the decision rule.

We have established that in states where player and designer preferences are aligned, it is optimal not to obfuscate information. In states where this is not the case, it depends on the precision of the players' prior information what contribution probabilities can be achieved. If q is above a threshold, then the designer can simply enforce the first-best solution to her objective. Otherwise, at least one player's obedience constraint is binding, such that the contribution probability will be different from the first-best solution. In our case, it was the contribution constraint of the low type that was binding since this type derives lower levels of utility from the game overall.

If the designer wanted to change the mechanism as well, she could pick an entry fee mechanism that would allow for higher contribution probabilities because of less free-riding opportunities.

This analysis has shown that the efficiency of public goods provision with social preferences can be improved without any monetary transfers, only by giving action recommendations. For future research, it would be interesting to extend this to a framework that includes other types of social preferences.

References

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Akerlof, G. A. and Kranton, R. E. (2010). Identity economics: How our identities shape our work, wages, and well-being. *Princeton University Press.*

Alonso, R. and Câmara, O. (2016a). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory*, 165:672–706.

Alonso, R. and Câmara, O. (2016b). Persuading voters. *American Economic Review*, 106(11):3590–3605.

Arieli, I. and Babichenko, Y. (2016). Private Bayesian persuasion. *Discussion paper, Technion.*

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1):67–96.

Aumann, R. J., Maschler, M., and Stearns, R. E. (1995). *Repeated games with incomplete information.* MIT press.

Balliet, D., Wu, J., and De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140(6):1556.

Bergemann, D., Brooks, B. A., and Morris, S. (2016). Optimal auction design in a common value model. *Discussion paper, Yale University, University of Chicago and Princeton University.*

Bergemann, D., Heumann, T., and Morris, S. (2015). Information and volatility. *Journal of Economic Theory*, 158:427–465.

Bergemann, D. and Hörner, J. (2018). Should first-price auctions be transparent? *Forthcoming in American Economic Journal: Microeconomics.*

Bergemann, D. and Morris, S. (2016a). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522.

Bergemann, D. and Morris, S. (2016b). Information design, Bayesian persuasion, and Bayes correlated equilibrium. *American Economic Review*, 106(5):586–91.

Bergemann, D. and Morris, S. (2018). Information design: A unified perspective. *Forthcoming in Journal of Economic Literature*.

Bergemann, D. and Pesendorfer, M. (2007). Information structures in optimal auctions. *Journal of Economic Theory*, 137(1):580–609.

Bergemann, D. and Wambach, A. (2015). Sequential information disclosure in auctions. *Journal of Economic Theory*, 159:1074–1095.

Blanes i Vidal, J. and Möller, M. (2016). Project selection and execution in teams. *The RAND Journal of Economics*, 47(1):166–185.

Bowles, S. and Hwang, S.-H. (2008). Social preferences and public economics: Mechanism design when social preferences depend on incentives. *Journal of Public Economics*, 92(8):1811–1820.

Celik, G. and Peters, M. (2016). Reciprocal relationships and mechanism design. *Canadian Journal of Economics*, 49(1):374–411.

Che, Y.-K. and Hörner, J. (2018). Optimal design for social learning. *Forthcoming in Quarterly Journal of Economics*.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on MovieLens. *American Economic Review*, 100(4):1358–1398.

Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1):431–457.

Cowen, T. (1992). *Public goods and market failures: A critical examination*. Transaction Publishers.

Daskalakis, C., Papadimitriou, C., and Tzamos, C. (2016). Does information revelation improve revenue? In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 233–250. ACM.

Diamantaras, D., Cardamone, E. I. I., and Campbell, K. A. C. (2009). *A toolbox for economic design*. Macmillan.

Doval, L. and Ely, J. (2016). Sequential information design. *Working Paper, Northwestern University*.

Eckel, C. C. and Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392.

Ely, J., Frankel, A., and Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1):215–260.

Ely, J. C. (2017). Beeps. *American Economic Review*, 107(1):31–53.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717–1722.

Gailmard, S. and Palfrey, T. R. (2005). An experimental comparison of collective choice procedures for excludable public goods. *Journal of Public Economics*, 89(8):1361–1398.

Gentzkow, M. and Kamenica, E. (2014). Costly persuasion. *American Economic Review*, 104(5):457–462.

Gentzkow, M. and Kamenica, E. (2016). A Rothschild-Stiglitz approach to bayesian persuasion. *American Economic Review*, 106(5):597–601.

Gentzkow, M. and Kamenica, E. (2017a). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior*, 104:411–429.

Gentzkow, M. and Kamenica, E. (2017b). Competition in persuasion. *The Review of Economic Studies*, 84(1):300–322.

Goldstein, I. and Huang, C. (2016). Bayesian persuasion in coordination games. *American Economic Review*, 106(5):592–596.

Groves, T. and Ledyard, J. (1977). Optimal allocation of public goods: A solution to the "free rider" problem. *Econometrica*, 45:783–809.

Guala, F. and Filippin, A. (2016). The effect of group identity on distributive choice: Social preference or heuristic? *The Economic Journal*, 127(602):1047–1068.

Hedlund, J. (2015). Persuasion with communication costs. *Games and Economic Behavior*, 92:28–40.

Hertel, G., Niedner, S., and Herrmann, S. (2003). Motivation of software developers in Open Source projects: An internet-based survey of contributors to the Linux kernel. *Research Policy*, 32(7):1159–1177.

Hörner, J. and Skrzypacz, A. (2016). Learning, experimentation and information design. *Working Paper, Yale University and Stanford University.*

Jehle, G. A. and Reny, P. J. (2001). *Advanced microeconomic theory, 2001.* Addison Wesley Longman.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Keser, C. and Van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, 102(1):23–39.

Kolotilin, A. (2015). Experimental design to persuade. *Games and Economic Behavior*, 90:215–226.

Kolotilin, A. (2016). Optimal information disclosure: A linear programming approach. *Forthcoming in Theoretical Economics.*

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Kremer, I., Mansour, Y., and Perry, M. (2014). Implementing the wisdom of the crowd. *Journal of Political Economy*, 122(5):988–1012.

Li, H. and Shi, X. (2017). Discriminatory information disclosure. *American Economic Review*, 107(11):3363–3385.

Mathevet, L., Perego, J., and Taneva, I. (2017). On information design in games. *Working Paper, New York University.*

McCabe, K. A., Rassenti, S. J., and Smith, V. L. (1996). Game theory and reciprocity in some extensive form experimental games. *Proceedings of the National Academy of Sciences*, 93(23):13421–13428.

Olson, M. (1965). *Logic of collective action: Public goods and the theory of groups.* Harvard University Press.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3):137–158.

Ostrovsky, M. and Schwarz, M. (2010). Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics*, 2(2):34–63.

Reeson, A. F. and Tisdell, J. G. (2008). Institutions, motivations and public goods: An experimental test of motivational crowding. *Journal of Economic Behavior and Organization*, 68(1):273–281.

Sell, J. and Wilson, R. K. (1991). Levels of information and contributions to public goods. *Social Forces*, 70(1):107–124.

Shang, J. and Croson, R. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal*, 119(540):1422–1439.

Taneva, I. (2015). Information design. *Discussion paper, University of Edinburgh.*

Zhang, J. and Zhou, J. (2016). Information disclosure in contests: A Bayesian persuasion approach. *The Economic Journal*, 126(597):2197–2217.

Appendix

Proof that $\overline{\theta}$'s expected contribution constraint holds when $\underline{\theta}$'s contribution constraint is satisfied

We show that, when the low type's contribution constraint (Equation (8)) is satisfied, then the high type's contribution constraint (6) automatically holds. We do this by showing that the expected utility from contributing for the low type is smaller or equal to the expected utility from contributing for the high type.

$$\begin{aligned}\Psi(-c+z) + (1-\Psi)(1-q)\left(y_a(-c-y) + (\alpha_1 - y_a)(v_{\underline{\theta}} - c - x)\right)\\ \leq \Psi(-c+z) + (1-\Psi)(1-q)\left(y_a(-c-y) + (\alpha_2 - y_a)(v_{\overline{\theta}} - c - x)\right)\end{aligned}$$

$$\begin{aligned}(1-\Psi)(1-q)\left(y_a(-c-y) + (\alpha_1 - y_a)(v_{\underline{\theta}} - c - x)\right)\\ \leq (1-\Psi)(1-q)\left(y_a(-c-y) + (\alpha_2 - y_a)(v_{\overline{\theta}} - c - x)\right)\end{aligned}$$

$$y_a(-c-y) + (\alpha_1 - y_a)(v_{\underline{\theta}} - c - x) \leq y_a(-c-y) + (\alpha_2 - y_a)(v_{\overline{\theta}} - c - x)$$

$$\underbrace{(\alpha_1 - y_a)}_{\geq 0}\underbrace{(v_{\underline{\theta}} - c - x)}_{<0} \leq \underbrace{(\alpha_1 - y_a)}_{\geq 0}\underbrace{(v_{\overline{\theta}} - c - x)}_{>0} \tag{40}$$

Thus, the LHS is negative and the RHS is positive, so the inequality is always satisfied. We can therefore state that when the decision rule is such that the low type's expected utility from contributing is positive, the high type will also have a positive utility from contributing.

Remarks on the Bayesian Nash Equilibria

VCM

We have already discussed the threshold for both types always contributing ($q = \frac{5}{6}$) in the main text. The *mixed strategy equilibrium* is found by setting:

$$\frac{1}{3}(2\underline{\alpha} - \frac{1}{2}(1 - \underline{\alpha})) + \frac{2}{3}(1 - q)(-2\overline{\alpha} - \frac{1}{2}(1 - \overline{\alpha})) = \frac{1}{3}\underline{\alpha} + \frac{2}{3}(1 - q)\overline{\alpha}$$

$$\frac{1}{3}(3\overline{\alpha} + \frac{1}{2}(1 - \overline{\alpha})) + \frac{2}{3}(1 - q)(-1\underline{\alpha} + \frac{1}{2}(1 - \underline{\alpha})) = \frac{2}{3}\overline{\alpha} + \frac{4}{3}(1 - q)\underline{\alpha}$$

That is, the expected utility from contributing (LHS) must be equal to the expected utility from not contributing (RHS). Solving this system yields a solution for $\frac{11}{14} < q < \frac{5}{6}$:

$$\underline{\alpha} = \frac{(-3 + 2q)(-9 + 10q)}{137 + 140(-2 + q)q}$$

$$\underline{\alpha} = \frac{-33 + 4(16 - 7q)q}{137 + 140(-2 + q)q}$$

Pure strategy equilibria can be found as follows:

- Always contribute as a low type, never contribute as a high type: This implies setting $\underline{\alpha} = 1$ and $\overline{\alpha} = 0$ in the constraints. The expected payoffs must be such that the low type gets a higher utility from contributing while the high type gets higher utility from not contributing for this situation to be an equilibrium.

$$\frac{1}{3}(2) + \frac{2}{3}(1 - q)(-\frac{1}{2}) \geq \frac{1}{3}$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1 - q)(-1) \leq \frac{4}{3}(1 - q)$$

  Both of these constraints are satisfied for $0 < q < \frac{11}{12}$.

- Always contribute as a high type, never contribute as a low type: We reverse the situation compared to the previous point by setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 1$:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1-q)(-2) \leq \frac{2}{3}(1-q)$$

$$\frac{1}{3}(3) + \frac{2}{3}(1-q)(\frac{1}{2}) \geq \frac{4}{3}(1-q)$$

These conditions are satisfied for $0 < q < 1$.

- Both types never contribute: Setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 0$ requires:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1-q)(-\frac{1}{2}) \leq 0$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1-q)(\frac{1}{2}) \leq 0$$

Clearly, the second constraint is not satisfied for any possible value of q.

Entry fee

For the entry fee mechanism, we go through the same steps to find the BNE. The *mixed strategy equilibrium* is found by setting:

$$\frac{1}{3}(2\underline{\alpha} - \frac{1}{2}(1-\underline{\alpha})) + \frac{2}{3}(1-q)(-2\overline{\alpha} - \frac{1}{2}(1-\overline{\alpha})) = 0$$

$$\frac{1}{3}(3\overline{\alpha} + \frac{1}{2}(1-\overline{\alpha})) + \frac{2}{3}(1-q)(-1\underline{\alpha} + \frac{1}{2}(1-\underline{\alpha})) = 0$$

In this case, there is no mixed strategy equilibrium for $0 < q < 1$. Regarding *pure strategy equilibria* we find:

- Always contribute for both types ($\underline{\alpha} = \overline{\alpha} = 1$) is an equilibrium when:

$$\frac{1}{3}(2) + \frac{2}{3}(1-q)(-2) \geq 0$$

$$\frac{1}{3}(3) + \frac{2}{3}(1 - q)(-1) \geq 0$$

Thus, we look for the threshold by setting the contribution probabilities to 1 and solving for q. Both constraints are satisfied for $q > \frac{1}{2}$.

- Always contribute as a low type, never contribute as a high type: This implies setting $\underline{\alpha} = 1$ and $\overline{\alpha} = 0$ in the constraints:

$$\frac{1}{3}(2) + \frac{2}{3}(1 - q)(-\frac{1}{2}) \geq 0$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1 - q)(-1) \leq 0$$

Both of these constraints are satisfied for $0 < q < \frac{3}{4}$.

- Always contribute as a high type, never contribute as a low type: We reverse the situation compared to the previous point by setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 1$:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1 - q)(-2) \leq 0$$

$$\frac{1}{3}(3) + \frac{2}{3}(1 - q)(\frac{1}{2}) \geq 0$$

These conditions are satisfied for $0 < q < 1$.

- Both types never contribute: Again, this is never an equilibrium. Setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 0$ requires:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1 - q)(-\frac{1}{2}) \leq 0$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1 - q)(\frac{1}{2}) \leq 0$$

This never holds for $0 < q < 1$.

Only $\overline{\theta}$ produces the good on his own

We first find the *mixed strategy equilibrium*:

$$\frac{1}{3}(2\underline{\alpha} - (1 - \underline{\alpha})) + \frac{2}{3}(1 - q)(-2\overline{\alpha} - (1 - \overline{\alpha})) = \frac{2}{3}(1 - q)\overline{\alpha}$$

$$\frac{1}{3}(3\overline{\alpha} + \frac{1}{2}(1 - \overline{\alpha})) + \frac{2}{3}(1 - q)(-1\underline{\alpha} + \frac{1}{2}(1 - \underline{\alpha})) = \frac{2}{3}\overline{\alpha}$$

Solving this system yields a solution for $\frac{1}{2} < q < \frac{15-\sqrt{33}}{16}$:

$$\underline{\alpha} = \frac{9 + 2q(-9 + 4q)}{3(7 + 8(-2 + q)q)}$$

$$\underline{\alpha} = \frac{1}{2}(-1 + \frac{1}{7 + 8(-2 + q)q})$$

Pure strategy equilibria:

- Always contribute for both types ($\underline{\alpha} = \overline{\alpha} = 1$) is an equilibrium when:

$$\frac{1}{3}(2) + \frac{2}{3}(1 - q)(-2) \geq \frac{2}{3}(1 - q)$$

$$\frac{1}{3}(3) + \frac{2}{3}(1 - q)(-1) \geq \frac{1}{3}(2)$$

  These constraints both hold for $q > \frac{2}{3}$.

- Always contribute as a low type, never contribute as a high type: Set $\underline{\alpha} = 1$ and $\overline{\alpha} = 0$ in the constraints:

$$\frac{1}{3}(2) + \frac{2}{3}(1 - q)(-\frac{1}{2}) \geq 0$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1 - q)(-1) \leq 0$$

  Note that these are the same constraints as in the entry fee mechanism so they are again satisfied for $0 < q < \frac{3}{4}$.

- Always contribute as a high type, never contribute as a low type: We reverse the situation by setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 1$:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1-q)(-2) \leq \frac{2}{3}(1-q)$$

$$\frac{1}{3}(3) + \frac{2}{3}(1-q)(\frac{1}{2}) \geq \frac{4}{3}(1-q)$$

These conditions are satisfied for $0 < q < 1$.

- Both types never contribute: Setting $\underline{\alpha} = 0$ and $\overline{\alpha} = 0$ requires:

$$\frac{1}{3}(-\frac{1}{2}) + \frac{2}{3}(1-q)(-\frac{1}{2}) \leq 0$$

$$\frac{1}{3}(\frac{1}{2}) + \frac{2}{3}(1-q)(\frac{1}{2}) \leq 0$$

Again, this is never satisfied as the high type would have an incentive to deviate in this case.