Savva, Foivos (2020) *Essays on institutions and implementation.* PhD thesis.

http://theses.gla.ac.uk/81834/

Essays on institutions and implementation

Thesis by Foivos Savva

Submitted in fulfilment of the requirements of the
Degree of Doctor of Philosophy

Adam Smith Business School, College of Social Sciences
University of Glasgow
United Kingdom
20/5/2020

This page is intentionally left blank.

To my partner Natalia, for empowering me to pursue my dreams.

This page is intentionally left blank.

Abstract

This thesis studies three issues in the field of implementation theory. In the first chapter, I examine the implementation of social choice rules under strong Nash equilibrium, when agents do not only care about the final outcomes, but also have a small intrinsic preference for honesty. Specifically, an agent is partially honest if she breaks ties in favour of a truthful strategy, when she faces indifference between outcomes. I present sufficient conditions for implementation in such cases and provide applications in matching and bargaining environments.

In the second chapter, I study the issue of decentralization from the implementation perspective. In most cases of institution design, a social planner is forced to operate in a decentralized manner, by designing distinct institutions that deal with different issues or sectors, over which agents may have complementarities in their preferences. By utilizing the notion of a rights structure, I consider a two-sector environment and examine the possibilities that arise in implementation when the social planner can condition the rights structure of one sector to the one of the other. We distinguish two cases, one when a sector constitutes an institutional constraint (constrained conditional implementation), and one where both sectors can be objects of design (conditional implementation). I characterize the social choice rules that are implementable in the first case, while in the second case I provide sufficient conditions for implementation. My results outline the difficulties of implementation in decentralized environments. As applications, I include some possibility results. First I prove the implementability of a weaker version of the stable rule in a constrained matching environment with partners and projects and second, I prove the implementability of the weak Pareto rule in a multi-issue environment with lexicographic preferences.

In the third chapter, I extend the positive results obtained in Dutta and Sen (2012) to the framework of rights structures. I show that the well-known unanimity condition is sufficient for implementation in such an environment when there is at least one partially honest agent.

Alternative thesis format

This thesis is in alternative format and includes the following papers:

(i) Savva, F., 2018. Strong implementation with partially honest individuals. Journal of Mathematical Economics, 78, pp.27-34.

(ii) Savva, F., 2019. Conditional Rights and Implementation. Available at SSRN: https://ssrn.com/abstract=3466832.

(iii) Savva, F., 2020. A note on partially honest implementation with rights structures.

Contents

List of Tables

List of Figures

Acknowledgements

First of all, I thank my family for their endless supply of love and my friends for their support during this endeavour.

I would also like to express my gratitude to my supervisors Michele Lombardi, Takashi Hayashi and Anna Bogomolnaia, for their invaluable advice and guidance. Additionally, I am especially thankful to Hisayuki Yoshimoto who supported me throughout my studies.

I have been fortunate enough to study in an inspiring and supportive environment at the University of Glasgow, interacting with brilliant people who have directly or indirectly left their mark upon my research. In particular, I would like to thank Hervé Moulin, Yiannis Vailakis, John Levy and Francesca Flamini.

I am also grateful to my friends and colleagues for their support and feedback. This list includes but is not limited to: Adhiraj Singh Rathore, Anastasia Papadopoulou, Arsenii Paramonov, Ashraf Haque Mahfuze, Damiano Turchet, Dimitrios Koutsoupakis, Elizaveta Victorova, Hualin Li, Jerome Santoul, Johanna Tiedemann, Laura Comincini, Li Nie, Moritz Mosenhauer, Nikolas Zivanas, Rohan Chowdhury, Prateek Chandra Bhan and Spyridon Lazarakis.

Last but not least, I would like to acknowledge the funding source for this research: This work was supported by the Economic and Social Research Council [ES/J500136/1].

Author's declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Foivos Savva
17/5/2020

Chapter 1

Introduction

Implementation theory provides a sharp tool that allows us to formally study the design of institutions. Its success mostly lies on the surprising generality of the results obtained, with minimal assumptions or structure on the environment. Its aim is to formally study and understand the relationship between institutions and social goals that is, how different institutional structures affect the attainable outcomes that a society or group can achieve. The theory, in its early stage, can be considered a by-product of, or has at least been inspired by, the socialist debate in the 1920s, when scholars such as von Mises, Lange and Lerner debated over whether a decentralized economic system would outperform a centralized one and vice versa.

Hurwicz, inspired by the debate years later, laid the formal foundations that could allow for a systematic and formal comparison of economic systems with his contributions Hurwicz (1960) and Hurwicz (1972). Since these early contributions, the literature has flourished and has provided a very detailed outline on what a social group can aim to achieve. Hurwicz's contribution boils down to an immensely important aspect of institution design, that of incentive compatibility. Essentially, a well-functioning institution is required to align individual incentives towards the social goal that it is designed to implement.

This specific property has a game-theoretic flavour. Indeed, from the implementation theoretic viewpoint, an institution is a game-theoretic device[1]. This abstraction on one hand constitutes an enormous simplification on the nature of an institution. On the other hand though it allows us to study the arising incentive issues at a very fine level, while maintaining a very general environment[2].

A typical implementation theoretic model consists of the following features:

[1] In most cases, a game form, which consists of a strategy space for each agent and an outcome function mapping outcomes to strategy profiles.

[2] Even though more applied models exist as well.

(i) A fixed set of agents, the participants of the society.

(ii) A fixed set of social outcomes. These represent the outcomes that are relevant for the social decision.

(iii) A set of preference profiles, each one specifying for every agent a preference ordering on the set of outcomes.

(iv) A social choice rule, that specifies for each preference profile, a set of outcomes that are considered socially optimal.

The interpretation of the previous setting is as follows: A hypothetical social planner, or the set of agents at an ex ante stage, desire to implement a social choice rule, which depends on their actual preferences. The problem is that, the social planner does not have that information, or equivalently, this information is not publicly verifiable at a later stage. In addition, the social planner cannot naively rely on the agents to truthfully reveal their preferences, as there might be individual incentives to manipulate the social decision. Therefore, the social planner has to design a game-theoretic device, which represents the institution, such that, in the equilibrium of the game induced by the institution and the actual preference profile, the socially optimal outcomes are realized.

From the previous narrative, two main research paths have been developed that try to answer different questions and with varying degrees of generality. First, fix an environment (agents, possible preference profiles and social choice rule). Now consider a particular game-theoretic equilibrium concept. Are there any game theoretic devices that can implement the desired rule in this solution concept? What are their properties? The other direction which is more general takes a step back and asks the following question: Fix an environment and a solution concept, but not a social choice rule. What can a social planner implement in this environment, with some game-theoretic device? In other words, what is the set of implementable social choice rules?

Given these broad research agendas, a substantial literature has been developed[3] that has significantly enhanced our understanding of institution design. Nevertheless, new problems in implementation theory emerge, specifically on the intersection with emerging fields such as Epistemic Game Theory, Behavioural Economics, Cooperative Game Theory and Coalition Theory or Matching and Market Design among others. It is these new advances that have inspired this thesis.

In chapter 2, I shed some light on the intersection of implementation theory and Behavioural Economics. Specifically, I examine the implementation of social choice

[3] I provide more in-depth literature reviews on specific strands of the literature in the individual chapters.

rules in strong equilibrium, when agents are partially honest, that is, when agents prefer not to lie, when the welfare that they derive from the outcome is not at stake. While this particular motive has been studied in other implementation concepts such as Nash, undominated strategies, Bayesian etc., I outline the robustness of partial honesty to stronger equilibrium concepts. Specifically, I show how the sufficient conditions for strong implementation derived in Korpela (2013) can be substantially weakened when agents adopt this weak honesty motive. I also provide applications of my results in matching and bargaining environments, showing the permissiveness of partial honesty in the implementation framework.

In chapter 3, I take a different approach to implementation. Specifically, I follow the recent contribution of Koray and Yildiz (2018) in studying implementation with rights structures. In this way, contrary to the standard literature where noncooperative game forms are utilized, the social planner allocates rights to agents and coalitions, that allow them to change the status quo state, with the specific details of the strategic interaction being left unspecified. The corresponding implementation concept is cooperative game-theoretic in nature. Utilizing this approach, I study implementation when there are two relevant sectors or social issues, that is, the social choice rule assigns pairs of outcomes for each preference profile, while the environment exhibits institutional constraints. First, I examine constrained conditional implementation, where the institution in one sector is fixed and not an object of design. In this setting, the social planner has to condition the rights structure of the designed sector to the fixed one, in order to leverage incentives and implement the desired social choice rule. I provide a complete characterization of the implementable social choice rules in this setting. Second, I examine conditional implementation, where both sectors are objects of design, but the planner is constrained to operate in a decentralized environment, in that she has to design one institution that is decisive for each sectoral outcome, with some (incomplete) communication taking place between them. In this case I provide sufficient conditions for implementation. Applications of my results include a matching environment with partners and projects and an environment with lexicographic preferences.

In chapter 4, I return back to the issue of motives, but now from the rights structure implementation viewpoint. I formulate partial honesty in a rights structure implementation theoretic framework, and I derive sufficient conditions for implementation, when there exists at least a partially honest agent. The results in this chapter compliment the rights structure approach, by providing an analogue of the theorem by Dutta and Sen (2012) in this framework.

Due to the alternative thesis format, each chapter is written as a research paper or note, with separate literature review, and is self-contained.

Chapter 2

Strong implementation with partially honest individuals

2.1 Introduction

Implementation theory studies the relationship between social goals and institutions[1]. Specifically, it aims to examine the effect of institutional design to the attainment of socially desirable outcomes. For example, suppose that a group of people have agreed on the desirable social outcomes as a function of their preferences. How can they make sure that they can indeed obtain those outcomes, when some or all of them may potentially benefit by misrepresenting their preferences? They thus have to rely on designing an institution (in other words, mechanism or game form) through which they will interact, that will ensure the optimality of the outcomes reached through this interaction. More formally, for any collective choice rule that assigns some socially optimal outcomes as a function of individual preferences, implementation is achieved when, for any profile of preferences, the set of socially optimal outcomes coincides with the set of outcomes attained in the equilibrium of the game induced by the mechanism.

While most of the classic literature on the subject relies on the assumption that agents have a purely *consequentialist* nature, that is, they only care about the final outcomes, the strand of behavioural implementation theory typically assumes that agents may also have *procedural* concerns. One recent subfield in particular, takes into account the fact that agents may have an intrinsic preference for honesty. This weak honesty motive is usually modelled in the following manner: Suppose that an agent is indifferent between two outcomes. Then she will strictly prefer to obtain an outcome with a truthful message rather than with an untruthful

[1]For a comprehensive survey of the main results in the literature of implementation theory see Jackson (2001).

one. This type of rationale is typically referred to as *partial honesty* or *minimal honesty* and can be supported by the experimental findings of Hurkens and Kartik (2009) for example, who show that subjects either are always honest, or tend to lie only when they gain by doing so. Despite being rather weak, partial honesty is shown to bear a significant positive effect for the set of implementable rules and limitations imposed by *Maskin-monotonicity*[2] in paticular. In their seminal paper, Dutta and Sen (2012) show that in the presence of at least one partially honest agent in the society, *Maskin-monotonicity* is no longer a necessary condition for Nash implementation and *No Veto Power* alone becomes sufficient for three or more agents.

Overall, the results on Nash implementation with partial honesty have been positive. An important question that remains unanswered though is whether these possibilities can be extended to other, possibly stronger, equilibrium concepts. For example, in many situations, the social planner cannot exclude the possibility of pre play communication between the agents and thus the mechanism may be vulnerable to group deviations. In such settings the natural solution concept to use is strong Nash equilibrium[3] *à la* Aumann (1960), that is robust to deviations by any possible coalition of agents.

The current paper identifies sufficient conditions for strong implementation when all agents are partially honest. Instead of a full characterization, we chose to follow the work of Korpela (2013) in providing simple sufficient conditions that have a more intuitive appeal and are generally easier to check in applications. First, we identify sufficient conditions for strong implementation when all agents are partially honest and prove their sufficiency. Specifically, we show that if a social choice rule satisfies *Weak Pareto Optimality* (WPO), *Universally Worst Alternative* (UWA) and *Weak Pareto Dominance* (WPD), then it can be implemented in strong equilibrium. In this way we achieve a relaxation in the condition of Korpela (2013), namely the *Axiom of Sufficient Reason* (ASR). Our new condition, WPD roughly requires the following to be true: if an outcome a is optimal at some state, and if there exists another outcome b, such that all agents weakly prefer b to a with at least one agent being indifferent between them, then b must be optimal as well. WPD is implied by ASR, therefore our condition is weaker. Next, we provide two applications of our results, in bargaining and pure matching environments. More specifically, we show that the man-optimal (or woman-optimal) solution in a pure matching environment as well as the Nash bargaining solution in a cake-cutting

[2]Maskin (1999) in his seminal paper identified a condition now known as *Maskin-monotonicity* as necessary and almost sufficient for Nash implementation. It roughly says that if an optimal outcome at some state does not fall in even one person's ranking when switching to another state, then it should still be selected as optimal. A formal definition will be given later.

[3]From now on we will use the terms strong equilibrium and strong Nash equilibrium interchangably. The same applies for the respective implementation concepts.

environment are both strongly implementable, when agents are partially honest.

The remainder of the paper is organized as follows: In section 2.2, we review the relevant literature. In section 2.3, we present the basic implementation setting and formal definitions. In section 2.4, we provide the definitions of our conditions, our main theorem and some additional results. Section 2.5 consists of our two applications. Finally, in section 2.6 we conclude by discussing our results and providing some points for further research. The proof of our main theorem is in the appendix.

2.2 Related Literature

The problem of strong implementation has primarily been studied by Maskin (1979). Moulin and Peleg (1982) study strong implementation with the introduction of effectivity functions, that represent the power structure of the mechanism. They show that there exists a class of strongly implementable social choice rules that satisfy *No Veto Power* (NVP), beyond the Pareto correspondence, contrary to what was initially proposed by Maskin (1979). A complete characterization of strongly implementable social choice rules is due to Dutta and Sen (1991). Suh (1996) generalizes the latter result by allowing the planner to possibly exclude some coalition formation *ex ante*, so in this more general setting not all coalitions are feasible. If the planner though cannot obtain such information, the relevant implementation concept is double implementation in Nash and strong equilibrium. Suh (1997) provides general results in this case as well. While complete characterizations are of high theoretical significance, they can be hard to apply to more specific settings. This motivates the more recent work by Korpela (2013) to identify simple sufficient conditions for strong implementation.

On the issue of partial honesty in implementation, the pioneering work of Dutta and Sen (2012) shows that *No Veto Power* (NVP) alone becomes sufficient for Nash implementation in the presence of at least one partially honest agent[4]. Their results are generalized by Lombardi and Yoshihara (2019), who provide a full characterization of Nash implementable rules in the presence of partial honesty, for both unanimous and non-unanimous social choice rules. In more applied settings, Kartik et al. (2014) focus on environments with economic interest and identify sufficient conditions for implementation in two rounds of iterative deletion of strictly dominated strategies by "simple" mechanisms, without utilizing the usual *canonical mechanisms*[5]. On restricted domains with private goods, Doghmi and Ziad

[4]In contrast with the case of no partial honesty, where NVP along with *Maskin-monotonicity* are sufficient. The well-known result is due to Maskin (1999).

[5]Jackson (1992) criticizes the use of canonical mechanisms in implementation theory as too permissive due to their unbounded strategy spaces. Instead, he derives a necessary condition

(2013) provide more positive results for Nash implementation. In other solution concepts with complete information, Saporiti (2014) shows that with partial honesty strategy-proofness is necessary and sufficient for secure implementation, which essentially requires implementation in dominant strategies and Nash equilibrium. Hagiwara (2019) also shows that NVP is sufficient with at least one, and unanimity is sufficient with at least two partially honest agents for double implementation in Nash and undominated Nash equilibria. Finally, the limitations of partial honesty in Nash implementation are outlined in Lombardi and Yoshihara (2018) who explore under which conditions partially honest Nash implementation is equivalent to Nash implementation, and in Adachi (2017).

Partial honesty can yield positive results in incomplete information environments as well. For example, Matsushima (2008) shows that incentive compatibility is sufficient for implementation in strong iterative dominance and Korpela (2014) proves that incentive compatibility and NVP are sufficient for implementation in Bayes Nash equilibrium. Studies with alternative solution concepts include Ortner (2015), who provides more positive results with partial honesty in fault-tolerant Nash equilibrium[6] and stochastically stable equilibrium.

The issue of implementation with partial honesty nevertheless can be put in the broader context of implementation with motives, where it is typically assumed that agents may also give significance to motives as procedural concerns, apart from the final outcomes. Along this line of research, it is worth mentioning a concept related to partial honesty, namely that of "social responsibility". In Lombardi and Yoshihara (2017), the effect of social responsibility is explored with regards to natural implementation[7]. Hagiwara et al. (2017) utilize a similar concept of social responsibility for strategy space reduction with an outcome mechanism for Nash implementation. In a different environment, Doğan (2017) shows that the unique socially optimal allocation of objects to agents can be Nash implemented, when at least three agents have a social responsibility motive. Some general results on motives as tie-breaking rules with regards to Nash implementation are in Kimya (2017). Other significant contributions to the literature of motives in implementation include Glazer and Rubinstein (1998), Corchón and Herrero (2004) and Bierbrauer and Netzer (2016).

for implementation with bounded mechanisms in undominated strategies. In the same context, Mukherjee et al. (2017) provide a full characterization when all agents are partially honest.

[6]Fault-tolerant Nash equilibrium was first introduced in the implementation literature by Eliaz (2002) as an equilibrium concept which is robust to the bounded rationality of a number of agents.

[7]Specifically, they show that the Walrasian correspondence, although it violates *Maskin-monotonocity*, can be implemented via a market-type mechanism, where agents announce prices and consumption bundles. Like in the case of Kartik et al. (2014), no tail-chasing construction is used.

2.3 Preliminaries

Our society consists of a finite set of individuals $N = \{1, ..., n\}$ with $|N| = n \geq 3$. By $C \subseteq N$ we will denote a coalition of agents. The set of all possible social outcomes is denoted by A and we typically assume that $|A| \geq 2$. Each agent i is endowed with a preference ordering (complete, reflexive and transitive binary relation) over A that is denoted by R_i. We denote the set of all such possible orderings for i by $\mathcal{R}_i$ and, as usual, by P_i and I_i we denote the asymmetric and symmetric part of R_i respectively. Define $\mathcal{R} \equiv \times_{i \in N} \mathcal{R}_i$ with a typical element $R = (R_1, ..., R_n)$ which we call a *preference profile* or simply, *state*. For each $i \in N$ let $L_i(a, R) = \{b \in A | aR_i b\}$ be agent i's *lower contour set* of outcome a in state R. A *social choice rule* (SCR) f is a correspondence $f : \mathcal{R} \rightrightarrows A$ such that for all $R \in \mathcal{R}, \emptyset \neq f(R) \subseteq A$. A *social choice function* (SCF) is a single-valued SCR. For any $R \in \mathcal{R}$, we call $f(R)$ the set of f-optimal outcomes in state R.

A mechanism G is a pair (S, g), which consists of a strategy space $S = \times_{i \in N} S_i$, with S_i being the set of available strategies for each $i \in N$, and an outcome function $g : S \to A$, that maps each strategy profile $s = (s_1, ..., s_n) \in S$ to an outcome in A. As usual, let (s_i', s_{-i}) be the strategy profile where agent i plays the strategy s_i' while all $j \neq i$ play s_j. In a similar manner, let $(s_C', s_{N \setminus \{C\}})$ be the strategy profile where all $i \in C$ play s_i', and all $j \in N \setminus C$ play s_j. We also define the range of a mechanism G as $g(S) = \{a \in A | a = g(s) \text{ for some } s \in S\}$. Now let Γ be the set of all possible mechanisms, and $\Gamma^* = \{G \in \Gamma | g(S) = A\}$, that is, Γ^* is the set of all mechanisms whose range is equal to the set of social outcomes. Any mechanism G with a preference profile R define a normal form game (G, R). We focus on the case of complete information where the state R is common knowledge among the agents, while not to the planner.

In our setting, we assume that agents do not only care about the social outcomes, but also give some importance (although small) to the procedure that leads to those outcomes. More specifically, we assume that agents are *partially honest* in the following sense: If an agent is indifferent between two outcomes and she can attain those outcomes with two different strategies with one being "honest" and the other being "dis-honest", then she strongly prefers to follow the honest strategy. More formally, for honesty to be meaningful in our setting, we should restrict the set of possible mechanisms such that the strategy set of each $i \in N$ is $S_i = \mathcal{R} \times M_i$. That is, each agent is required to announce a preference profile $R \in \mathcal{R}$ and an arbitrary message $m_i \in M_i$. Then, given a mechanism G, for any $i \in N$ we define i's *truthful correspondence* as $T_i^G : \mathcal{R} \rightrightarrows S_i$ such that for each agent i, state R and message m_i, $T_i^G(R) = \{R\} \times M_i$. The truthful correspondence represents the truthful strategies for each agent i in state R, which essentially consist of announcing the "true" state. We now define agent i's *extended preferences* on the strategy space S as follows. Given a vector of truthful correspondences

$T^G = (T_1^G, ..., T_n^G)$, for all $i \in N$ and $R \in \mathcal{R}$, define $\succeq_i^R$ as a complete, transitive and reflexive binary relation on S. An extended preference profile in state R is denoted by $\succeq^R = (\succeq_1^R, ..., \succeq_n^R)$. We are now ready to proceed to the formal definition of partial honesty.

Given a mechanism G, an agent i is *partially honest* if $\forall s_i, s_i' \in S_i, \forall s_{-i} \in S_{-i}$:

- $[s_i \in T_i^G(R), s_i' \notin T_i^G(R)$ and $g(s_i, s_{-i}) R_i g(s_i', s_{-i})] \Rightarrow (s_i, s_{-i}) \succ_i^R (s'_{-i}, s_{-i})$.
- In all other cases, $g(s_i, s_{-i}) R_i g(s_i', s_{-i}) \iff (s_i, s_{-i}) \succeq_i^R (s'_{-i}, s_{-i})$

An agent i is *not partially honest* if $\forall s_i, s_i' \in S_i, \forall s_{-i} \in S_{-i}$:

- $g(s_i, s_{-i}) R_i g(s_i', s_{-i}) \iff (s_i, s_{-i}) \succeq_i^R (s'_{-i}, s_{-i})$

In other words, an agent cares about honesty in a lexicographic manner: First she "consults" her ordering over outcomes, and if she is indifferent between some, she consults her ordering over strategies, strongly preferring the honest strategies if they exist. That is, her partial honesty serves the purpose of a *tie-breaking rule* when she faces indifference. On the other hand, an agent that is not partially honest cares only about the outcomes and does not give any significance to her strategies.

Notice that a mechanism G with an extended preference profile $\succeq^R$ in state R define an (extended) game in normal form $(G, \succeq^R)$. Finally, we assume that in our society there can be partially honest and not partially honest agents and we denote the set of partially honest agents by H. For the planner however, we only assume that he knows the class of all conceivable sets of partially honest agents, $\mathcal{H} \subseteq 2^N \setminus \{\emptyset\}$, without knowing which set is the actual one.

Regarding the solution concept, since we assume that players are allowed to collude, the equilibrium notion that we use is *strong equilibrium*. Formally, $s \in S$ is a strong equilibrium in the game $(S, g, \succeq^R)$, if for all $C \subseteq N$ and $s'_C \in S_C$, there exists an agent $i \in C$ such that $(s_C, s_{N \setminus C}) \succeq_i^R (s'_C, s_{N \setminus C})$. In other words, a strategy profile is a strong equilibrium if there is no coalition that can deviate from it and make all of its members strictly better off. Let the set of strong equilibria of $(S, g, \succeq^R)$ be $SE(G, \succeq^R) = \{s \in S | \text{s is a strong equilibrium in } (G, \succeq^R)\}$. We say that mechanism G implements the SCR f in strong equilibrium, if in any state $R \in \mathcal{R}$, $g(SE(G, \succeq^R)) = f(R)$, that is, if in any state, the set of outcomes obtained through the strong equilibria of the extended game coincides with the set of socially optimal outcomes. The SCR f is strongly implementable if there exists a mechanism that implements it in strong equilibrium.

The previous formal setting can be interpreted as follows. First of all, the SCR represents the collective choice rule that our society utilizes in order to make collective decisions. It can also be interpreted as the constitution of the society designed in an *ex ante* stage. A mechanism on the other hand represents the institution through which the agents in the society interact with each other, that is, it determines the rules and the outcomes of the interaction. A hypothetical benevolent social planner wishes to implement the SCR, however, he does not know the true state, hence, he relies on the agents in order to obtain this information. On the other hand, truthful revelation of the state may not be in the best interests of some agents. Therefore, the goal of the social planner is to construct a mechanism that will lead to the optimal according to the SCR outcome, for any realization of the agents' preferences, that is, for any preference profile. For the strong implementation of the SCR we thus require any optimal outcome to be attainable by some strong equilibrium and any strong equilibrium to lead to an optimal outcome.

2.4 Results

In this section, we present our main results. Before proceeding though, it would be helpful first to review the result of Korpela (2013). This will enable us to outline more clearly the weakening of the sufficient conditions for strong implementation when we adopt the partial honesty assumption. The conditions are the following:

Holocaust Alternative (HA): $\exists a_H \in A$, such that:

- $\forall R \in \mathcal{R}, a_H \notin f(R)$, and,
- $\forall R \in \mathcal{R}, \forall a \in A \setminus \{a_H\}, a \notin L_i(a_H, R)$.

Weak Pareto Optimality (WPO): $\forall R \in \mathcal{R}, f(R) \subseteq wPO(A, R)$, where $wPO(A, R) = \{a \in A | \nexists b \in A \text{ such that } \forall i \in N, bP_i a\}$.

Axiom of Sufficient Reason (ASR): $\forall R, R' \in \mathcal{R}, \forall a \in f(R), \forall b \in A$:

$$\forall i \in N, L_i(a, R) \subseteq L_i(b, R') \Rightarrow b \in f(R').$$

Intuitively, one can imagine **HA** as the worst alternative for all agents in any state, that cannot ever be selected as an optimal outcome. It is a significant restriction on the preference domain, however, it is meaningful in various applications. It essentially allows us to overcome more involved conditions such as Condition γ of Dutta and Sen (1991). **WPO** restricts the range of the SCR to weakly Pareto optimal outcomes. It is well-known from Maskin (1979) that weak

Pareto optimality in the range of the mechanism is also a necessary condition for strong implementation.

ASR can be interpreted as follows: Let an outcome a be selected as f-optimal for some preference profile R. Now imagine an outcome b and profile R' such that for all agents, every outcome that was ranked weakly below a in R is also ranked weakly below b in R'. Then, b should be f-optimal in R'. In other words, if every reason for a to be f-optimal in R is also a reason for b to be f-optimal in R', and a is indeed selected as an optimal outcome in R, then b should be selected as an optimal outcome in R' as well. It is useful to note that **ASR** is stronger than *Maskin-monotonicity* (**MON**) and *Unanimity* (**U**) as it implies both. We review the formal definitions below:

Maskin-Monotonicity (MON): $\forall R, R' \in \mathcal{R}, \forall i \in N, \forall a \in f(R)$:

$$\forall i \in N, L_i(a, R) \subseteq L_i(a, R') \Rightarrow a \in f(R').$$

Unanimity (U): $\forall R \in \mathcal{R}, \forall a \in A$:

$$\forall i \in N, A \subseteq L_i(a, R) \Rightarrow a \in f(R).$$

For example, note that we obtain **MON** if in the definition of **ASR** we set $b = a$. To see that it implies **U**, suppose that **ASR** holds, and for some state R and outcome a we have that for all i, $A \subseteq L_i(a, R)$. Then, for any state R' and any outcome $c \in f(R')$ it trivially holds that for all i, $L_i(c, R') \subseteq A \subseteq L_i(a, R)$, and from **ASR**, $a \in f(R)$ is obtained. We are now ready to present Korpela's theorem:

Theorem 1 (Korpela, 2013)**.** If a SCR f satisfies **HA**, **WPO** and **ASR** then it is strongly implementable.

Theorem 1 makes no assumptions with regards to the partial honesty motive. Its significance lies on the simplicity and intuitive appeal of the conditions. Now proceeding to our results, we will utilize the following assumption that summarizes the knowledge of the social planner with regards to the number of partially honest agents in the society.

Assumption 1: All agents in N are partially honest and the planner knows that.

Assumption 1 has been extensively used in implementation problems. Examples include Kartik and Tercieux (2012), Korpela (2014), Matsushima (2008), Mukherjee et al. (2017), Ortner (2015) and Saporiti (2014). As in the case of the Dutta and Sen (2012) in Nash implementation, our goal is to examine the effect of the presence of partially honest agents on the strong implementation problem.

Moreover, we aim to determine whether partial honesty bears analogous significant impact in the case of strong implementation as in Nash implementation, given that the sufficient conditions for the former are much stronger than in the case of the latter. In fact, by assuming that all agents are partially honest we manage to derive sharp and significant results. For our first result, we identify sufficient conditions for strong implementation when all agents are partially honest. Our key condition is the following[8]:

Weak Pareto Dominance (WPD): $\forall R \in \mathcal{R}, \forall a \in f(R), \forall b \in A$, if:

- $\exists j \in N, aI_jb$, and
- $\forall i \in N \setminus \{j\}, bR_ia$,

then $b \in f(R)$.

The intuition behind our condition is the following: Suppose that a is an f-optimal outcome at state R. Then, if there exists an outcome b such that everyone weakly prefers b to a, with at least one agent being indifferent between them, then b must be selected as f-optimal as well[9]. Another way to look at **WPD** is as an "expansion" of the set of socially optimal outcomes in each state, so as to include all unanimously weakly preferred, or indifferent outcomes. The latter interpretation also has a strong normative appeal. Notice that **WPD** is implied by **ASR**. To see this simply set $R = R'$ in the definition of **ASR** which makes **WPD** true. Another interesting fact with regards to **WPD** is that together with **WPO**, it implies **U**, which will prove to be particularly useful in our main result. This is stated formally in Proposition 1 below.

Proposition 1. If a SCR f satisfies **WPO** and **WPD**, then it satisfies **U**.

Proof. Consider a SCR f that satisfies both **WPO** and **WPD**. Also, consider a state $R \in \mathcal{R}$ and an outcome $a \in A$ such that $\forall i \in N, \forall b \in A, aR_ib$, so that the premises of **U** are satisfied. If $a \in f(R)$, then we are done. Suppose that this is not the case. Then, since $f(R) \neq \emptyset$, there must exist an outcome $c \in A$ such that $c \in f(R)$. Since $\forall i \in N, \forall b \in A, aR_ib$, we must have that aR_ic. Now suppose that $\forall i \in N, aP_ic$. This however cannot be the case as **WPO** is violated. Therefore, there must exist an agent $j \in N$ such that aI_jc. However, for all $i \in N \setminus \{j\}$

[8] We are grateful to an anonymous referee for motivating us to pursue a weakening of the condition that we initially presented in our working paper.

[9] In general, we can exclude the possibility of a being strictly Pareto dominated by b by the **WPO** condition which, apart from using it as part of our sufficient condition, we also show it to be necessary for partially honest strong implementation in the range of the mechanism. See Proposition 2.

it holds that aR_ic. But then, **WPD** implies $a \in f(R)$, a contradiction. This completes the proof.

□

Next, we present the second part of our sufficient condition, a weakening of **HA**, the *Universally Worst Alternative*. It is particularly useful as it is satisfied in various interesting environments as shown in our applications section. We state it formally below:

Universally Worst Alternative (UWA): $\exists a_W \in A$, such that $\forall R \in \mathcal{R}, \forall i \in N, \forall a \in f(R), aP_ia_W$.

So, a **UWA** is strictly worse than any socially optimal outcome for any agent and state and is never selected as socially optimal itself. It is easy to see that it is implied by **HA**, as any **HA** is also a **UWA**[10]. Now, **UWA**, **WPO** and **WPD** become sufficient for strong implementation when all agents are partially honest, which is stated in our main theorem:

Theorem 2. Suppose that Assumption 1 holds. If a SCR f satisfies **UWA**, **WPO**, and **WPD**, then it is strongly implementable.

Proof. See apendix. □

Regarding the proof, we utilize the mechanism of Korpela (2013). Each agent is called to announce an outcome, a state, a positive integer and whether she raises a flag or not. We essentially show that because of Assumption 1, there cannot be any strong equilibria where an agent is announcing a state different from the true one, as in such a case, due to the nature of the outcome function, there would exist profitable deviations motivated by partial honesty. Then, we show that our conditions are sufficient to guarantee that a socially optimal outcome is a strong equilibrium and that any strong equilibrium leads to a socially optimal outcome.

Several points are worth noting in this particular theorem. First, **WPD** constitutes a significant weakening of the **ASR** which reduces to a Pareto related condition. This is quite interesting since we were able to dispose of **MON**, or any variation of it from our sufficient conditions. In fact, we only utilise "intra-state" conditions, that is, conditions that restrict the socially optimal set with regards to the same state, rather than "inter-state" ones. The second point to note is that **WPO** is also a necessary condition for partially honest strong implementation, given that the range of the mechanism coincides with the set of alternatives[11]. We formally prove the statement in Proposition 2 below. Finally, notice that if we

[10]For other uses of **UWA** see Moore and Repullo (1990), or Jackson et al. (1994).

[11]This assumption is crucial for the necessity of **WPO**.

only allow for linear orderings[12], **WPD** holds trivially (Proposition 3) and it becomes redundant as a sufficient condition. Below we provide the formal statements and appropriate proofs and in Corollary 1 we state a characterization theorem of strongly implementable SCRs for the case of linear preferences when agents are partially honest.

Proposition 2. Let Assumption 1 hold and f be strongly implementable by a mechanism $G \in \Gamma^*$. Then f satisfies **WPO**.

Proof. Let the premises hold. To derive a contradiction, suppose that f does not satisfy **WPO**. This implies that for some $R \in \mathcal{R}$, there exists $a \in f(R)$ such that $a \notin wPO(A, R)$. So, there must exist $b \in A$ such that $\forall i \in N, bP_i a$. Now, since f is strongly implementable, there exists a strong equilibrium $s \in S$ such that $g(s) = a$. So, $\forall C \subseteq N, \forall s'_C \in S_C, \exists j \in C, (s_C, s_{N \setminus C}) \succeq_j^R (s'_C, s_{N \setminus C})$. Since $G \in \Gamma^*$, we are allowed to consider $C = N$ and $g(s') = b$. Then, we have that $s \succeq_j^R s'$ and for j it holds that:

- $s \sim_j^R s'$ (1), or
- $s \succ_j^R s'$ (2)

If (1) holds, then $g(s) = aI_j b = g(s')$, but also $bP_j a$, a contradiction. If (2) holds, we have either $g(s) = aP_j b = g(s')$ and $bP_j a$, a contradiction, or $a = g(s) I_j g(s'), s_i \in T_j^G(R)$ and $s'_i \notin T_j^G(R)$ which also contradicts $bP_j a$. So, our initial statement that f does not satisfy **WPO** cannot hold. This completes the proof. □

Proposition 3. If $\mathcal{R}^A = \mathcal{L}$, then any SCR f satisfies **WPD**.

The proof of Proposition 3 is straightforward, as one can notice that if there exists an agent that is indifferent between a socially optimal alternative a and an outcome b, as dictated in the premise of **WPD**, then, by the linear preference assumption, a must be equal to b and the condition holds vacuously. We are now ready to proceed with our Corollary:

Corollary 1. Let $\mathcal{R}^A = \mathcal{L}$ and Assumption 1 hold. If a SCR f satisfies **UWA**, then it is strongly implementable by a mechanism $G \in \Gamma^*$ if and only if it satisfies **WPO**.

Proof. Immediate implication of Theorem 2 and Propositions 2 and 3. □

[12] Formally, let $\mathcal{L}_i$ be the set of all linear, that is, complete, transitive and antisymmetric, orders on A for each agent i and let $\mathcal{L} \equiv \times_{i \in N} \mathcal{L}_i$. Let the space of admissible preferences be $\mathcal{R}^A$. So, in this case we set $\mathcal{R}^A = \mathcal{L}$.

Corollary 1 provides a characterization of the strongly implementable social choice rules with linear preferences, when there exists a **UWA** and all agents are partially honest. Essentially, in this case **WPO** is a necessary and sufficient condition for strong implementation[13].

2.5 Applications

In this section we provide applications of our Theorem 2. Our first application is in pure matching environments, that is, one-to-one matching environments where for every agent, staying unmatched is not feasible, or it is the worst possible alternative in any state. For example, a manager in a firm might want to match people from two groups with different abilities in pairs, in order to undertake projects. In this case it might be reasonable to assume that staying unmatched is not feasible (as it might lead to redundancies). We show that when all agents are partially honest, the man-optimal (or woman-optimal) stable solution is strongly implementable. This is to be compared with the results of Tadenuma and Toda (1998), who show that with more than three agents in each group, while the whole stable solution in pure matching problems is Nash implementable, no single-valued subsolution of it is. Lombardi and Yoshihara (2019) show that partial honesty can resolve this issue for Nash implementation, as the man-optimal (or woman-optimal) solution become Nash implementable in this case. With regards to strong implementation, Shin and Suh (1996) present a mechanism for strong implementation of the stable rule in one-to-one matching problems and the implementability of the stable rule in pure marriage problems is shown in Korpela (2013).

Our second application is in bargaining environments. We show that when all agents are partially honest, the Nash bargaining solution is strongly implementable. In general, it is known that the Nash bargaining solution is not Nash implementable, due to the result by Vartiainen (2007*a*). However, Lombardi and Yoshihara (2019) again show that it can be implemented with partial honesty. Our results extend theirs to the strong equilibrium concept.

2.5.1 Pure Matching Environments

We start by defining the formal pure matching environment. Let M, W be two fixed finite sets, such that $|M| = |W| \geq 2$ and $M \cap W = \emptyset$. For all $i \in M$, P_i is a linear order on $W \cup \{i\}$, and for all $i \in W$, P_i is a linear order on $M \cup \{i\}$. A matching is a function $\mu : M \cup W \to M \cup W$ such that for any $i \in M \cup W$ the following hold:

[13] We thank an anonymous referee for pointing us to the possibility of this characterization theorem.

- $i \in M$ & $\mu(i) \neq i \Rightarrow \mu(i) \in W$,
- $i \in W$ & $\mu(i) \neq i \Rightarrow \mu(i) \in M$, and
- $\mu(\mu(i)) = i$.

Let $\mathcal{M}$ be the set of all matchings. We now extend the relation P_i to $\mathcal{M}$ by defining a new relation R_i as follows:

$$\forall i \in M \cup W, \forall \mu, \mu' \in \mathcal{M}, \mu R_i \mu' \iff \mu(i) P_i \mu'(i) \text{ or } \mu(i) = \mu'(i)$$

Let the set of all preferences over $\mathcal{M}$ of each agent i be $\mathcal{R}_i$. We then define $\mathcal{R} \equiv \times_{i \in M \cup W} \mathcal{R}_i$. As usual, $R \in \mathcal{R}$ denotes a preference profile. Now we make the following assumption, which makes our environment one of pure matching:

Assumption 2: $\forall m \in M, \forall w \in W, \forall \mu \in \mathcal{M}, w P_m m$ & $m P_w w$.

A solution (or SCR) is a correspondence $\varphi : \mathcal{R} \rightrightarrows \mathcal{M}$ such that for all $R \in \mathcal{R}, \varphi(R) \subseteq \mathcal{M}$. A pair $(m, w) \in M \times W$ *blocks* $\mu \in \mathcal{M}$ in $R \in \mathcal{R}$ if $w P_m \mu(m)$ and $m P_w \mu(w)$. A matching $\mu \in \mathcal{M}$ is *stable* in $R \in \mathcal{R}$, if there is no pair $(m, w) \in M \times W$ such that (m, w) blocks μ in R. Let $S(R)$ be the set of all stable matchings in $R \in \mathcal{R}$. The stable matching rule is a rule $f^S : \mathcal{R} \rightrightarrows \mathcal{M}$ such that for every $R \in \mathcal{R}$, $f^S(R) = S(R)$. We say that $\mu^M \in \mathcal{M}$ is the *man-optimal* stable matching in state $R \in \mathcal{R}$ if $\mu^M \in S(R)$ and for every $\mu' \in S(R)$ and $m \in M$, we have that $\mu^M(m) P_m \mu'(m)$, or $\mu^M(m) = \mu'(m)$. The man-optimal stable rule f^M is a function $f^M : \mathcal{R} \to \mathcal{M}$ such that for every $R \in \mathcal{R}$, $f(R) = \mu^M$. In a similar manner, we can define the woman-optimal stable matching and rule. We now proceed by stating our possibility result for the pure matching environment.

Proposition 4. Let Assumptions 1 and 2 hold. Then, the man-optimal stable rule f^M is strongly implementable.

Proof. It suffices to show that f^M satisfies **UWA**, **WPO** and **WPD**,

Claim 1: f^M satisfies **UWA**.

Proof. By the construction of the pure matching environment, we have assumed that staying single is the worst alternative for every $i \in M \cup W$. So, we can set $a_W = \mu_W$, where for all $i \in M \cup W$, $\mu_W(i) = i$. So, our environment satisfies **UWA** [14]. □

Claim 2: f^M satisfies **WPO**.

[14] The pure matching environment actually satisfies the stronger condition **HA** as shown in Korpela (2013).

Proof. Suppose not. Consider $R \in \mathcal{R}$ such that $\mu = f^M(R)$ and suppose there exists $\mu' \in \mathcal{M}$ with $\mu' \neq \mu$ such that $\forall i \in M \cup W, \mu'(i) P_i \mu(i)$. Then, there exists $(m, w) \in M \times W$ such that $\mu'(m) = w \neq \mu(m)$ and $\mu'(w) = m \neq \mu(w)$. Consequently, the pair (m, w) would block matching μ, which contradicts its stability. Therefore, f^M satisfies **WPO**.

□

Claim 3: f^M satisfies **WPD**.

Proof. Consider $R \in \mathcal{R}$ and let $f^M(R) = \mu^M$. Now suppose there exists $\mu \in \mathcal{M}$ such that:

- $\exists j \in N, \mu^M I_j \mu$, and
- $\forall i \in N \setminus \{j\}, \mu R_i \mu^M$

Since the man-optimal stable rule f^M is a function, it suffices to show that $\mu = \mu^M$. Now, without loss of generality let $j = m \in M$. For m it holds that $\mu I_m \mu^M$, which implies $\mu(m) = \mu^M(m)$. Let $\mu(m) = w$. Then necessarily it must be the case that $\mu(w) = \mu^M(w)$ and thus $\mu I_w \mu^M$. Now if for all $i \in M \cup W \setminus \{m, w\}$ it also holds that $\mu(i) = \mu^M(i)$, then $\mu = \mu^M$ and we are done. Suppose that this is not the case. So, there exists $i \in M \cup W \setminus \{m, w\}$ such that $\mu(i) \neq \mu^M(i)$. Again, without loss of generality, assume that $i = m' \in M$. Then, it must be that $\mu(m') P_{m'} \mu^M(m')$. Let $\mu(m') = w'$. Now, for w' it is also true that $m' P_{w'} \mu^M(w')$. However, this contradicts the stability of the man-optimal stable matching μ^M, as the couple (m', w') would block it. Therefore, we conclude that $\mu = \mu^M$ and **WPD** holds.

□

By Claims 1, 2, 3 and Theorem 2, we have that the man-optimal stable solution is strongly implementable. This completes the proof.

□

2.5.2 Bargaining Environments

For the definition of the bargaining environment we chose to follow the work of Vartiainen (2007*a*), to whom we refer for the detailed formulation. Let $N = \{1, 2, ..., n\}$ be the set of players. The set of outcomes is $A = \{(a_1, ..., a_n) \in \mathbb{R}^n_+ | \sum_{i=1}^n a_i \leq 1\}$. Let the set of possible types of each agent $i \in N$ be Θ. For each $\theta_i \in \Theta$, $v_i(\cdot, \theta_i) : [0, 1] \to \mathbb{R}$ is agent i's strictly monotonic and continuous utility function. Let Θ_0 be the normalized set of types for each i such that $\Theta_0 = \{\theta_i \in \Theta | v_i(0, \theta_i) = 0\}$. Let Δ be the set of all probability distributions on A. So, for any outcome $p \in \Delta$ and agent $i \in N$, $v_i(p, \theta) = \int_A v_i(a_i, \theta_i) dp(a)$ is the utility function of i defined on Δ. We also set the disagreement points $\mathbf{d} = \mathbf{0}$. The Nash solution is a SCR $f^N : \Theta_0^n \rightrightarrows \Delta$ such that $\forall \theta \in \Theta_0^n, f^N(\theta) = \text{argmax}_{p\in\Delta}\Pi_{i=1}^n v_i(p, \theta_i)$. Notice that our environment satisfies **UWA**, since we have assumed strictly monotonic utility functions and in any Nash solution all agents get positive amounts of the good. This allows us to set $a_W = \mathbf{d} = \mathbf{0}$.

Proposition 5. Let Assumption 1 hold. Then, the Nash solution f^N is strongly implementable.

Proof. Since the Nash solution satisfies weak Pareto optimality by definition, and our environment satisfies **UWA**, it suffices to show only that f^N satisfies **WPD**.

Claim 4: f^N satisfies **WPD**.

Proof. Consider $\theta \in \Theta_0^n$ such that $p \in f^N(\theta)$. Now, let $q \in \Delta$ be such that $\exists j \in N, v_j(q, \theta_j) = v_j(p, \theta_j)$ and $\forall i \in N \setminus \{j\}, v_i(q, \theta_i) \geq v_i(p, \theta_i)$. If $q = p$, then we are done. Suppose that $q \neq p$. If now for all $i \in N \setminus \{j\}$ it is also the case that $v_i(q, \theta_i) = v_i(p, \theta_i)$, then it must be that $q \in\text{argmax}_{p\in\Delta}\Pi_{i=1}^n v_i(p, \theta_i)$. Assume then that there exists an $i \in N \setminus \{j\}$ such that $v_i(q, \theta_i) > v_i(p, \theta_i)$. But this contradicts that $p \in \text{argmax}_{p\in\Delta}\Pi_{i=1}^n v_i(p, \theta_i)$. So, it is true that f^N satisfies **WPD**. □

By Claim 4, Theorem 2 and the fact that the Nash bargaining solution satisfies **UWA** and **WPO**, we conclude that it is strongly implementable. This completes the proof. □

We have shown that the Nash solution satisfies our sufficient conditions and is thus strongly implementable when all agents are partially honest. For this result we relied on the ordinality of the environment. Note for example that **U** is not satisfied by the egalitarian solution in an environment where interpersonal comparisons are allowed, preferences are not strictly monotone and there is more than one good[15]. This implies that our Theorem 2 cannot be applied in this case.

2.6 Concluding remarks

We have provided a sufficiency theorem for strong implementation when all agents are partially honest. Our goal was to extend the positive results that have been obtained in partially honest Nash implementation to the solution concept of strong equilibrium. Our sufficient conditions are much stronger than in the case of Nash implementation and this is due to the much more demanding solution concept, as well as due to the attempt to provide simple sufficient conditions rather than a complete characterization.

[15] For studies in bargaining theory in this type of environment see Roemer (1988).

As applications of our main theorem, we showed that the man-optimal (or woman-optimal) stable rule in a pure matching environment as well as the Nash solution in a bargaining environment with strictly monotone preferences are both strongly implementable when all agents are partially honest. However, as noted before, both these rules are not strongly implementable when there are no partially honest agents, therefore our results show the expansion of strongly implementable rules when the motive of minimal honesty is assumed.

In our view, the applications of our theorems provide an insight into the possibilities that arise in implementation theory when non-consequentialist motives are taken into account. They also emphasize the importance of procedural concerns in mechanism design and social choice theory. An interesting problem for further research which we aim to tackle, is closing the gap between our necessary and suffficient conditions. In fact, the *Non-emptiness* condition of Dutta and Sen (1991) is necessary in our case as well and we conjecture that it could constitute part of a sufficient condition, given that the mechanism is appropriately modified. In that way, the domain restriction of UWA could be avoided and more clear-cut results could be obtained. Finally, along the same line, it would be intriguing to study under which conditions partially honest strong implementation is equivalent to strong implementation.

Acknowledgements

The current research was conducted as part of my PhD thesis at the University of Glasgow. I want to thank my supervisors, Michele Lombardi, Takashi Hayashi and Anna Bogomolnaia for their invaluable comments and support. I am also grateful to two anonymous referees whose comments greatly improved the results of the paper. Finally, I would like to thank Arunava Sen and the participants of the 14th meeting of the Society for Social Choice and Welfare for their comments and suggestions. This work was supported by the Economic and Social Research Council [ES/J500136/1].

Appendix

Mechanism

For the proof of Theorem 2 we will utilize the following mechanism $G = (S, g)$:

For all $i \in N$, $S_i = A \times \mathcal{R} \times \{NF, F\} \times \mathbb{N}_+$. The outcome function g is defined as follows:

1) If $\forall i \in N, s_i = (a, R, NF, \cdot)$ and $a \in f(R)$, then $g(s) = a$.
2) If $\exists C \subset N, \forall i \in N \setminus C, s_i = (a, R, NF, \cdot)$ with $a \in f(R)$, and $\forall j \in C, s_j = (a^j, R^j, F, n^j)$, then:

- If $k = \min\{\text{argmax}_{j\in C} n^j\}$ and $a^k \in \cup_{j\in C} L_j(a, R)$, then $g(s) = a^k$
- Otherwise, $g(s) = a$

3) If $\forall i \in N, s_i = (a^i, R^i, F, n^i)$, then $k = \min\{\text{argmax}_{j\in N} n^j\}$ and set $g(s) = a^k$.
4) If none of the above apply, set $g(s) = a_W$.

Proof of Theorem 2

We will show that any SCR f that satisfies our premises, namely **UWA**, **WPO** and **WPD** can be implemented by mechanism G and we break the proof into two parts:

Part 1: $\forall R \in \mathcal{R}, f(R) \subseteq SE(R)$

Let the true state be R^*. Consider the strategy profile where $\forall i \in N, s_i = (a, R^*, NF, \cdot)$ and $a \in f(R^*)$. If $j \in N$ deviates to rule 2 she will obtain any $b \in L_j(a, R^*)$. So, $g(S_j, s_{N\setminus\{j\}}) = L_j(a, R^*)$. If any $C \subset N$ deviates to rule 2, the obtained outcome will be in $L_j(a, R^*)$ for at least one $j \in C$. If N deviate to rule 3, there cannot be an improvement for all $i \in N$ since f satisfies **WPO**. Finally, there is no profitable deviation by any coalition to rule 4, since, by definition of the **UWA**, a_W is ranked strictly worse to any socially optimal outcome, by all agents. Therefore, s is a strong equilibrium in R^*.

Part 2: $\forall R \in \mathcal{R}, SE(R) \subseteq f(R)$

Let the true state be R^*. We proceed by first proving three useful claims:

Claim 1*: There is no strong equilibrium under rule 1 where $\forall i \in N, R^i \neq R^*$.

Proof. Suppose there exists a strong equilibrium under rule 1, where $\forall i \in N, s_i = (a, R, NF, \cdot)$ with $a \in f(R)$ and $R \neq R^*$. By rule 1 the outcome is a. Then, $\forall i \in N, s_i \notin T_i^G(R^*)$, so, any $i \in N$ can deviate to $s'_i = (a, R^*, F, n^i) \in T_i^G(R^*)$ inducing rule 2 while announcing the true state and not changing the outcome. Therefore, s cannot be a strong equilibrium. □

Claim 2*: There is no strong equilibrium under rule 2 where $\exists i \in N \setminus C$ such that $R^i \neq R^*$.

Proof. Suppose there exists a strong equilibrium under rule 2 where $\exists i \in N \backslash C, s_i = (a, R, NF, \cdot)$ with $a \in f(R), R \neq R^*$, and $\forall j \in C, s_j = (a^j, R^j, F, n^j)$ and let $g(s) = b$. Then, we have that $s_i \notin T_i^G(R^*)$. We break the proof into two cases:

Case 1: $|N \setminus C| \geq 2$

- If $b = a$: Then, since by definition $a \in L_i(a, R)$ holds, i can play $s_i' = (a, R^*, F, n^i) \in T_i^G(R^*)$ with a sufficiently high integer without changing the outcome and become strictly better off by Rule 2.
- If $b \neq a$: Then, again, since $b \in \cup_{j \in C} L_j(a, R)$ it must hold that $b \in \cup_{j \in C \cup \{i\}} L_i(a, R)$, so agent i can play $s_i' = (b, R^*, F, n^i) \in T_i^G(R^*)$ with a sufficiently high integer without changing the outcome and become strictly better off by Rule 2.

Case 2: $N \setminus C = \{i\}$

In this case i can play $s_i' = (b, R^*, F, n^i) \in T_i^G(R^*)$ with a sufficiently high integer without changing the outcome and become strictly better off by Rule 3.

Therefore, there is no strong equilibrium under rule 2, where for some $i \in N \setminus C, R^i \neq R^*$. □

Claim 3*: There is no strong equilibrium under rule 2 where $\exists i \in C$, with $R^i \neq R^*$.

Proof. Suppose this is not the case, that is, there exists a strong equilibrium under rule 2 such that $\exists i \in C$, with $R^i \neq R^*$. Also, by Claim 2*, we have established that in any strong equilibrium that falls in Rule 2, $\forall j \in N \setminus C, R^j = R^*$. So, we consider a case where $\forall j \in N \setminus C, s_j = (a, R^*, NF, \cdot)$ with $a \in f(R^*)$ and $\forall k \in C, s_k = (a^k, R^k, F, n^k)$ such that $R^k \neq R^*$ for some $i \in C$, that is, $\exists i \in C$ such that $s_i \notin T_i^G(R^*)$. Moreover, let $g(s) = b$. Now we take two mutually exclusive cases:

Case 1: $|C| \geq 2$

- If $b = a$, then, since we have that $a \in L_i(a, R^*)$ by definition, agent i can play $s_i' = (a, R^*, F, n^i) \in T_i^G(R^*)$ with a sufficiently high n^i inducing rule 2 without changing the outcome and becoming strictly better off.
- If $b = a^l \neq a$, where $l = \min\{\text{argmax}_{j \in C} n^j\}$, we distinguish two cases:
  - $l \neq i$: In this case, since $a^l \in \cup_{j \in C} L_j(a, R^*)$, agent i can deviate to $s_i' = (b, R^*, F, n^i) \in T_i^G(R^*)$, win the integer game for a sufficiently high integer without affecting the outcome, and thus become better off by rule 2.
  - $l = i$: Again, $a^l \in \cup_{j \in C} L_j(a, R^*)$, so i can play $s_i' = (b, R^*, F, n^i) \in T_i^G(R^*)$ and again become better off by rule 2.

Case 2: $C = \{i\}$.

- If $b = a$, then i can deviate to $s'_i = (a, R^*, NF, \cdot) \in T_i^G(R^*)$ inducing Rule 1 and become better off by announcing the truth.
- If $b \neq a$, then it must be that $b = a^i$. So, since $b \in L_i(a, R^*)$, i can revert to truth-telling by playing $s'_i = (b, R^*, F, n^i) \in T_i^G(R^*)$ and become better off by rule 2.

Therefore, there is no strong equilibrium under rule 2 where $\exists i \in C$ such that $R^i \neq R^*$. □

Claim 4*: There is no strong equilibrium under rule 3 where $\exists i \in N$, with $R^i \neq R^*$.

Proof. Suppose there exists a strong equilibrium under rule 3 where $\forall j \in N, s_j = (a^j, R^j, F, n^j)$, $g(s) = b$ and let $R^i \neq R^*$ for some $i \in N$, that is, $\exists i \in N$ such that $s_i \notin T_i^G(R^*)$. Then, i can deviate to $s'_i = (b, R^*, F, n^i) \in T_i^G(R^*)$ and obtain b while announcing the true state R^*, for a sufficiently high integer n^i. Therefore, s cannot be a strong equilibrium. □

Claim 5*: There is no strong equilibrium under rule 4.

Proof. Suppose on the contrary that there exists one, namely $s \in S$, with $g(s) = a_W$. So, $\forall C \subseteq N, \forall s'_C \in S_C, \exists i \in C, (s_C, s_{N\setminus C}) \succeq_i^{R^*} (s'_C, s_{N\setminus C})$. Consider the case where $C = N$ and let $g(s') = a \in f(R^*)$. Then, there exists $i \in N$ such that:

- $(s_C, s_{N\setminus C}) \succ_i^R (s'_C, s_{N\setminus C})$ (1), or
- $(s_C, s_{N\setminus C}) \sim_i^R (s'_C, s_{N\setminus C})$ (2).

Suppose (1) holds. Then, either $g(s) = a_W P_i a = g(s') \in f(R^*)$, which is a contradiction of **UWA**, or $g(s) = a_W I_i^* a = g(s') \in f(R^*), s_i \in T_i^G(R^*)$ and $s'_i \notin T_i^G(R^*)$, where we have a contradiction as well. If (2) holds, then $g(s) = a_W I_i^* a = g(s') \in f(R^*)$ and the same contradiction emerges. So, there is no strong equilibrium under rule 4 and this completes the proof. □

Corollary 2. Any strong equilibrium s of the mechanism G, falls under rules 1-3 and it also holds that $\forall i \in N, R^i = R^*$.

Proof. Immediate implication of Claims 1*-5*. □

By the above arguments, we can restrict attention to strong equilibria under rules 1, 2 or 3, where $\forall i \in N, R^i = R^*$. Consider a strong equilibrium under rule:

1. That is, $\forall i \in N, s_i = (a, R^*, NF, \cdot)$. Then $g(s) = a \in f(R^*)$.

2. That is, $\forall i \in N \setminus C, s_i = (a, R^*, F, \cdot)$ with $a \in f(R^*)$, and $\forall j \in C, s_j = (a^j, R^*, F, n^j)$. Let $g(s) = b$. We distinguish two cases:

   $|N \setminus C| \geq 2$: Then, it must be that $\forall i \in N \setminus C, g(S_i, s_{N\setminus\{i\}}) = \cup_{j \in C \cup \{i\}} L_j(a, R^*)$ and $\forall j \in C, g(S_j, s_{N\setminus\{j\}}) = \cup_{j\in C} L_j(a, R^*)$, from Rule 2. For s to be a strong equilibrium, it must hold that $\forall i \in N \setminus C, L_i(a, R^*) \subseteq \cup_{j\in C\cup\{i\}} L_j(a, R^*) \subseteq L_i(b, R^*)$ and, $\forall j \in C, L_j(a, R^*) \subseteq \cup_{j\in C} L_j(a, R^*) \subseteq L_j(b, R^*)$. So, for any $i \in N$ we have that $L_i(a, R^*) \subseteq L_i(b, R^*)$. However, since $a \in f(R^*)$, from **WPO**, it cannot be the case that $\forall i \in N, bP_i^* a$. So there must exist $j \in N$ such that $aI_j^* b$. From **WPD** it follows that $b \in f(R^*)$.

   $N \setminus C = \{i\}$: Then, for i it must hold that $g(S_i, s_{N\setminus\{i\}}) = A$ from rule 3, and $\forall j \in C$ it must hold that $g(S_j, s_{N\setminus\{j\}}) = \cup_{j\in C} L_j(a, R^*)$ by rule 2. For s to be a strong equilibrium, it must hold that $\forall i \in N \setminus C, L_i(a, R^*) \subseteq A \subseteq L_i(b, R^*)$ and $\forall j \in C, L_j(a, R^*) \subseteq \cup_{j\in C} L_j(a, R^*) \subseteq L_j(b, R^*)$. So for all $i \in N$ it holds that $L_i(a, R^*) \subseteq L_i(b, R^*)$. As before, from **WPO** and the fact that $a \in f(R^*)$, there must exist $j \in N$ such that $aI_j^* b$. Again, from **WPD** we must have that $b \in f(R^*)$.

3. That is, $s_i = (a^i, R^*, F, n^i), \forall i \in N$ and let $g(s) = b$. Then, $\forall i \in N$, it must hold that $g(S_i, s_{N\setminus\{i\}}) = A$. Now, for s to be a strong equilibrium it must be that $\forall i \in N, A \subseteq L_i(b, R^*)$. Then, from **WPO**, **WPD** and Proposition 1, it must hold that $b \in f(R^*)$.

This completes the proof.

Chapter 3

Conditional rights and implementation

3.1 Introduction

Decentralized systems are prevalent in modern societies, from governance and administration, to markets and provision. It is now widely recognized at least since the work of Hayek, that decentralization allows for more efficient use of information that is dispersed across the economy. His analysis in Hayek (1980) has shaped economic thought by providing a theoretic foundation for decentralization against central planning. Nevertheless, the performance of decentralized systems in realizing collective goals by providing appropriate individual incentives is not always clear.

The motivation for this study comes from the fact that in many real-life cases, decentralization is not a matter of design or choice for a social planner. Arguments in favour of decentralization such as informational processing constraints, accountability, or balance of power, dictate that economic design has to work around that fact, by taking it as a constraint. For example, when a government designs a scheme for the provision of a public good, it has to take the private goods markets as given. In other instances, such as when a new academic institution is established, informational constraints might dictate that the allocation of instructors and GTAs (Graduate Teaching Assistants) to courses should be handled by different administrative departments. Similarly, in a collective bargaining agreement, for tractability, the issue of wages and the issue of work conditions may be handled by different bargaining councils and procedures. Finally, in the World Trade Organization (WTO), negotiations take place by subject.

On the other hand, society's goals might be concerned with the overall fairness, efficiency, etc. When a government runs procurement auctions for public projects, overall fairness may be compromised, if a firm acquires all projects. In other words, from a welfare perspective, the social choice might not be independent across sectors.

The problem becomes even more interesting given that agents may in general have non-separable preferences over sectoral outcomes. To give an example, imagine a university where each course must be allocated to an instructor/tutor pair. Instructors have preferences over pairs of courses and tutors, and tutors have preferences over pairs of courses and instructors. Suppose that tutor w is very experienced in course h, but is totally incompetent in assisting with course $\bar{h}$, while tutor $\bar{w}$ is an overall decent tutor in both courses. Then, the preferences of an instructor m could be represented as:

$$(w, h) \succ_m (\bar{w}, \bar{h}) \succ_m (\bar{w}, h) \succ_m (w, \bar{h})$$

Of course, these preferences are not separable, as the ranking of w and $\bar{w}$ depends on the choice of the course. Similarly, in the procurement auction example, firms may have complementarities in their preferences on public projects, due to economies of scale.

To give an intuition of our model and the questions that we attempt to answer, consider the above scenario with the instructor/tutor pair. Now suppose that different administrative departments have to be responsible for the allocation of tutors and instructors to courses respectively. A relevant question that we answer later in this paper is, can the university design such decentralized institutions, but at same time achieve overall stability, that is, reach a matching where no pair of instructor/tutor can block?

Naturally, in our model we consider two sectors and we ask two questions: First, suppose that one of these two institutions, say sector 2, is fixed, that is, we consider it an *institutional constraint*. Can we design the institution in sector 1 by taking into account sector 2 such that we implement the desirable outcomes in both sectors, and which rules are implementable in this case? This is relevant when the institution in the fixed sector represents some inalienable rights or some power distribution that the designer cannot affect.

Our first exercise provides a good understanding of the problem of conditionality in institution design. Therefore, for our second question, we are able to push the concept even further. In this case the social planner has to operate in a decentralized environment, which implies the existence of different institutions that deal with different issues. Therefore, she has to design two institutions *conditional* on each other. Moreover, each institution is decisive on its respective sectoral outcome. This case is relevant when decentralization is simply given, or desirable for other reasons, such as informational efficiency and accountability. We ask which rules are implementable in this manner and implicitly show the difficulties that arise in such case.

In order to answer these questions, we utilize the notion of *rights structures*, introduced by Koray and Yildiz (2018) to formalize the idea of an institution. In this setting, the social planner endows agents or coalitions with rights over changing the state. In particular, she designs three objects: (i) a state space, which provides the conditionals on which rights are tailored upon, (ii) an effectivity correspondence, that specifies, for any possible ordered pair of states, the family of coalitions that have the *right* to change the state from one to the other, and finally, (iii) an outcome function, which maps states to outcomes. The interpretation of the setting is that a hypothetical social planner wishes to implement a particular *social choice rule* (SCR) that assigns a set of socially optimal outcomes to any possible preference profile of the agents. However, she does not know the true preference profile. Thus, she has to design a rights structure that will implement the desired SCR. As usual, implementation is achieved when all outcomes realized in equilibrium are socially optimal, and all socially optimal outcomes can be realized through some equilibrium.

The equilibrium concept that we use for all of our results is a generalization of the γ-*equilibrium*[1] to the two-sector environment. Roughly speaking, a state is a γ-equilibrium if any coalition that has an incentive to change the state (in any one, or in both sectors "simultaneously") does not have the power to do so and vice versa, any coalition that has the right to change the state would not benefit from this change.

Implementation with rights structures possesses various appealing properties. First, it can be considered as a natural way of abstracting away the complex interactions that occur in a society, and focus on the most significant aspects, while at the same time, it bears a straightforward analogue to the exercise of rights in the real world. Second, in some cases it is difficult or even impossible for the planner to specifically plan in advance for any possible strategy that the agents may choose to play in a noncooperative mechanism. This is highly important when the constitutional power of the planner over the agents is limited. In these settings, rights structures deal with these shortcomings in a natural manner by allowing for cooperative game-theoretic solutions, where the details of the interaction among agents are left unspecified. Finally, implementation with rights does not utilize classic implementation devices such as integer or modulo games that have unnatural characteristics, and allows for neat characterizations that outline the essence of the implementation problem.

A crucial aspect of our model is intersectoral communication between institutions. In order to formalize it, we introduce the notion of *conditional rights structures*. The idea is to condition one sector's effectivity correspondence and/or

[1] Korpela (2013) refer to this solution concept as *core-equilibrium*, as it is very similar to the solution concept of the core.

outcome function to features of the other sector. The interpretation is that, a coalition's effectiveness to change the state in a particular sector is also conditional on the current state in the other sector. The idea is quite natural. For example, one's ability to exercise their right for welfare benefits depends on whether they are already employed and vice versa. The implications of the intersectoral communication on our model are substantial. On one hand, we are able to escape the inevitable decomposability of the social choice rule as in Hayashi and Lombardi (2017). On the other hand, given the incomplete nature of this communication, our results outline the restrictiveness of decentralization, especially when the social choice in one sector is not independent of the social choice in the other. Nevertheless, we feel that our approach provides a natural foundation for conditionality in constitution design.

To answer our first question when one institution is fixed, we fully characterize the social choice rules that are *constrained conditionally implementable*, given a fixed rights structure in sector 2. Our characterization consists of two conditions which are together necessary and sufficient. The first one, condition $\mathcal{E}$ bears a twofold role: First, it guarantees that the fixed rights structure is well-behaved, while ensuring the existence of suitable punishment-outcomes in sector 1 for possible deviations in sector 2. Second, it incorporates a *constrained-monotonicity* condition, similar to the one in Hayashi and Lombardi (2019). It is essentially a variation of the well-known *Maskin-monotonicity*[2] that takes into account the outcomes for which coalitions are effective for in sector 2. Our second condition, *constrained unanimity*, is a variant of the usual unanimity condition in the same manner. Therefore, our results are immediately comparable to Korpela et al. (2018), as we provide "constrained" versions of their conditions.

Our second task we answer by studying *conditional implementation*. We provide a condition called *projection-monotonicity* which is necessary. It is weaker than Maskin-monotonicity, as it requires an outcome to remain socially optimal, when there is no preference reversal between it and any other outcome that can be constructed as combination from the *projections* of the lower contour sets. This outlines the difficulties that decentralization poses on the implementation problem. We then proceed by providing simple sufficient conditions for conditional implementation under a domain restriction.

As application of constrained conditional implementation, we consider a matching environment, where matching occurs in triplets between a set of managers, a set of workers and a set of projects. We assume that the sector where projects are assigned to workers is fixed and show that, under some natural assumptions, the

[2]Maskin-monotonicity is due to the seminal study of Maskin (1999). It roughly says that if an outcome x is selected as socially optimal in a preference profile R, but not in R', then there must have been a preference reversal for at least one agent between x and some other outcome in R'. Maskin-monotonicity is necessary and almost sufficient for implementation in Nash equilibrium.

social planner can indeed overcome this constraint and implement the stable rule, when managers are the "owners" of the projects. To the best of our knowledge, this is the first study of implementability in such environment. Our application of conditional implementation is a multi-issue environment with lexicographic preferences. We show that the rule that assigns the weakly Pareto efficient outcomes can be conditionally implemented.

The contribution of this paper boils down to two issues: First, we extend the recent literature on decentralized mechanism design by incorporating cooperative implementation concepts, such as rights structures. Apart from extending the notion of γ-equilibrium to a two-sector setting, our rights structure approach naturally describes the exercise of conditional and unconditional rights in the real world. Secondly and most importantly, by utilizing rights structures, we study the new problem of conditional implementation, where the planner is forced to design one institution for each issue, and not a centralized institution to deal with all issues "simultaneously". We show that some of the difficulties of constrained implementation persist even in this case, as it is not a lot more permissive. Our results are complimented with examples and applications.

The remainder of the paper is organized as follows: In section 3.2 we provide a review of the relevant literature, in section 3.3 we present the formal environment and the basic definitions and in section 3.4 we present our conditions and implementation results. In section 3.5 we present our applications and in section 3.6 we discuss our results and conclude. The proofs of the main theorems are in the appendix.

3.2 Literature

Decentralization has been studied extensively in the mechanism design literature. For an in-depth literature review we refer to Mookherjee (2006). Even in the abstract implementation theory framework though, decentralization is not a new issue. Since its very infant stage, implementation theory has been used as a formal way to study institutions and the comparison of different economic systems. We distinguish two main strands:

Direct vs indirect game forms: According to this early literature, a centralized institution is represented by a *direct game form*, where agents communicate only their information to the central authority. An early but powerful result established independently by Gibbard (1973) and Satterthwaite (1975) implies that no non-trivial social choice rule can be implemented by a direct game form in a *strategy-proof* manner. A reaction to this negative result was the relaxation of the manipulation desiderata. Maskin (1999) established some positive results on what is now well-known as *Nash implementation of social choice correspondences* with an *indirect game form*, which can be interpreted as a decentralized institution, where agents' messages are more complex and arbitrary. While some positive results in strategy-proofness have been recovered for restricted preference domains, in general domains Nash implementation has produced some remarkably powerful and general theorems. Even though it is out of the scope of this paper to survey the entire literature on implementation with indirect game forms[3], we mention some important contributions: Maskin (1999) and Moore and Repullo (1990) for Nash implementation, Abreu and Sen (1990) and Vartiainen (2007*b*) for subgame perfect implementation, and Mezzetti and Renou (2017) for repeated Nash implementation. For incomplete information environments, see Jackson (1991) for example.

Multiple issues/sectors: Another relatively more recent way in which decentralization appears in implementation theory is through the issue of multiple sectors. Decentralization in this case manifests in the existence of several institutions that deal with different social or economic issues, rather than a centralized institution dealing with all issues. A seminal contribution along this line of research is Breton and Sen (1999) who study strategy-proofness in a multi-issue environment. They show that with separable strict preferences and a sufficiently rich domain, strategy-proof social choice functions must essentially be decomposable. In a more recent contribution, Hayashi and Lombardi (2017) study whether the usual partial equilibrium analysis that is prevalent in the economic design literature is innocuous. They show that in the presence of multiple sectors that can only conceive of separable preferences and no intersectoral communication, the possibilities for implementation are limited. In Hayashi and Lombardi (2019), the paper most closely to ours, the issue of implementation with institutional constraints is studied with the use of noncooperative game forms. Our constrained conditional implementation results are complimentary to theirs in the rights structure framework however, we also expand our scope to conditional implementation.

Our contribution also falls into the literature of implementation using effectivity functions. Early but classic contributions in this area include Moulin and Peleg (1982) and Greenberg (1990). In Peleg and Winter (2002), an effectivity function E represents the constitution of the society. In this case, they study conditions for a noncooperative game form to Nash implement a SCR, while the effectivity function derived from the game form has to be compatible with E. More recently, Koray and Yildiz (2018) introduced the notions of a rights structure and code of rights[4],

[3]For more extensive reviews of implementation theory the reader can see Jackson (2001), Palfrey (2002) or Maskin and Sjöström (2002). For a more recent overview, one can see Corchón (2015).

[4]A code of rights is a rights structure, where the set of possible states coincides with set of possible outcomes and the outcome function is the identity map.

where implementation is achieved in a cooperative game-theoretic framework[5]. They provide necessary and sufficient conditions for implementation with rights structures on the full domain of preferences. Korpela et al. (2018) provide some more results on implementation with codes of rights, by outlining the value of utilizing coalitions in economic design.

Another related to our paper field is the modeling of rights in the social choice literature. Since the impossibility result of Sen (1970), a substantial literature has emerged that attempts to formulate the exercise of rights in a formal game-theoretic framework, with the use of effectivity functions or noncooperative game forms, such as Gibbard (1974), Gärdenfors (1981), Hammond (1996), Peleg (1997) and Deb et al. (1997). In The more recent contribution of McQuillin and Sugden (2011), they model rights with the use of games in transition function form, which is much closer to the notion of a rights structure.

3.3 Environment

3.3.1 Primitives

The set of agents is N, with $|N| \geq 3$, while the set of all possible coalitions of agents is denoted by $\mathcal{N} = 2^N$ and the set of all non-empty coalitions by $\mathcal{N}_0$. The set of all possible outcomes is denoted by $X^1 \times X^2$, where X^1 is the set of outcomes in sector 1, while X^2 is the set of outcomes in sector 2. For every agent i, we define a complete reflexive and transitive binary relation on $X^1 \times X^2$, denoted by R_i, which represents agent i's preferences over X. As usual, by P_i we denote the asymmetric part of R_i. Let the set of all possible preferences for i be $\mathcal{R}_i$. The set of all possible preference profiles is denoted by $\mathcal{R} \equiv \times_{i\in N}\mathcal{R}_i$, with a typical element R. By slightly abusing notation, we extend the relation R to coalitions, such that for all $K \in \mathcal{N}_0$, (x^1, x^2) and $(y^1, y^2) \in X^1 \times X^2$, we write $(x^1, x^2)R_K(y^1, y^2)$ if and only if there exists $i \in K$ such that $(x^1, x^2)R_i(y^1, y^2)$. Now we define, for any agent $i \in N$, outcome $(x^1, x^2) \in X^1 \times X^2$ and preference profile $R \in \mathcal{R}$, $L_i((x^1, x^2), R) \equiv \{(y^1, y^2) \in X^1 \times X^2 | (x^1, x^2)R_i(y^1, y^2)\}$ as agent i's *lower contour set* with respect to outcome (x^1, x^2) in preference profile R. Then, for any $K \in \mathcal{N}_0$, $(x^1, x^2) \in X^1 \times X^2$ and $R \in \mathcal{R}$, $L_K((x^1, x^2), R) \equiv \cup_{i\in K} L_i((x^1, x^2), R)$.

A *social choice rule* (SCR) is a correspondence $\phi : \mathcal{R} \rightrightarrows X^1 \times X^2$, such that, for any $R \in \mathcal{R}$, $\emptyset \neq \phi(R) \subseteq X^1 \times X^2$. By $\phi(\mathcal{R})$ we denote the image of ϕ, that is $\phi(\mathcal{R}) \equiv \{(x^1, x^2) \in X^1 \times X^2 | (x^1, x^2) \in \phi(R)$, for some $R \in \mathcal{R}\}$.

[5]For some earlier similar results on implementation with cooperative game-theoretic devices see Miyagawa (2002). An earlier notion of a rights structure is due to Sertel (2002).

3.3.2 Constrained conditional rights structures

A *rights structure* for sector 2 is defined as $\Gamma^2 \equiv (S^2, h^2, \gamma^2)$, and consists of a state space S^2, an outcome function $h^2 : S^2 \to X^2$, and an effectivity correspondence $\gamma^2 : S^2 \times S^2 \rightrightarrows \mathcal{N}$. Let $\mathcal{G}^2$ be the set of all possible rights structures for sector 2. Then, given $\Gamma^2 \in \mathcal{G}^2$, a *constrained* (from Γ^2) *rights structure* $\Gamma^1 \equiv (S^1, h^1, \gamma^1)$ for sector 1 consists of a state space S^1, a *constrained outcome function* $h^1 : S^1 \times S^2 \to X^1$, and a *constrained effectivity correspondence*[6] $\gamma^1(\gamma^2) : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$. A pair $\Gamma^1 \times \Gamma^2 \equiv \Gamma^c$ is called a *constrained conditional rights structure.* A convenient assumption that we make is that for any $K \in \mathcal{N}_0$ and $(s^1, s^2) \in S^1 \times S^2$, we have $K \in \gamma^i(s^1, s^1; s^2) \cap \gamma^2(s^2, s^2)$, which allows us to define our equilibrium concept more compactly. In essence, this assumption simply states that any coalition is effective for no change or, equivalently, inaction is a possibility. For clarity, we summarize in the following table:

Object	Sector 2	Sector 1
State space	S^2	S^1
Outcome function	$h^2 : S^2 \to X^2$	$h^1 : S^1 \times S^2 \to X^1$
Effectivity correspondence	$\gamma^2 : S^2 \times S^2 \rightrightarrows \mathcal{N}$	$\gamma^1 : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$

Table 3.1: Constrained conditional rights structures

The intuition behind the notation is straightforward. While the given institution in sector 2 utilizes information only from the same sector, the institution in sector 1 processes information from sector 2 as well. This information is expressed in the conditionality of the outcome function and effectivity correspondence of sector 1.

3.3.3 Conditional rights structures

We define a pair of *conditional rights structures* as $\Gamma \equiv \Gamma^1 \times \Gamma^2$, where $\Gamma^1 = (S^1, h^1, \gamma^1)$ and $\Gamma^2 = (S^2, h^2, \gamma^2)$. S^1 and S^2 are the state spaces, $h^1 : S^1 \times S^2 \to X^1$ and $h^2 : S^2 \times S^1 \to X^2$ are the *conditional outcome functions*, and $\gamma^1 : S^1 \times S^2 \times S^2 \rightrightarrows \mathcal{N}$ and $\gamma^2 : S^2 \times S^2 \times S^1 \rightrightarrows \mathcal{N}$ are the *conditional effectivity correspondences* for sectors 1 and 2 respectively. We make a similar assumption about inaction being a possibility in this setting as well and we summarize in the following table:

The interpretation here is that the social planner is forced to design an institution that is decisive for each sectoral outcome, in a decentralized manner.

[6] To ease notation, we will omit writing γ^1 as function of γ^2 for the rest of the paper.

Object	Sector 1	Sector 2
State space	S^1	S^2
Outcome function	$h^1 : S^1 \times S^2 \to X^1$	$h^2 : S^2 \times S^1 \to X^2$
Effectivity correspondence	$\gamma^1 : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$	$\gamma^2 : S^2 \times S^2 \times S^1 \rightrightarrows \mathcal{N}$

Table 3.2: Conditional rights structures

The incomplete communication that takes place between the two institutions is expressed through conditionality. We proceed with our equilibrium notions.

3.3.4 Equilibrium notions

Our equilibrium notions are generalizations of the (weak) core equilibrium notion for the two sector setting.

Definition 3.3.1. Given a pair of constrained conditional rights structures $\Gamma^1 \times \Gamma^2$, a pair of states $(s^1, s^2) \in S^1 \times S^2$ is a γ-*equilibrium* in preference profile $R \in \mathcal{R}$ if there is no (t^1, t^2) and $K \in \mathcal{N}_0$, with $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2)$ such that $(h^1 \times h^2)(t^1, t^2) P_K (h^1 \times h^2)(s^1, s^2)$.

The equilibrium concept for conditional rights structures is very similar:

Definition 3.3.2. Given a pair of conditional rights structures, $\Gamma^1 \times \Gamma^2$, a pair of states $(s^1, s^2) \in S^1 \times S^2$ is a γ-*equilibrium* in preference profile $R \in \mathcal{R}$ if there is no (t^1, t^2) and $K \in \mathcal{N}_0$, with $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2; t^1)$ such that $(h^1 \times h^2)(t^1, t^2) P_K (h^1 \times h^2)(s^1, s^2)$.

A comment with regards to our equilibrium concept. As the reader may notice, we allow deviations by any coalition in one or both sectors "simultaneously". In fact, as it will become obvious later, it is this exact feature that makes our environment interesting. Now let $C(\Gamma^1 \times \Gamma^2, R) = \{(s^1, s^2) \in S^1 \times S^2 | (s^1, s^2)$ is a γ-equilibrium in profile $R\}$ be the set of γ equilibrium states in R.

3.3.5 Implementation concepts

The concept of implementation that we use is summarized in the following definition:

Definition 3.3.3. A pair of (constrained) conditional rights structures $\Gamma^1 \times \Gamma^2$ γ-*implements* the SCR ϕ if, for all $R \in \mathcal{R}$, we have that $\phi(R) = (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$.

A SCR ϕ is γ-implementable, or simply implementable, if it can be γ-implemented by a (constrained) conditional rights structure $\Gamma^1 \times \Gamma^2$. In the case of constrained conditional implementation of course, Γ^2 is given.

3.3.6 Examples

In this section we provide a few examples to compliment our formal definitions and provide a more intuitive and graphic view of our model.

First, we show an example of a constrained conditional rights structure. Let $X^1 = \{x^1, y^1\}$, $X^2 = \{x^2, y^2\}$ and $N = \{1, 2\}$. Now consider Γ^2 fixed as follows:

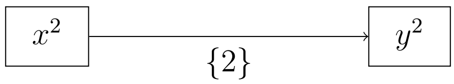

Figure 3.1: Example, fixed rights structure

In the above graph, rectangles represent the states, while arrows represent effectivity. So, in this case, only agent 2 has the power to change x^2 to y^2 and this is something that the planner has to take as given. Now consider the following two possible preference profiles:

R		R'	
1	2	1	2
(y^1, x^2)	(x^1, x^2)	(x^1, x^2)	(y^1, y^2)
(x^1, x^2)	(x^1, y^2)	(y^1, x^2)	(x^1, y^2)
(x^1, y^2)	(y^1, x^2)	(y^1, y^2)	(y^1, x^2)
(y^1, y^2)	(y^1, y^2)	(x^1, y^2)	(x^1, x^2)

Table 3.3: Example of constrained rights structure, preferences

Suppose that the SCR ϕ is such that $\phi(R) = \{(x^1, y^2), (y^1, x^2), (y^1, y^2)\}$, while $\phi(R') = \{(x^1, y^2), (y^1, y^2)\}$. So the social planner, needs to devise a rights structure Γ^1 for sector 1, such that the constrained conditional rights structure $\Gamma^1 \times \Gamma^2$ implements ϕ. Consider the constrained conditional rights structure below:

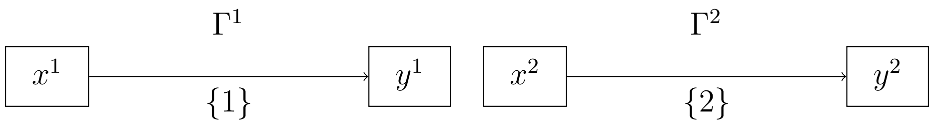

Figure 3.2: Constrained rights structures

That is, agent 1 has the right to change x^1 to y^1, regardless of the status quo outcome in sector 2. Now notice the following: (y^1, x^2), (x^1, y^2) and (y^1, y^2) are all

γ-equilibria in R. However, (x^1, x^2) is not, since agent 1 would like to move to y^1, as he prefers (y^1, x^2) to (x^1, x^2). Now let us consider profile R'. (y^1, y^2) is still a γ-equilibrium in R', but notice that (x^1, y^2) is not, since 1 would like to move from x^1 to y^1 and get (y^1, y^2) which he prefers to (x^1, y^2). In such case, implementation fails, since $(x^1, y^2) \in \phi(R')$.

Now consider the following rights structure for sector 1 instead:

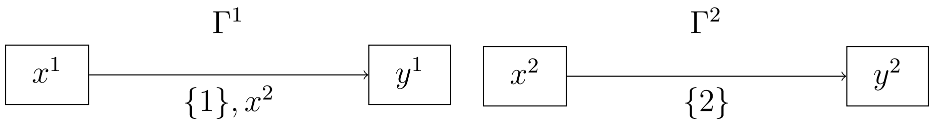

Figure 3.3: Constrained rights structures

In this case, agent 1 has the right to change x^1 to y^1, **only if** the status quo[7] in sector 2 is x^2. The reader will notice in this case that (x^1, y^2) is a γ-equilibrium in R' as well and implementation is restored.

The previous example outlined the power of conditionality: In the first case, the rights structure in sector 1 did not take into account any information about sector 2. This information was expressed in the form of a conditional right. Agent 1's right to move from x^1 to y^1 was conditional on the status quo in sector 2 being x^2.

To also illustrate the difficulties that arise with constrained conditional implementation, consider the following example: As before, we have $X^1 = \{x^1, y^1\}$ and $X^2 = \{x^2, y^2\}$ and a fixed rights structure for sector 2 as in Figure 3.1. Now consider the following preference profile:

R	
1	2
(x^1, x^2)	(y^1, y^2)
(y^1, x^2)	(y^1, x^2)
(y^1, y^2)	(x^1, y^2)
(x^1, y^2)	(x^1, x^2)

Table 3.4: Example of constrained rights structure, preferences

Suppose that the social planner desires to implement ϕ , were $\phi(R) = (y^1, x^2)$. Notice that no matter how she designs Γ^1, agent 2 can always change x^2 to y^2, as

[7]Notice that in the graph, below the arrow, x^2 appears next to $\{1\}$. This means that agent 1 is entitled to change x^1 to y^1, only if x^2 is status quo in sector 2.

y^2 is strictly better for him for any choice of outcome in sector 1. This makes the SCR ϕ non-implementable, even with one preference profile. This example hints towards our necessary conditions for constrained conditional implementation that we explore in the next section.

Finally, to show how the planner can solve the above problem by redesigning the rights structure in sector 2, we will show a conditional rights structure that implements the above SCR with two preference profiles, as shown below:

R		R'	
1	2	1	2
(x^1, x^2)	(y^1, y^2)	(x^1, x^2)	(y^1, y^2)
(y^1, x^2)	(y^1, x^2)	(y^1, y^2)	(y^1, x^2)
(y^1, y^2)	(x^1, y^2)	(y^1, x^2)	(x^1, y^2)
(x^1, y^2)	(x^1, x^2)	(x^1, y^2)	(x^1, x^2)

Table 3.5: Example of conditional rights structure, preferences

Let $\phi(R) = (y^1, x^2)$ as before and $\phi(R') = (y^1, y^2)$. Now consider the following conditional rights structure:

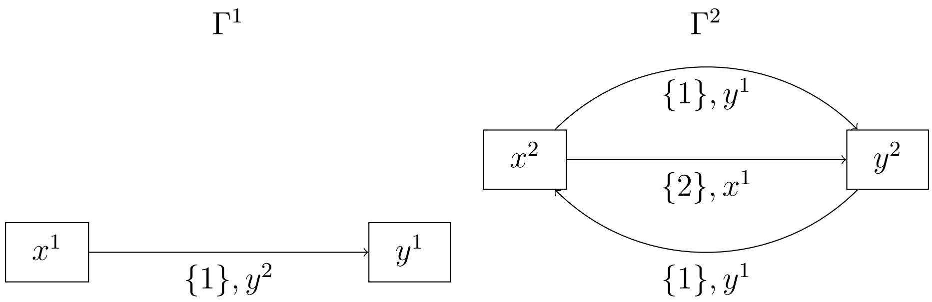

Figure 3.4: Conditional rights structures

In R, (y^1, x^2) is the only γ-equilibrium: From (x^1, x^2) agent 2 can move to (x^1, y^2), from (x^1, y^2) 1 moves to (y^1, y^2) and from (y^1, y^2) 1 can move to (y^1, x^2). On the other hand, no one can block (y^1, x^2).

Similarly, in R', the only equilibrium is (y^1, y^2). Notice that from (y^1, x^2) 1 moves to (y^1, y^2) which was not the case in R. Therefore, the above conditional rights structures implement ϕ.

3.4 Results

3.4.1 Constrained conditional implementation

Further definitions

Before we proceed to our results on constrained conditional implementation a few further definitions are in order. Given a pair of constrained conditional rights structures $\Gamma^1 \times \Gamma^2$, we define the range of Γ^2 as $h^2(S^2) \equiv \{h^2(s^2) \in X^2 | s^2 \in S^2\}$ and the range of Γ^1 as $h^1(S^1, S^2) \equiv \{h^1(s^1, s^2) \in X^1 | (s^1, s^2) \in S^1 \times S^2\}$. The range of $\Gamma^1 \times \Gamma^2$ is $(h^1 \times h^2)(S^1, S^2) \equiv \{(h^1(s^1, s^2), h^2(s^2)) \in X^1 \times X^2 | (s^1, s^2) \in S^1 \times S^2\}$. For any coalition $K \in \mathcal{N} \setminus \{\emptyset\}$ and state $s^2 \in S^2$, let $E(K, s^2) \equiv \{h^2(t^2) \in X^2 | t^2 \in S^2$ and $K \in \gamma^2(s^2, t^2)\}$ be the sector 2 outcomes that coalition K can induce from state s^2. Now, for any $K \in \mathcal{N}_0$ and $(s^1, s^2) \in S^1 \times S^2$ we define:

$$E(K, (s^1, s^2)) \equiv \{(h^1(t^1, t^2), h^2(t^2)) \in X^1 \times X^2 | K \in \gamma^1(s^1, t^1; t^2) \cap \gamma^2(s^2, t^2)\}.$$

That is, $E(K, (s^1, s^2))$ is the set of the attainable outcomes for coalition K from the state pair (s^1, s^2). Notice that for any $(x^1, x^2) \in E(K, (s^1, s^2))$, $x^2 \in E(K, s^2)$. The previous definitions allow us to define γ-equilibrium more compactly: A pair (s^1, s^2) is a γ-equilibrium in $R \in \mathcal{R}$, if for any $K \in \mathcal{N}_0$, we have $E(K, (s^1, s^2)) \subseteq L_K((h^1 \times h^2)(s^1, s^2), R)$.

Finally, for any set $T \subseteq X^1 \times X^2$, $i \in \{1, 2\}$ and $Z^i \subseteq X^i$, let $proj_{Z^i}\{T\}$ be the projection of T onto Z^i.

Necessary conditions

A helpful notion that will allow us to present our conditions more compactly is the notion of the constrained lower contour set. Given Γ^2, for any $(x^1, x^2) \in X^1 \times X^2$, $K \in \mathcal{N} \setminus \{\emptyset\}$, $s^2 \in S^2$ and $R \in \mathcal{R}$, let the sector 2 *constrained lower contour set* $L_K((x^1, x^2), s^2, R) \equiv \{(y^1, y^2) \in X^1 \times E(K, s^2) | (x^1, x^2) R_K (y^1, y^2)\}$ be the set of outcomes that K can induce from state s^2 that are weakly worse for K with respect to (x^1, x^2) in profile R. In addition, for any $Z^1 \times Z^2 \subseteq X^1 \times X^2$, let $L_K((x^1, x^2), \bar{s}^2, R)|_{Z^1 \times Z^2}$, be the restriction of $L_K((x^1, x^2), \bar{s}^2, R)$ to $Z^1 \times Z^2$. Now we are ready to present our necessary conditions.

Definition 3.4.1. Let Γ^2 be given. The SCR ϕ satisfies *Condition $\mathcal{E}$* if there exists a set $Y^1 \subseteq X^1$, with $proj_{X^1}\phi(\mathcal{R}) \subseteq Y^1$ and for all $R \in \mathcal{R}$ and $(x^1, x^2) \in \phi(R)$, there exists $\bar{s}^2((x^1, x^2), R) \equiv \bar{s}^2 \in S^2$ such that:

(i) $h^2(\bar{s}^2) = x^2$,

(ii) for all $t^2 \in S^2$ such that $\gamma^2(\bar{s}^2, t^2) \neq \emptyset$, we have that $h^2(t^2) \in \bigcap_{K \in \gamma^2(\bar{s}^2, t^2)} proj_{X^2}\{L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}\}$, and

(iii) for any $R' \in \mathcal{R}$, if for all $K \in \mathcal{N}_0$ we have $L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)} \subseteq L_K((x^1, x^2), \bar{s}^2, R')$, then it must be that $(x^1, x^2) \in \phi(R')$.

In our next proposition, we show that condition $\mathcal{E}$ is necessary for constrained conditional implementation.

Proposition 6. If a SCR ϕ is constrained conditionally implementable in γ-equilibrium, then it satisfies condition $\mathcal{E}$.

Proof. Consider Γ^2 as given. Suppose ϕ is constrained implementable by Γ^1 and let the true preference profile be $R \in \mathcal{R}$ with $(x^1, x^2) \in \phi(R)$. Define $Y^1 \equiv proj_{X^1}\{(h^1 \times h^2)(S^1, S^2)\}$. Obviously, from constrained implementability, it must be that $proj_{X^1}\phi(\mathcal{R}) \subseteq Y^1 \subseteq X^1$.

We will first prove part (i). By implementability, there exists a pair of states $(\bar{s}^1, \bar{s}^2) \in S^1 \times S^2$, such that $(h^1 \times h^2)(\bar{s}^1, \bar{s}^2) = (x^1, x^2)$ and $(\bar{s}^1, \bar{s}^2) \in C(\Gamma^1 \times \Gamma^2, R)$, that is, for all $K \in \mathcal{N}_0$, $E(K, (\bar{s}^1, \bar{s}^2)) \subseteq L_K((x^1, x^2), R)$. Set $\bar{s}^2 \equiv \bar{s}^2((x^1, x^2), R)$ and notice that $h^2(\bar{s}^2) = x^2$. This completes the proof of part (i).

We will now prove part (ii). For the sake of contradiction, suppose that there exists $t^2 \in S^2$, such that $h^2(t^2) \notin \bigcap_{K \in \gamma^2(\bar{s}^2, t^2)} proj_{X^2}\{L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}\}$. So, there exists $K \in \gamma^2(\bar{s}^2, t^2)$, $K \neq \emptyset$, such that $h^2(t^2) \notin proj_{X^2}\{L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}\}$. This implies that, for any $y^1 \in Y^1$, $(y^1, h^2(t^2)) \notin L_K((x^1, x^2), \bar{s}^2, R)$. Now take any (t^1, t^2), such that $K \in \gamma^1(\bar{s}^1, t^1; \bar{s}^2)$. It follows that $K \in \gamma^1(\bar{s}^1, t^1; \bar{s}^2) \cap \gamma^2(\bar{s}^2, t^2)$, while $(h^1(t^1, t^2), h^2(t^2)) \notin L_K((x^1, x^2), R)$. This however contradicts that $(x^1, x^2) \in (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$ and this completes the proof of (ii).

Moving to part (iii), consider a profile $R' \in \mathcal{R}$, such that for all $K \in \mathcal{N}_0$, $L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)} \subseteq L_K((x^1, x^2), \bar{s}^2, R')_{Y^1 \times h^2(S^2)}$ is true, yet, $(x^1, x^2) \notin \phi(R')$. Then, by constrained conditional implementability, it must be that $(x^1, x^2) \notin C(\Gamma^1 \times \Gamma^2, R')$, so there must exist $(t^1, t^2) \in S^1 \times S^2$ and a coalition $K \in \mathcal{N}_0$, such that $K \in \gamma^1(\bar{s}^1, t^1; \bar{s}^2) \cap \gamma^2(\bar{s}^2, t^2)$ and $(h^1 \times h^2)(t^1, t^2) \equiv (z^1, z^2) \notin L_K((x^1, x^2), R')$. By our assumption then we have $(z^1, z^) \notin L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}$. Obviously, $(z^1, z^2) \in Y^1 \times h^2(S^2)$, so it must be that $(z^1, z^2) \notin L_K((x^1, x^2), \bar{s}^2, R)$. But then, we have that $K \in \gamma^1(\bar{s}^1, t^1; \bar{s}^2) \cap \gamma^2(\bar{s}^2, t^2)$, while $(z^1, z^2) \notin L_K((x^1, x^2), R)$, which clearly contradicts that $(x^1, x^2) \in (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$. This completes the proof of (iii) and concludes the proof of Proposition 6.

□

Our condition contains three statements. Statement (i) is straightforward given the definition of γ-implementability: it simply states that, for a constrained conditionally implementable SCR ϕ, if (x^1, x^2) is ϕ-optimal for some $R \in \mathcal{R}$, then there must exist a "supporting" state $\bar{s}^2 = \bar{s}^2((x^1, x^2), R)$ such that $h^2(\bar{s}^2) = x^2$. Statement (ii) can be considered as a "punishment" condition that our environment must satisfy. We illustrate using the figure below:

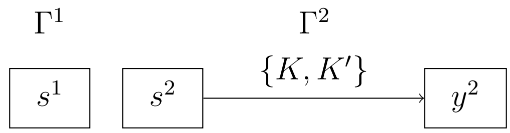

Figure 3.5: Condition $\mathcal{E}$, Rights structures

Consider the rights structures Γ^1 and Γ^2 as shown in Figure 1. Suppose that $(x^1, x^2) \in \phi(R)$ for some $R \in \mathcal{R}$, and $s^2 = \bar{s}^2((x^1, x^2), R)$, that is, $h^2(s^2) = x^2$. Now, if we hope to implement ϕ, it must be that deviations from s^2 in sector 2 are deterred when the preference profile is R. Indeed, suppose the status-quo in sector 1 is s^1. Notice that coalition K is effective for a change from s^2 to t^2. For this change to not be profitable, it must be the case that $(h^1(s^1, t^2), h^2(t^2)) \in L_K((x^1, x^2), s^2, R)$. Note though that this condition must also hold for any coalition that is effective for a change from s^2, so here it must be that $(h^1(s^1, t^2), h^2(t^2)) \in \cap_{i \in \{K, K'\}} L_i((x^1, x^2), s^2, R)$. This is the essence of our condition $\mathcal{E}$(ii).

How is our $\mathcal{E}$(iii) condition different from the well-known Maskin-monotonicity condition[8]? First we present Maskin-monotonicity formally in our environment and then we use an example to illustrate.

Definition 3.4.2. A SCR ϕ satisfies Maskin-monotonicity (with respect to a set Y), if for any $R, R' \in \mathcal{R}$, $(x^1, x^2) \in \phi(R)$, the following implication holds:

$$[\forall K \in \mathcal{N}_0, L_K((x^1, x^2), R) \cap Y \subseteq L_K((x^1, x^2), R')] \Rightarrow (x^1, x^2) \in \phi(R').$$

First, our condition only requires the nestedness of the lower contour sets with respect to a subset of $X^1 \times X^2$. Secondly and most importantly, the lower contour sets are restricted by the state in Γ^2. To illustrate, consider the following setting. Let $X^1 = \{x^1\}$, $X^2 = \{x^2, y^2, z^2\}$, $X^1 \times X^2 = \{(x^1, x^2), (x^1, y^2), (x^1, z^2)\}$, $\mathcal{R} = \{R, R'\}$ and the social choice rule ϕ defined as follows:

$$\phi(R) = \{(x^1, x^2), (x^1, z^2)\}, \ \phi(R') = \{(x^1, z^2)\}.$$

In the table we depict the rankings of agents 1 and 2 in preference profiles R and R' respectively and on the figure we have the rights structure Γ^2. The rectangles are the possible states, that is $S^2 = X^2$ and the arrows represent effectivity. For example, $\gamma^2(x^2, z^2) = \{1\}$. Notice that when moving

[8]Maskin-monotonicity as stated in Maskin (1999) is somewhat simpler than this version which is due to Korpela et al. (2018) who state it in terms of coalitions. Of course, if the nestedness holds for any coalition, it holds for individuals as well.

R_1	R_2	R'_1	R'_2
(x^1, x^2)	(x^1, z^2)	(x^1, y^2)	(x^1, z^2)
(x^1, y^2)	(x^1, x^2)	(x^1, x^2)	(x^1, x^2)
(x^1, z^2)	(x^1, y^2)	(x^1, z^2)	(x^1, y^2)

Table 3.6: Condition $\mathcal{E}$, Preferences

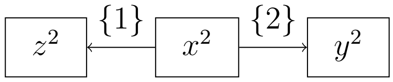

Figure 3.6: Condition $\mathcal{E}$, Γ^2

from R to R', $(x^1, x^2) \in \phi(R)$ falls in agent 1's ranking as there is a preference reversal with respect to (x^1, y^2), therefore the premise of Maskin-monotonicity is not fulfilled and it holds vacuously. However, (x^1, x^2) is not an attainable outcome for agent 1. Indeed, if we look at the constrained lower contour sets we have that $L_1((x^1, x^2), x^2, R) \subseteq L_2((x^1, x^2), x^2, R')$ and $L_2((x^1, x^2), x^2, R) \subseteq L_2((x^1, x^2), x^2, R')$, but $(x^1, x^2) \notin \phi(R')$, hence $\mathcal{E}$(iii) is violated and ϕ cannot be constrained conditionally implemented. Our condition has a very similar flavour to the one in Hayashi and Lombardi (2019), however it is tailored to our particular environment of implementation with rights structures.

The following Lemma is easily obtained from Proposition 1 and will be particularly useful for the proof of Theorem 4.

Lemma 3. Let the SCR ϕ satisfy condition $\mathcal{E}$. Then, for all $R \in \mathcal{R}$, $(x^1, x^2) \in \phi(R)$ and $t^2 \in S^2$, $\bigcap_{K \in \gamma^2(\bar{s}^2, t^2)} proj_{X^1}\{L_K((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}\} \neq \emptyset$.

Proof. Suppose that the premises are true and ϕ satisfies Condition $\mathcal{E}$. Now, consider $R \in \mathcal{R}$, $(x^1, x^2) \in \phi(R)$ and $t^2 \in S^2$ such that $\bigcap_{K \in \gamma^2(\bar{s}^2((x^1,x^2),R),t^2)} proj_{X^1}\{L_K ((x^1, x^2), \bar{s}^2, R)|_{Y^1 \times h^2(S^2)}\} = \emptyset$. This implies that there exists $K \in \gamma^2(\bar{s}^2, t^2)$, such that for all $y^1 \in Y^1$, $(y^1, h^2(t^2)) \notin L_K((x^1, x^2), \bar{s}^2, R)$. However, this contradicts $\mathcal{E}$(ii).

□

We proceed with our second condition, which we also explain using an example.

Definition 3.4.3. Let Γ^2 be given. A SCR ϕ satisfies *constrained unanimity* with respect to $Y^1 \subseteq X^1$ if we have $proj_{X^1}\{\phi(\mathcal{R})\} \subseteq Y^1$ and for all $R \in \mathcal{R}$, $x^1 \in Y^1$ and $s^2 \in S^2$, the following is true:

$$[\forall K \in \mathcal{N}_0, Y^1 \times E(K, s^2) \subseteq L_K((x^1, h^2(s^2)), s^2, R)] \Rightarrow (x^1, h^2(s^2)) \in \phi(R).$$

To illustrate, consider the following example:

$N = \{1, 2\}$, $X^1 = S^1 = \{x^1\}$, $X^2 = S^2 = \{x^2, y^2, z^2\}$ and the preferences in profile R and the rights structures Γ^1 and Γ^2 are as follows:

R_1	R_2
(x^1, z^2)	(x^1, z^2)
(x^1, x^2)	(x^1, y^2)
(x^1, y^2)	(x^1, x^2)

Table 3.7: Constrained unanimity, Preferences

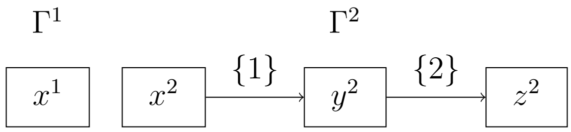

Figure 3.7: Constrained unanimity, Γ^1 and Γ^2

Notice that both agents prefer (x^1, z^2) to (x^1, x^2). However, by constrained unanimity we must have $(x^1, x^2) \in \phi(R)$. This is because z^2 is not attainable from x^2 by any coalition. Furthermore, since (x^1, x^2) is not top-ranked in the range of the outcome functions, it would not have to be socially optimal if we applied the standard unanimity concept as in Korpela et al. (2018). We now show that constrained unanimity is a necessary condition for constrained implementation with rights structures.

Proposition 7. Let Γ^2 be given. If a SCR ϕ is constrained conditionally implementable, then it satisfies constrained unanimity.

Proof. Let Γ^2 be given and suppose that ϕ is constrained conditionally implementable by Γ^1. First, let $Y^1 \equiv proj_{X^1}\{(h^1 \times h^2)(S^1, S^2)\}$. Now, consider $x^1 \in Y^1$. By the definition of Y^1, there exists $(s^1, s^2) \in S^1 \times S^2$, such that $h^1(s^1, s^2) = x^1$. Now suppose that for all $K \in \mathcal{N}_0$, we have $Y^1 \times E(K, s^2) \subseteq L_K((x^1, h^2(s^2)), s^2, R)$, yet, for the sake of contradiction, $(x^1, h^2(s^2)) \notin \phi(R)$. Then, by constrained conditional implementability, there exists $(t^1, t^2) \in S^1 \times S^2$ with $(t^1, t^2) \neq (s^1, s^2)$ and $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2)$, such that $(h^1(t^1, t^2), h^2(t^2))P_K(x^1, h^2(s^2))$. Notice though that $h^1(t^1, t^2) \in Y^1$ and $h^2(t^2) \in E(K, s^2)$. But, from out assumption, we have that for all $K \in \mathcal{N}_0$, $Y^1 \times E(K, s^2) \subseteq L_K((x^1, h^2(s^2)), s^2, R)$, a contradiction. □

Sufficient conditions

In the previous section, we proved that condition $\mathcal{E}$, as well as constrained unanimity are necessary conditions for constrained conditional implementation with rights structures. In this section we will show that they are also sufficient.

Theorem 4. If a SCR ϕ satisfies condition $\mathcal{E}$ and constrained unanimity, then it is constrained-implementable in γ-equilibrium.

Proof. See the Appendix. □

For the proof we use a canonical rights structure in the spirit of Koray and Yildiz (2018) and Korpela et al. (2018). Our novelty lies on the fact that we construct Γ^1 *conditionally* on Γ^2. Specifically, we design h^1 such that it takes as input not only the state in sector 1, but also the state in sector 2, that is $h^1 : S^1 \times S^2 \to X^1$. This essentially captures the flow of information between sectors. Furthermore, we design the effectivity correspondence $\gamma^1 : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$ so as to depend on three "issues": (i) pairs of (s^1, t^1), (ii) the status quo state in sector 2, s^2, and (iii) γ^2. What this construction essentially aims to capture is that whether a coalition K is effective for a change from s^1 to t^1, depends on what they are capable of obtaining from the status quo in sector 2, s^2. Our full characterization is a corollary of Proposition 6 and Theorem 4:

Corollary 3. Given a rights structure Γ^2 for sector 2, a SCR ϕ is constrained conditionally implementable if and only if it satisfies condition $\mathcal{E}$ and constrained unanimity.

Proof. Implication of Proposition 6 and Theorem 4. □

3.4.2 Conditional implementation

Further definitions

Similarly to the previous section on constrained conditional implementation, given a pair of conditional rights structures $\Gamma^1 \times \Gamma^2$, we define the attainable set as:

$$E(K, (s^1, s^2)) = \{(h^1(t^1, t^2), h^2(t^1, t^2)) \in X^1 \times X^2 | K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2; s^1)\}.$$

Necessary conditions

We first introduce our necessary conditions. Our first one is the well-known unanimity condition (as in Korpela et al. (2018)), which we show to be necessary in our case.

Definition 3.4.4. A SCR ϕ satisfies *unanimity* with respect to a set Y, if there exists $Y \supseteq \phi(\mathcal{R})$, such that for any $(x^1, x^2) \in Y$ and $R \in \mathcal{R}$:

$$\text{for all } K \in \mathcal{N}_0,\ Y \subseteq L_K((x^1, x^2), R) \Rightarrow (x^1, x^2) \in \phi(R).$$

Proposition 8. Let the SCR ϕ be conditionally implementable. Then it satisfies unanimity.

Proof. Suppose that the premises hold. Let $Y \equiv (h^1 \times h^2)(S^1, S^2) \supseteq \phi(\mathcal{R})$ and suppose there exists $(x^1, x^2) \in Y$ such that for all $K \in \mathcal{N}_0$, $Y \subseteq L_K((x^1, x^2), R)$, for some $R \in \mathcal{R}$. Now, since $(x^1, x^2) \in Y$, we have that $(x^1, x^2) = (h^1 \times h^2)(s^1, s^2)$, for some $(s^1, s^2) \in S^1 \times S^2$. Suppose that $(x^1, x^2) \notin \phi(R)$. Then, $(s^1, s^2) \notin C(\Gamma^1 \times \Gamma^2, R)$, so there must exist $(t^1, t^2) \in S^1 \times S^2$ and $K \in \mathcal{N}_0$, such that $(h^1 \times h^2)(t^1, t^2) P_K (x^1, x^2)$. But $(h^1 \times h^2)(t^1, t^2) \in Y \subseteq L_K((x^1, x^2), R)$, by our assumption, which is a contradiction. □

The fact that unanimity is necessary for conditional implementation is not surprising. In a setting where the social planner can design a "centralized" rights structure $\Gamma = (S, h, \gamma)$ and set $X = X^1 \times X^2$ as the set of outcomes, Korpela et al. (2018) have shown that unanimity is necessary for implementation. We now introduce a variant of Maskin-monotonicity, which is also necessary in our case.

Definition 3.4.5. A SCR satisfies *projection-monotonicity* with respect to a set Y, if there exists $Y \subseteq X^1 \times X^2$, such that for all $R, R' \in \mathcal{R}$ and $(x^1, x^2) \in \phi(R)$:

$$\text{for all } K \in \mathcal{N}_0,\ [proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\}] \cap Y \subseteq L_K((x^1, x^2), R') \Rightarrow (x^1, x^2) \in \phi(R').$$

Projection-monotonicity roughly says the following: Take $(x^1, x^2) \in \phi(R)$, for some $R \in \mathcal{R}$. Now consider a new profile $R' \in \mathcal{R}$ where $x \notin \phi(R')$. Then, there must exist an outcome (y^1, y^2) such that $(x^1, x^2) R_K (y^1, y^2)$ and $(y^1, y^2) P'_K (x^1, x^2)$ (preference reversal) for some $K \in \mathcal{N}$, where (y^1, y^2) is a combination of sectoral outcomes y^1 and y^2 that are in the respective projections of the lower contour sets of K for (x^1, x^2) in R. Below we prove that it is necessary for conditional implementation.

Proposition 9. If a SCR ϕ is conditionally implementable, then it must satisfy projection-monotonicity.

Proof. Consider a conditionally implementable SCR ϕ by a conditional rights structure $\Gamma^1 \times \Gamma^2$. Set $Y \equiv (h^1 \times h^2)(S^1, S^2)$. Suppose that for all $K \in \mathcal{N}_0$, $proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\} \cap Y \subseteq L_K((x^1, x^2), R')$. for some $R, R' \in \mathcal{R}$, where $(x^1, x^2) \in \phi(R)$. Suppose that $(x^1, x^2) \notin \phi(R')$. Since

$(x^1, x^2) \in \phi(R)$, there exists $(s^1, s^2) \in S^1 \times S^2$, such that $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$, where $(h^1 \times h^2)(s^1, s^2) = (x^1, x^2)$. Moreover, since $(x^1, x^2) \notin \phi(R')$, there exists $(t^1, t^2) \in S^1 \times S^2$ and $K \in \mathcal{N}_0$, with $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2; s^1)$, such that $(h^1 \times h^2)(t^1, t^2) P'_K (h^1 \times h^2)(s^1, s^2) = (x^1, x^2)$. By our assumptions then, it must be either that $(h^1 \times h^2)(t^1, t^2) \notin Y$, which is rejected by the very definition of Y, or $(h^1 \times h^2)(t^1, t^2) \notin proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\}$. This implies that either (i) $h^1(t^1, t^2) \notin proj_{X^1}\{L_K((x^1, x^2), R)\}$ or that (ii) $h^2(t^1, t^2) \notin proj_{X^2}\{L_K((x^1, x^2), R)\}$. Suppose that (i) is true. Then, for all $y^2 \in X^2$, we have that $(h^1(t^1, t^2), y^2) \notin L_K((x^1, x^2), R)$. Take $y^2 = h^2(t^1, t^2)$. In such case though, because $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2; s^1)$, we have a contradiction that $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$. A similar argument applies for (ii). This completes the proof. □

An interesting lemma is the following:

Lemma 5. If a SCR ϕ satisfies Maskin-monotonicity with respect to a set Y, then it also satisfies projection-monotonicity with respect to Y.

Proof. Suppose that a SCR ϕ satisfies Maskin-monotonicity with respect to Y and consider $R, R' \in \mathcal{R}$ and $(x^1, x^2) \in \phi(R)$, such that for all coalitions $K \in \mathcal{N}_0$, $[proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\}] \cap Y \subseteq L_K((x^1, x^2), R')$. Now take $(y^1, y^2) \in L_K((x^1, x^2), R) \cap Y$. Clearly, we have that $(y^1, y^2) \in [proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\}] \cap Y$, so $L_K((x^1, x^2), R) \cap Y \subseteq [proj_{X^1}\{L_K((x^1, x^2), R)\} \times proj_{X^2}\{L_K((x^1, x^2), R)\}] \cap Y \subseteq L_K((x^1, x^2), R')$. By Maskin-monotonicity we must have $(x^1, x^2) \in \phi(R')$ and we established that projection-monotonicity is also true. This completes the proof. □

In the previous lemma we showed that Maskin-monotonicity implies projection-monotonicity. The converse of our lemma is not true, as shown in the following example:

$X^1 = \{x^1, y^1, z^1\}$, $X^2 = \{x^2, y^2, z^2\}$ and $X = X^1 \times X^2$.

R_i	R'_i
(z^1, z^2)	(y^1, y^2)
(y^1, y^2)	(x^1, x^2)
(x^1, x^2)	(z^1, z^2)
(z^1, y^2)	(z^1, y^2)
(x^1, y^2)	(x^1, y^2)
(y^1, z^2)	(y^1, z^2)
(x^1, z^2)	(x^1, z^2)
(y^1, x^2)	(y^1, x^2)

Table 3.8: Projection-monotonicity, Preferences

Consider the above preference ranking for agent i and let $\phi(R) = \{(x^1, x^2)\}$, while $\phi(R') = \{(y^1, y^2)\}$. Now notice that:

- $L_i((x^1, x^2), R) = \{(x^1, x^2), (z^1, y^2), (x^1, y^2), (y^1, z^2), (x^1, z^2), (y^1, x^2)\}$
- $L_i((x^1, x^2), R') = L_i((x^1, x^2), R) \cup \{(z^1, z^2)\}$.
- $proj_{X^1}\{L_i((x^1, x^2), R)\} \times proj_{X^2}\{L_i((x^1, x^2), R)\} = X^1 \times X^2$.

Therefore, we have $L_i((x^1, x^2), R) \subseteq L_i((x^1, x^2), R')$ and Maskin-monotonicity would dictate that $(x^1, x^2) \in \phi(R')$, but this is not the case. Hence ϕ is not Maskin-monotonic. However, $proj_{X^1}\{L_i((x^1, x^2), R)\} \times proj_{X^2}\{L_i((x^1, x^2), R)\} \nsubseteq L_i((x^1, x^2), R')$, so the premise of projection-monotonicity is not satisfied and it holds vacuously.

Sufficient conditions

In this section, we present our sufficiency theorem, which will conclude our study of conditional implementation. Before introducing our conditions, we need a few further definitions.

For any $(x^1, x^2) \in X^1 \times X^2$, $K \in \mathcal{N}_0$ and $R \in \mathcal{R}$, we define:

$$L_K^{x^2}((x^1, x^2), R) \equiv \{(y^1, y^2) \in L_K((x^1, x^2), R) | y^2 = x^2\}$$

That is, $L_K^{x^2}((x^1, x^2), R)$ is the outcomes that are in the lower contour set of (x^1, x^2) for K in R, such that the outcome in sector 2 is x^2. Similarly we can define $L_K^{x^1}((x^1, x^2), R)$. We are now ready to state our main sufficient condition for conditional implementation:

Definition 3.4.6. A SCR ϕ satisfies *P*-monotonicity* with respect to a set Y, if there exists $Y \subseteq X^1 \times X^2$ where $\phi(\mathcal{R}) \subseteq Y$ such that for all $R, R' \in \mathcal{R}$ and $(x^1, x^2) \in \phi(R)$:

$$\text{for all } K \in \mathcal{N}_0,\ [proj_{X^1}\{L_K^{x^2}((x^1,x^2),R)\} \times proj_{X^2}\{L_K^{x^1}((x^1,x^2),R)\}] \cap Y \subseteq L_K((x^1,x^2),R') \Rightarrow (x^1,x^2) \in \phi(R').$$

For our sufficiency theorem, we need one more condition. This is a domain restriction, which is fairly easy to check in applications.

Definition 3.4.7. An environment satisfies *domain restriction-I* (**DR-I**) if for any $i \in N$, $R \in \mathcal{R}$, $(x^1, x^2) \in \phi(R)$ and $(y^1, y^2) \in proj_{X^1}\{\phi(\mathcal{R})\} \times proj_{X^2}\{\phi(\mathcal{R})\}$,

$$\{(x^1, y^2), (y^1, x^2)\} \subseteq L_i((x^1, x^2), R) \Rightarrow (y^1, y^2) \in L_i((x^1, x^2), R)$$

Notice that **DR-I** is stated in terms of individual agents however, we can actually work with a weaker version that includes coalitions. For the purpose of this paper, we will keep with the agent-based restriction. **DR-I** restricts the complementarity in preferences among sectors but it does not imply separability of preferences[9]. This is shown by the following example:

R_i
(x^1, x^2)
(y^1, x^2)
(y^1, y^2)
(x^1, y^2)

Table 3.9: **DR-I**, Preferences

Let $\phi(R) = \{(x^1, x^2)\}$. Notice that $\{(x^1, y^2), (y^1, x^2)\} \subseteq L_i((x^1, x^2), R)$ and $(y^1, y^2) \in L_i((x^1, x^2), R)$. Separability of preferences though would require that $(x^1, y^2) R_i (y^1, y^2)$, which is not the case.

P*-monotonicity and unanimity under **DR-I** become sufficient for conditional implementation. The combination of P*-monotonicity and **DR-I** guarantees that if there are no deviations in any one sector alone, then there will not be any deviation in both sectors "simultaneously", which is essentially the weak point of conditional implementation. We proceed with our last theorem:

Theorem 6. Suppose that **DR-I** holds. If a SCR ϕ satisfies P*-monotonicity and unanimity, then it is conditionally implementable.

Proof. The proof is in the Appendix. □

[9]Preferences are separable if $(x^1, x^2) R_i (y^1, y^2)$ implies that for all $z^2 \in X^2$, $(x^1, x^2) R_i (y^1, z^2)$.

3.5 Applications

3.5.1 Stable matching with partners and projects

Environment

As application of our Theorem 4, we consider a matching environment with partners and projects as in Combe (2017), where matching occurs in triplets, from three disjoint sets[10]. Our example of allocating GTAs and instructors to courses falls into this setting. In most cases, this matching is decentralized in the sense that the allocation of courses to instructors, is separated from the allocation of courses to GTAs. Hence, it is natural to wonder whether we can achieve stability in this case, when for example the rights structure for course allocation to instructors is something fixed. Other possible interpretations of this environment could include project-partner allocations in a firm, where for example a manager has to be matched with a worker in order to undertake a particular project, or assignment of male and female police officers to different patrolling duties[11]. What makes our environment "pure", is our assumption that no project can be left vacant or, equivalently, every agent prefers having a partner and/or a project to not having any[12].

The setting follows closely from Combe (2017) and is as follows: Let M, W, H be three fixed and disjoint sets, such that $|M| = |W| = |H| \geq 2$. We can interpret M as the set of managers, W as the set of workers and H as the set of projects. For any $I \in \{M, W\}$, an I-matching is a function $\mu_I : H \to I \cup \{\emptyset\}$, such that $h^1 \neq h^2 \implies \mu_I(h^1) \neq \mu_I(h^2)$. Let the set of all M and W-matchings be $\mathcal{M}_M$ and $\mathcal{M}_W$, respectively.

Now, a *matching* is a function $\mu : H \to [M \cup \{\emptyset\}] \times [W \cup \{\emptyset\}]$, such that for any $h \in H$, $proj_M\{\mu(h)\} \in \mathcal{M}_M$ and $proj_W\{\mu(h)\} \in \mathcal{M}_W$. We will also use the notion of the inverse image of μ_M and μ_W, that is, for any $I \in \{M, W\}$ and $i \in I$, we will write $\mu_I^{-1}(i) = h$, if $\mu_I(h) = i$ and $\mu_I^{-1}(i) = \emptyset$, if $\mu_I(h) = \emptyset$. Let the set of all matchings be $\mathcal{M}$. For any $I \in \{M, W\}$ and $i \in I$, P_i is a linear order on $(J \cup \{\emptyset\}) \times (H \cup \{\emptyset\})$, where $J \in \{M, W\} \backslash I$. We make the following assumption:

Assumption 1: For any $I \in \{M, W\}$, $i \in I$, $J \in \{M, W\} \setminus I$, $j, k \in J$ and $h, \bar{h} \in H$,

$$(j, h) P_i (\emptyset, \emptyset),\ (j, h) P_i (k, \emptyset) \text{ and } (j, h) P_i (\emptyset, \bar{h}).$$

[10] A similar environment can be found in Nicolo et al. (2019).

[11] This could be for example part of an affirmative action plan. In the UK there is significant effort to have a balanced workforce with respect to gender.

[12] Implementability with noncooperative mechanisms in one-to-one matching environments is studied in Tadenuma and Toda (1998), Korpela (2013) and Savva (2018).

Our assumption dictates that any agent prefers to be assigned to a project with a partner than not being assigned a project or a partner at all. For any $i \in \{m, w\}$, we extend the relation P_i to the set of all matchings as follows: For all $I \in \{M, W\}$, $J \in \{M, W\} \setminus I$, $i \in I$ and $\mu, \bar{\mu} \in \mathcal{M}$:

$$\mu \succ_i \bar{\mu} \iff (\mu_J(\mu_I^{-1}(i)), \mu_I^{-1}(i)) P_i (\bar{\mu}_J(\bar{\mu}_I^{-1}(i)), \bar{\mu}_I^{-1}(i)), \text{ and}$$

$$\mu \sim_i \bar{\mu} \iff (\mu_J(\mu_I^{-1}(i)), \mu_I^{-1}(i)) = (\bar{\mu}_J(\bar{\mu}_I^{-1}(i)), \bar{\mu}_I^{-1}(i))$$

The above expression simply states that the preferences over matchings are extensions of the preferences over objects and partners, in the usual sense: an agent prefers a matching over another if and only if she prefers her partner and project under this matching to her partner and project in the other matching. As usual, $\succsim_i$ represents the reflexive part of $\succ_i$ with the usual interpretation. The set of all possible preferences over $\mathcal{M}$ for any $i \in M \cup W$ is $\mathcal{R}_i$ and the set of all possible preference profiles is $\mathcal{R}$.

An *ownership structure* is a function $\theta : [M \cup \{\emptyset\}] \times [W \cup \{\emptyset\}] \times H \to M \cup W$, such that for any $(m, w, h) \in [M \cup \{\emptyset\}] \times [W \cup \{\emptyset\}] \times H$, $\theta(m, w, h) \in \{m, w\}$. Let the set of all ownership structures be Θ. Now, for any $I \in \{M, W\}$, θ_S^I is a *strong I-ownership structure* if for any $h \in H$, $\mu_I(h) \in I$ and for all $h \in H$ and $\mu_J \in \mathcal{M}_J$, $\theta_S^I(\mu_I(h), \mu_J(h), h) = \mu_I(h)$. Essentially, a strong I-ownership structure specifies that every project is assigned to an I member and the owner of any triplet is an I member as well. Now we define a concept of stability given an ownership structure.

Definition 3.5.1. $\mu \in \mathcal{M}$ is *individually rational* in profile $R \in \mathcal{R}$, if for any $I \in \{M, W\}$, $J \in \{M, W\} \setminus I$ and $i \in I$, we have $(\emptyset, \mu_I^{-1}(i)) P_i (\mu_J(\mu_I^{-1}(i)), \mu_I^{-1}(i))$.

Definition 3.5.2. For any $\theta \in \Theta$, $R \in \mathcal{R}$ and $\mu \in \mathcal{M}$, a *blocking triplet* $(m, w, h) \in M \times W \times H$ *with respect to* θ in profile R, is such that (1) **and** either one of the (2)-(4) is true:

(1) $(w, h) P_m (\mu_W(\mu_M^{-1}(m)), \mu_M(m))$ and $(m, h) P_w (\mu_M(\mu_W^{-1}(w)), \mu_W(w))$.

(2) $\mu_M^{-1}(m) = h$ and $\theta(m, \mu_W(h), h) = m$.

(3) $\mu_W^{-1}(w) = h$ and $\theta(\mu_M(h), w, h) = w$.

(4) $\mu_M(h) = \mu_W(h) = \emptyset$.

To give the intuition behind the last definition, a blocking triplet (m, w, h) is such that either one of the following is true:

- m is assigned to h and is the owner while both m and w prefer each other with h than their current assignments.
- w is assigned to h and is the owner while both m and w prefer each other with h than their current assignments.
- h is vacant while both m and w prefer each other with h than their current assignments.

Definition 3.5.3. For any $\theta \in \Theta$ and $R \in \mathcal{R}$, a matching $\mu \in \mathcal{M}$ is *strongly stable with respect to* θ, if it is individually rational and there exists no blocking triplet $(m, w, h) \in M \times W \times H$ with respect to θ.

Notice that, because of Assumption 1, individual rationality is never binding in our problem.

In this environment, a social choice rule is a correspondence $\varphi : \mathcal{R} \rightrightarrows \mathcal{M}$, such that for any $\succsim \in \mathcal{R}$, $\emptyset \neq \varphi(\succsim) \subseteq \mathcal{M}$. Given a strong I-ownership structure $\theta^I_S \in \Theta$, we define the stable rule given θ^I_S as $\varphi^S_I : \mathcal{R} \rightrightarrows \mathcal{M}$ such that for any $\succsim \in \mathcal{R}$, $\varphi^S_I(\succsim) = \{\mu \in \mathcal{M} | \mu$ is strongly stable with respect to θ^I_S in $\succsim\}$. While in general strongly stable matchings may not exist in this triple matching environment, the set of strongly stable matchings given a strong I-ownership structure is non-empty as shown in Combe (2017). Essentially, by augmenting the triple matching environment with an ownership structure, we restrict the possible blocking triplets by only allowing owners to block their assigned project.

Rights structure of W sector

Let us assume that the rights structure in the sector that assigns projects to workers is fixed, that is, it is not available for design. We will show that even when this is the case, we can still γ-implement the stable rule with respect to θ^W_S, by showing that the environment φ^S_W, Γ^W satisfies our conditions $\mathcal{E}$ and constrained unanimity. In particular, we assume that Γ^W is a *code of rights*, that is, a rights structure where the state space consists only of outcomes, and the outcome function is simply the identity map. Essentially, it is a more natural specification of a rights structure[13]. We formally define it below:

Let $\Gamma^W = (\mathcal{M}_W, \gamma^W)$, where $\gamma^W : \mathcal{M}_W \times \mathcal{M}_W \rightrightarrows 2^W$ is such that for all $w, w' \in W$, $\mu_W, \bar{\mu}_W \in \mathcal{M}_W$:

$$\{w\} \in \gamma^W(\mu_W, \bar{\mu}_W) \text{ if and only if there exists } \bar{h} \in H, \text{ such that } \mu_W(\bar{h}) = \emptyset, \\ \bar{\mu}_W(\bar{h}) = w, \text{ for all } \tilde{h} \in H \setminus \{h, \bar{h}\}, \bar{\mu}_W(\tilde{h}) = \mu_W(\tilde{h}), \text{ and } [\mu_W^{-1}(w) = h \text{ for some } \\ h \in H] \Rightarrow \bar{\mu}_W(h) = \emptyset.$$

[13]For more results in implementation with codes of rights, see Koray and Yildiz (2018) and Korpela et al. (2018).

Therefore, the code of rights Γ^W is such that it allows any worker to obtain a vacant project. We consider this a very natural code of rights.

Constrained conditional-implementation

In the next proposition we show that we can implement the stable rule under strong worker ownership, when the worker sector is fixed as above.

Proposition 10. Suppose that Γ^W is fixed. Then φ_W^S is constrained conditionally implementable in γ-equilibrium.

Proof. We will show that Γ^W and φ_W^S satisfy condition $\mathcal{E}$ and constrained unanimity.

Take $Y^1 \equiv \mathcal{M}_M \supseteq proj_{\mathcal{M}_M}\{\varphi_W^S(\mathcal{R})\}$. Now, consider $\succsim \in \mathcal{R}$ and let $(\mu_M, \mu_W) \in \varphi_W^S(\succsim)$.

First, notice that (i) from condition $\mathcal{E}$ is trivially satisfied, as we can set $\bar{s}^2((\mu_M, \mu_W), \succsim) \equiv (\mu_M, \mu_W)$.

Now we proceed to prove (ii). For the sake of contradiction, suppose there exists $\tilde{\mu}_W \in \mathcal{M}_W$ such that $\tilde{\mu}_W \notin \bigcap_{K\in\gamma^W(\mu_W,\tilde{\mu}_W)} proj_{\mathcal{M}_W}\{L_K((\mu_W, \mu_M), \mu_W, \succsim)\}$. This implies that there exists $K \in \gamma^W(\mu_W, \tilde{\mu}_W)$ such that for all $\tilde{\mu}_M \in \mathcal{M}_M$, $(\tilde{\mu}_M, \tilde{\mu}_W) \notin L_K((\mu_W, \mu_M), \mu_W, \succsim)$. By the construction of Γ^W, it must be that $K = \{w\}$, $\tilde{\mu}_W^{-1}(w) = h$ and $\mu_W^{-1}(w) = \emptyset$, for some $h \in H$. But, by Assumption 1 and since $|W| = |H|$, $(\mu_M, \mu_W) \notin \varphi_W^S$, so we have a contradiction.

Now let us prove (iii). Take $\succsim' \in \mathcal{R}$ such that for all $K \in \mathcal{N}_0$, $L_K((\mu_M, \mu_W), \mu_W, \succsim) \subseteq L_K((\mu_M, \mu_W), \mu_W, \succsim')$, but $(\mu_M, \mu_W) \notin \phi_W^S(\succsim')$. Then, there must exist $(m, w, h) \in M \times W \times H$, such that

$$(w, h)P'_m(\mu_W(\mu_M^{-1}(m)), \mu_M(m)),\ (m, h)P'_w(\mu_M(\mu_W^{-1}(w)), \mu_W(w)),$$

and either one of the following is true:

(1) $\mu_M^{-1}(m) = h$ and $\theta(m, \mu_W(h), h) = m$

(2) $\mu_W^{-1}(w) = h$ and $\theta(\mu_M(h), w, h) = w$.

(3) $\mu_M(h) = \mu_W(h) = \emptyset$.

(1) cannot be true due to strong W-ownership. (3) also cannot be the case as it implies $(\mu_M, \mu_W) \notin \varphi_W^S(\succsim)$ as well. Assume that (2) holds. Then, by our assumption about the constrained lower contour sets, it must be that $(w, h)P_m(\mu_W(\mu_M^{-1}(m)), \mu_M(m))$ and $(m, h)P_w(\mu_M(\mu_W^{-1}(w)), \mu_W(w))$, while $\mu_W^{-1}(w) = h$ and $\theta(\mu_M(h), w, h) = w$, which contradicts that $(\mu_M(h), \mu_W(h)) \in \varphi_W^S(\succsim)$. we conclude that our environment satisfies $\mathcal{E}$.

Finally, we will prove that our environment satisfies constrained unanimity. Suppose there exists $Y^1 \subseteq \mathcal{M}_M$ such that $proj_{\mathcal{M}_M}\{\varphi_W^S(\mathcal{R}\} \subseteq Y^1$ and there exist $\succsim \in \mathcal{R}$ and $(\mu_M, \mu_W) \in Y^1 \times \mathcal{M}_W$ such that for all $K \in \mathcal{N}_0$, $Y^1 \times E(K, \mu_W) \subseteq L_K((\mu_M, \mu_W), \mu_W, \succsim)$, but $(\mu_M, \mu_W) \notin \varphi_W^S(\succsim)$. Again, there must exist $(m, w, h) \in M \times W \times H$ for which it holds that:

$$(w, h)P_m(\mu_W(\mu_M^{-1}(m)), \mu_M(m)),\ (m, h)P_w(\mu_M(\mu_W^{-1}(w)), \mu_W(w)),$$

and either one of the following is true:

(1) $\mu_M^{-1}(m) = h$ and $\theta(m, \mu_W(h), h) = m$

(2) $\mu_W^{-1}(w) = h$ and $\theta(\mu_M(h), w, h) = w$.

(3) $\mu_M(h) = \mu_W(h) = \emptyset$.

Again (1) is rejected by our assumption of strong W-ownership and (2) contradicts Assumption 1. Suppose that (2) is true. Then we have a contradiction of our assumption. Therefore, our environment satisfies constrained unanimity. By Theorem 4, we conclude that φ_W^S is constrained implementable and this completes the proof.

□

The intuition of our positive result is derived from two features: First, the choice of the "natural" rights structure for the fixed sector and second, from the weak stability concept. Indeed, we conjecture that the implementation of ϕ_W^S would be a lot harder or even impossible with a different choice for a fixed rights structure. This also relates to the question of how much freedom coalitions are allowed to have during the matching process, but, most importantly, outlines the importance of well behaved fixed institutions for constrained implementation. Additionally, as shown by Combe (2017), if we strengthen the stability notion, the set of stable matchings might be empty. Such difficulty would carry over to the implementation problem as well.

Finally, a comment on this section vis-à-vis the contribution of Combe (2017), who also presents an algorithm for the implementation of the stable rule under one-sided ownership. It is straightforward to extend his algorithm in our setting. Indeed, fix Γ^W such that any worker can get a project if and only if it is vacant. Now we define Γ^M as follows: (i) Any manager is effective for acquiring a vacant project and, (ii) any manager with the permission of a worker can acquire the worker's project. It is easy to see that these rights structures implement the stable rule under strong manager ownership. However, we have shown a different result, that is, given Γ^W as above, we can implement the stable rule under strong worker ownership.

3.5.2 Multi-issue environment with lexicographic preferences

Setting

As application of our Theorem 6, we will prove the implementability of the weakly Pareto optimal rule, by restricting our attention to an important subclass of preferences, the *lexicographic preferences*[14]. We present the formal definition below:

Definition 3.5.4. R_i is a lexicographic preference if for all $j \in \{1,2\}$, there exists an order R_i^j on X^j and a linear order $\prec_i$ on $\{1,2\}$, such that for any $(x^1,x^2),(y^1,y^2) \in X^1 \times X^2$:

$$(x^1,x^2)R_i(y^1,y^2) \iff x^jR_i^jy^j \text{ for some } j \in \{1,2\} \text{ and for all } k \in \{1,2\} \text{ such that } y^kR_i^kx^k, \text{ there exists } j \prec_i k \text{ with } x^jR_i^jy^j.$$

Let the set of all lexicographic preferences for agent i be $\mathcal{R}_i^{lex} \subseteq \mathcal{R}_i$ and let $\mathcal{R}^{lex} \subseteq \mathcal{R}$ be the set of all possible lexicographic preference profiles. Now, the planner knows that the true profile R is an element of $\mathcal{R}^{lex}$, but she does not know which one. An interpretation is that the planner knows that agents have lexicographic preferences with respect to the issues, but she does not know the priority ordering $\prec_i$ of each agent. Next we define the *Weak Pareto optimal* set.

Definition 3.5.5. For each $R \in \mathcal{R}^{lex}$ and $Z \subseteq X^1 \times X^2$, let $WPO(R) = \{(x^1,x^2) \in X^1 \times X^2 |$ there is no $(y^1,y^2) \in Z$ such that for all $i \in N$, $(y^1,y^2)P_i(x^1,x^2)\}$.

Now, we define the *Weak Pareto optimal rule* as a correspondence $\phi^{WPO} : \mathcal{R}^{lex} \rightrightarrows X^1 \times X^2$, such that for all $R \in \mathcal{R}^{lex}$, $\phi^{WPO}(R) = WPO(R)$. We will show that ϕ^{WPO} is conditionally implementable. First, we establish that our environment satisfies **DR-I**.

Lemma 7. $\mathcal{R}^{lex}$ satisfies **DR-I**.

Proof. Suppose that $(x^1,x^2) \in X^1 \times X^2$ and that $\{(x^1,y^2),(y^1,x^2)\} \subseteq L_i((x^1,x^2),R)$, for some $R \in \mathcal{R}^{lex}$ and $(y^1,y^2) \in X^1 \times X^2$. Suppose that $(y^1,y^2) \notin L_i((x^1,x^2),R)$. Now, $\{(x^1,y^2),(y^1,x^2)\} \subseteq L_i((x^1,x^2),R)$ implies $x^1R_i^1y^1$ and $x^2R_i^2y^2$, which in turn, these two statements imply $(x^1,x^2)R_i(y^1,y^2)$. This however contradicts that $(y^1,y^2)P_i(x^1,x^2)$. Since the previous argument holds for any $(x^1,x^2) \in X^1 \times X^2$, it should also be true for any $(x^1,x^2) \in \phi^{WPO}(R)$. This concludes the proof. □

We have guaranteed that our environment satisfies **DR-I**. Now we show that ϕ^{WPO} satisfies P*-monotonicity and unanimity.

[14]For an early axiomatization of lexicographic preferences, see Fishburn (1975).

Proposition 11. ϕ^{WPO} satisfies P*-monotonicity and unanimity.

Proof. First we show that ϕ^{WPO} satisfies P*-monotonicity with respect to a set Y. Suppose that for all $K \in \mathcal{N}_0$, $[proj_{X^1}\{L_K^{x^2}((x^1, x^2), R)\} \times proj_{X^2}\{L_K^{x^1}((x^1, x^2), R)\}] \cap Y \subseteq L_K((x^1, x^2), R')$, for some $R, R' \in \mathcal{R}^{lex}$ and $(x^1, x^2) \in \phi^{WPO}(R)$. Suppose that $(x^1, x^2) \notin \phi^{WPO}(R')$. We claim that there must exist $(y^1, y^2) \in \phi^{WPO}(R')$, such that for all $j \in N$, $(y^1, y^2)P'_j(x^1, x^2)$. First, notice that since $(x^1, x^2) \notin \phi^{WPO}(R')$, there exists $(z^1, z^2) \in X^1 \times X^2$, such that for all $j \in N$, $(z^1, z^2)P'_j(x^1, x^2)$. If it is the case that $(z^1, z^2) \in \phi^{WPO}(R')$, then we are done. Suppose not. Then, there exists $(w^1, w^2) \in X^1 \times X^2$, such that for all $j \in N$, $(w^1, w^2)P'_j(z^1, z^2)P'_j(x^1, x^2)$. If $(w^1, w^2) \in \phi^{WPO}(R')$, then we are done. If not, then we can continue this reasoning and the existence of such (y^1, y^2) is guaranteed by the finiteness of the sets X^1 and X^2. Now, by our assumption, we have either $(y^1, y^2) \notin Y$ which is rejected since $Y \subseteq \phi(\mathcal{R})$, or $(y^1, y^2) \notin proj_{X^1}\{L_K^{x^2}((x^1, x^2), R)\} \times proj_{X^2}\{L_K^{x^1}((x^1, x^2), R)\}$. This implies either (i) $(y^1, x^2) \notin L_K((x^1, x^2), R)$ or (ii) $(x^1, y^2) \notin L_K((x^1, x^2), R)$. Suppose that (i) is true. Then, we have that for all $i \in N$, $(y^1, x^2)P_i(x^1, x^2)$, which contradicts that $(x^1, x^2) \in \phi^{WPO}(R)$. A similar argument holds for (ii) and this concludes the proof.

Finally, we show that ϕ^{WPO} satisfies unanimity with respect to some set Y such that $\phi^{WPO}(\mathcal{R}) \subseteq Y$. Let $Y \equiv \phi^{WPO}(\mathcal{R})$ and suppose that for all $K \in \mathcal{N}_0$, $Y \subseteq L_K((x^1, x^2), R)$, for some $R \in \mathcal{R}^{lex}$ and $(x^1, x^2) \in Y$. Now assume to the contrary that $(x^1, x^2) \notin \phi^{WPO}(R)$. Then, there exists $(y^1, y^2) \in Y$, such that for all $i \in N$, $(y^1, y^2)P_i(x^1, x^2)$. This however contradicts our premise and this concludes the proof. $\square$

We conjecture that our previous result would also hold in a multi-issue bargaining environment with lexicographic preferences, given that we assume strict monotonicity on the preferences and that $X^1 \subseteq \mathbb{R}$ and $X^2 \subseteq \mathbb{R}$ are closed.

3.6 Discussion of results and conclusion

3.6.1 Discussion

First, a comment with regards to the canonical rights structures that we use in the proofs of Theorems 4 and 6. It is well-known in implementation theory, that the proofs of the sufficiency theorems rely on complicated constructions, with unnatural characteristics, such as integer games or unbounded message spaces. Also, these constructions can be criticized for being too abstract.

Indeed the interpretation of the state space in a canonical rights structure is not very clear[15]. However, the aim of these constructions is not to describe a realistic mechanism design application, rather than characterize what is possible.

With respect to the first critique though, we claim that rights structures, by demanding implementation in a cooperative concept, abstract from the possibly problematic features of noncooperative interaction, especially when the social planner is interested in preventing coalitional deviations. This is particularly evident on the neat characterization theorems that one obtains when considering implementation via rights structures. Indeed, as Aumann (1987) observes:

> (...) when one does build negotiation and enforcement procedures explicitly into the model, then the results of a non-cooperative analysis depend very strongly on the precise form of the procedures, on the order of making offers and counteroffers, and so on. This may be appropriate in voting situations in which precise rules of parliamentary order prevail, where a good strategist can indeed carry the day. But problems of negotiation are usually more amorphous; it is difficult to pin down just what the procedures are. More fundamentally, there is a feeling that procedures are not really all that relevant; that it is the possibilities for coalition forming, promising and threatening that are decisive, rather than whose turn it is to speak. **Aumann (1987)**

3.6.2 Conclusion

We used implementation theory as a tool to study decentralization and tackled two issues in the general framework of rights structures: First, we characterized the set of implementable social choice rules when one sector is fixed and second, we provided some necessary and some sufficient conditions for implementation when both sectors are objects of design. The value of our theorems is also outlined with applications.

Our results show the limitations that decentralization poses when the social goal, as well as agent's preferences, are non-separable across several issues. More importantly, our findings outline that these limitations are robust to some (incomplete) intersectoral communication. This is evident when we compare our conditions with Korpela et al. (2018) for example, where implementation in a centralized environment is considered. This is because of a very fundamental reason: the absence of central authority, capable of monitoring "simultaneous" deviations.

On the other hand though, when we compare our results with Hayashi and Lombardi (2017), our results seem positive. This is because, by allowing for intersectoral communication, we do not demand any sort of decomposability on

[15] Koray and Yildiz (2018) provide a nice interpretation as an analogue of deviation-constrained mechanisms. For a further discussion, see Yildiz (2019).

the social choice rule. In this way, we feel that our results outline the value of conditionality in constitution design and provide a more realistic description of decentralized implementation.

Even though we tried to explore many directions in this paper, there are still various topics that are left for future research. For example, we have only studied implementation with rights structures, while it would be interesting to explore (constrained) conditional implementation with codes of rights. Another fruitful avenue for further research is to identify particular rights structures or mechanisms in decentralized environments that implement desirable rules.

Appendix

Proof of Theorem 4

Proof. Of course, we consider Γ^2 as given. First, suppose that a SCR ϕ satisfies Condition $\mathcal{E}$ and constrained unanimity with respect to a set $Y^1 \supseteq proj_{X^1}\{\phi(\mathcal{R})\}$. Now we define the constrained conditional rights structure $\Gamma^1 = (S^1, h^1, \gamma^1)$ as follows:

State space

Let $T^1 \equiv \{((x^1, x^2), R) \in X^1 \times X^2 \times \mathcal{R} | (x^1, x^2) \in \phi(R)\}$. Then we define $S^1 \equiv T^1 \cup Y^1$.

Outcome function

Let $h^1 : S^1 \times S^2 \to X^1$ be such that for all $(s^1, s^2) \in S^1 \times S^2$,

- If $s^1 = ((x^1, x^2), R)$ and $s^2 \neq \bar{s}^2((x^1, x^2), R) \equiv \bar{s}^2$, then $h^1(s^1, s^2) = y^1 \in \bigcap_{K \in \gamma^2(\bar{s}^2, s^2)} proj_{X^1}\{L_K((x^1, x^2), \bar{s}^2, R)\}$, where y^1 is such that $(y^1, h^2(s^2)) \in \bigcap_{K \in \gamma^2(\bar{s}^2, s^2)} L_K((x^1, x^2), \bar{s}^2, R)$.
- Otherwise, $h^1(((x^1, x^2), R), s^2) = h^1(x^1, s^2) = x^1$.

Effectivity correspondence

Let $\gamma^1 : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$ be such that for all $(s^1, t^1; s^2) \in S^1 \times S^1 \times S^2$:

1. If $s^1 = ((x^1, x^2), R)$ and $s^2 = \bar{s}^2((x^1, x^2), R)$, then for all $K \in \mathcal{N}_0$:

$$K \in \gamma^1(s^1, t^1; s^2) \iff \text{for all } t^2 \in S^2 \text{ such that } K \in \gamma^2(s^2, t^2), \text{ we have } (h^1 \times h^2)(s^1, s^2) R_K (h^1 \times h^2)(t^1, t^2).$$

2. Otherwise, for all $K \in \mathcal{N}_0$, $K \in \gamma^1(s^1, t^1; s^2)$.

In order to show that for any $R \in \mathcal{R}$, $\phi(R) = (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$, we break the proof in two parts:

Part 1: For all $R \in \mathcal{R}$, $\phi(R) \subseteq (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$.

Let $(x^1, x^2) \in \phi(R)$ for some $R \in \mathcal{R}$ and consider $s^1 = ((x^1, x^2), R)$, $s^2 = \bar{s}^2((x^1, x^2), R)$, the existence of the latter guaranteed by Condition $\mathcal{E}$(i). Then, from the outcome function, we have $(h^1(s^1, s^2), h^2(s^2)) = (x^1, x^2)$. Now, for the sake of contradiction, consider $K \in \mathcal{N}_0$ and suppose there exists $(t^1, t^2) \in S^1 \times S^2$, with $(t^1, t^2) \neq (s^1, s^2)$, such that $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2)$ and $(h^1 \times h^2)(t^1, t^2) P_K (x^1, x^2)$. We distinguish the following cases:

1. $t^1 = s^1$, $t^2 \neq s^2$ (K **"moves" only in sector 2**): Then, by the outcome function we have $(h^1 \times h^2)(t^1, t^2) = (y^1, h^2(s^2)) R_K (x^1, x^2)$ for all $K \in \gamma^2(s^2, t^2)$, where the existence of y^1 is guaranteed by Lemma 3. This contradicts our assumption that $(h^1 \times h^2)(t^1, t^2) P_K (x^1, x^2)$.

2. $t^1 \neq s^1$, $t^2 = s^2$ (K **"moves" only in sector 1**): Then, by the effectivity correspondence we have that $(x^1, x^2) R_K (h^1 \times h^2)(t^1, s^2)$, a contradiction.

3. $t^1 \neq s^1$, $t^2 \neq s^2$ (K **"moves" in both sectors**): Again, by the effectivity correspondence it must be that $(x^1, x^2) R_K (h^1 \times h^2)(t^1, t^2)$, a contradiction.

Thus, in all cases, $E(K, (s^1, s^2)) \subseteq L_K((x^1, x^2), R)$ and it holds that $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$. We now proceed to the second part of the proof.

Part 2: For all $R \in \mathcal{R}, (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R) \subseteq \phi(R)$:

Consider $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$. We will show that $(y^1, y^2) \in \phi(R)$. First of all, if $s^1 = ((y^1, y^2), R)$ and $s^2 = \bar{s}^2((y^1, y^2), R)$ there is nothing to prove. We then distinguish the following cases:

1. $s^1 = y^1$: Let $(h^1 \times h^2)(s^1, s^2) = (y^1, y^2)$. By the design of γ^1, we have that for all $K \in \mathcal{N}_0$ and $t^1 \in S^1$, $K \in \gamma^1(s^1, t^1; s^2)$. Therefore, for (s^1, s^2) to be a γ-equilibrium in R it must be that for all $K \in \mathcal{N}_0$, $proj_{X^1}\{\phi(\mathcal{R})\} \times E(K, s^2) \subseteq Y^1 \times E(K, s^2) \subseteq L((y^1, y^2), s^2, R)$ and by constrained unanimity we have $(y^1, y^2) \in \phi(R)$.

2. $s^1 = ((y^1, y^2), R')$. Then we can have two possible subcases:

(a) $s^2 \neq \bar{s}^2((y^1, y^2), R')$: Then, we have $h^1(s^1, s^2) = z^1$, such that $z^1 \in \bigcap_{K \in \gamma^2(\bar{s}^2((y^1,y^2),R),s^2)} proj_{X^1}\{L_K((y^1, y^2), R)\}$. Let $h^2(s^2) = z^2$. Now, by γ^1, for all $K \in \mathcal{N}_0$ and $t^1 \in S^1$, $K \in \gamma^1(s^1, t^1; s^2)$. Thus, as before, it must be that for any $K \in \mathcal{N}_0$, $proj_{X^1}\{\phi(\mathcal{R})\} \times E(K, s^2) \subseteq Y^1 \times E(K, s^2) \subseteq L((y^1, y^2), s^2, R)$. Again constrained unanimity dictates that $(z^1, z^2) \in \phi(R)$.

(b) $s^2 = \bar{s}^2((y^1, y^2), R')$: Then, $(h^1(s^1, s^2), h^2(s^2)) = (y^1, y^2)$. By the design of γ^1, we have that $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2)$ if and only if $(h^1 \times h^2)(t^1, t^2) \in L_K((y^1, y^2), \bar{s}^2, R')|_{Y^1 \times h^2(S^2)}$. For (s^1, s^2) to be a γ-equilibrium, we must have $L_K((y^1, y^2), \bar{s}^2, R')|_{Y^1 \times h^2(S^2)} \subseteq L_K((y^1, y^2), \bar{s}^2, R)$. By condition $\mathcal{E}$(iii) then, we have that $(y^1, y^2) \in \phi(R)$ as well. This concludes the proof of Theorem 4.

□

Proof of Theorem 6

Proof. Consider the following rights structure. We will show that it implements any SCR ϕ that satisfies P-monotonicity and unanimity with respect to a set $Y \subseteq X^1 \times X^2$ under the **DR-I** assumption.

State space

We define $T^1 = \{((x^1, x^2), R) \in Y \times \mathcal{R} | (x^1, x^2) \in \phi(R)\}$. Now let $S^1 \equiv T^1 \cup proj_{X^1}\{Y\}$. Similarly, $T^2 = \{((x^1, x^2), R) \in Y \times \mathcal{R} | (x^1, x^2) \in \phi(R)\}$ and $S^2 \equiv T^2 \cup proj_{X^2}\{Y\}$.

Outcome functions

The outcome functions are as follows:

- $h^1 : S^1 \times S^2 \to X^1$ such that, for any $s^1 \in \{((x^1, x^2), R), x^1\} \subseteq S^1$ and $s^2 \in S^2$, $h^1(s^1, s^2) = h^1(s^1) = x^1$.
- $h^2 : S^2 \times S^1 \to X^2$ such that, for any $s^2 \in \{((x^1, x^2), R), x^2\} \subseteq S^2$ and $s^1 \in S^1$, $h^2(s^1, s^2) = h^2(s^2) = x^2$.

Effectivity correspondences

We define the effectivity correspondences accordingly:

$\gamma^1 : S^1 \times S^1 \times S^2 \rightrightarrows \mathcal{N}$, such that for all $(s^1, t^1; s^2) \in S^1 \times S^1 \times S^2$:

1. If $s^1 = s^2 = ((x^1, x^2), R)$, then for all $K \in \mathcal{N}_0$, $K \in \gamma^1(s^1, t^1; s^2) \iff (h^1(t^1), x^2) \in L_K((x^1, x^2), R)$.

2. In all other cases, for all $K \in \mathcal{N}_0$, $K \in \gamma^1(s^1, t^1; s^2)$.

$\gamma^2 : S^2 \times S^2 \times S^1 \rightrightarrows \mathcal{N}$, such that for all $(s^2, t^2; s^1) \in S^2 \times S^2 \times S^1$:

1. If $s^1 = s^2 = ((x^1, x^2), R)$, then for all $K \in \mathcal{N}_0$, $K \in \gamma^2(s^2, t^2; s^1) \iff (x^1, h^2(t^2)) \in L_K((x^1, x^2), R)$.

2. In all other cases, for all $K \in \mathcal{N}_0$, $K \in \gamma^2(s^2, t^2; s^1)$.

We proceed to show that for all $R \in \mathcal{R}$, $\phi(R) = (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$. We break the proof into two parts:

Part 1: For all $R \in \mathcal{R}$, $\phi(R) \subseteq (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$.

Let the true state be $R \in \mathcal{R}$ and $s^1 = ((x^1, x^2), R) = s^2$. Then, $(h^1 \times h^2)(s^1, s^2) = (x^1, x^2) \in \phi(R)$. Consider $(t^1, t^2) \neq (s^1, s^2)$ and $K \in \gamma^1(s^1, t^1; s^2) \cap \gamma^2(s^2, t^2; s^1)$. Then, if $s^1 \neq t^1$ and $s^2 = t^2$, we have $(h^1(t^1), x^2) \in L_K((x^1, x^2), R)$. If $s^1 = t^1$ and $s^2 \neq t^2$ we have $(x^1, h^2(t^2)) \in L_K((x^1, x^2), R)$. Finally, if $s^1 \neq t^1$ and $s^2 \neq t^2$ we have that $(h^1(t^1), x^2) \in L_K((x^1, x^2), R)$ and $(x^1, h^2(t^2)) \in L_K((x^1, x^2), R)$. By **DR-I** we have that $(h^1(t^1), h^2(t^2)) \in L_K((x^1, x^2), R)$, and we have no possible profitable move by an arbitrary coalition K. Then $(x^1, x^2) \in (h^1 \times h^2) \circ C(\Gamma^1 \times \Gamma^2, R)$.

Part 2: For all $R \in \mathcal{R}$, $(h^1 \times h^2) \circ (\Gamma^1 \times \Gamma^2, R) \subseteq \phi(R)$.

Now, let the true state be $R \in \mathcal{R}$ and consider $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$. We have the following cases:

1. $s^1 = s^2 = ((x^1, x^2), R')$: Then, for (s^1, s^2) to be a γ-equilibrium, it must be that for all $K \in \mathcal{N}_0$, $E(K, (s^1, s^2)) \subseteq L_K((x^1, x^2), R)$. Now, for all $K \in \mathcal{N}_0$, $[proj_{X^1}\{L_K^{x^2}((x^1, x^2), R')\} \times proj_{X^2}\{L_K^{x^1}((x^1, x^2), R')\}] \cap Y \subseteq E(K, (s^1, s^2)) \subseteq L_K((x^1, x^2), R)$. Then, by P-monotonicity we have that $(x^1, x^2) \in \phi(R)$.

2. All other cases: Let $(h^1 \times h^2)(s^1, s^2) = (x^1, x^2)$. By the effectivity correspondence, for all $K \in \mathcal{N}_0$, $E(K, (s^1, s^2)) = Y$ and since $(s^1, s^2) \in C(\Gamma^1 \times \Gamma^2, R)$, we have $Y \subseteq L_K((x^1, x^2), R)$. By unanimity, we must have that $(x^1, x^2) \in \phi(R)$. This completes the proof.

□

Chapter 4

A note on partially honest implementation with rights structures

4.1 Introduction

Motives and behavioural traits have become increasingly prevalent in economic theory and it is not without a good reason. Indeed, they have shown to provide a richer analytical framework and explain a larger set of economic and social phenomena, that classical theory failed to tackle.

There is a substantial literature that explores the effect of behavioural biases or motives in implementation theory[1]. After the seminal contributions of Matsushima (2008) and Dutta and Sen (2012) who introduced a minimal honesty motive to the implementation problem, the theory has explored various equilibrium notions and specifications[2]. However, to the best of our knowledge, the literature has focused only on noncooperative implementation notions, that is, implementation by game forms, or simply mechanisms. This implies that the results rely heavily on the specific equilibrium concept.

Instead, in this note we follow the contribution of Koray and Yildiz (2018) who introduce the notion of a *rights structure*. In this setting, the social planner designs a state space, an outcome function that maps states to social outcomes, and an effectivity correspondence, by which she endows agents or coalitions with

[1]For a survey on the frontier of behavioural implementation and some relevant open questions, see Dutta (2019), as well as other contributions in the same volume.

[2]Just to name a few, Lombardi and Yoshihara (2019) for a full characterization for partially honest Nash implementation, Savva (2018) for strong Nash implementation, Korpela (2014) for Bayesian implementation, Saporiti (2014) for secure implementation, Hagiwara (2019) for double implementation, and Mukherjee et al. (2017) for elimination of undominated strategies.

rights to change the status quo state. In equilibrium, no coalition who has the power to change the state has any incentive to do so. Implementation of a social choice rule is achieved when the set of equilibrium outcomes coincide with the set of socially optimal outcomes that the rule specifies, for any possible preference profile.

The above implementation concept, by being cooperative in nature, admits characterizations that are not very specific to the equilibrium notion and how the game will actually be played. Moreover, the sufficiency theorems do not rely on complicated and unnatural constructions, such as integer and modulo games. Along this new line of research we mention also the contribution of Korpela et al. (2018) who provide a complete characterization of implementation with *codes of rights*[3].

In our study, we provide a sufficient condition for implementation with rights structures, when there exists at least one partially honest agent in the society, i.e. an agent who prefers to tell the truth, when the welfare she derives from the outcome is not at stake. Our condition is the well-known *unanimity* condition. In this way, we provide an analogue of the contribution of Dutta and Sen (2012) in the rights structure framework. In the next section we present our formal environment, while section 4.3 contains our results. Section 4.4 concludes.

4.2 Model

The society consists of a set of agents $N = \{1, ..., n\}$, where $n \geq 2$ and a (finite) set of social outcomes X. Each agent i is endowed with a weak preference relation R_i on X, where P_i and I_i is its strict and symmetric part respectively. The set of all possible preferences for each i is denoted $\mathcal{R}_i$. An n-tuple $R = (R_1, ..., R_n) \in \mathcal{R}_1 \times ... \times \mathcal{R}_n \equiv \mathcal{R}$, is called a preference profile.

A *social choice rule* ϕ is a correspondence $\phi : \mathcal{R} \rightrightarrows X$, such that for any R, $\phi(R) \subseteq X$ is nonempty. The image of ϕ is denoted by $\phi(\mathcal{R})$.

A *rights structure* is a triplet $\Gamma = (S, h, \gamma)$ is such that S is a state space, $h : S \to X$ is an outcome function that maps states to outcomes, and $\gamma : S \times S \rightrightarrows N$ is an effectivity correspondence such that, for any $(s, t) \in S \times S$, $\gamma(s, t)$ is the set of agents who *individually*[4] are effective to change the state from s to t. Finally, let $h(S) \equiv \{x \in X |$ there exists $s \in S$, $h(s) = x\}$ be the range of the rights structure.

In our setting, states contain more information than merely an outcome. It is

[3] A codes of rights is a rights structure where the state space is the set of outcomes and the outcome function is the identity map.

[4] Indeed, an effectivity correspondence can be defined more generally for coalitions. In our setting we consider *individual-based* rights structures. This is done without loss of generality for our weak-core solution concept as shown by Korpela et al. (2018).

this exact feature of the rights structure which the social planner can use as leverage to exploit the partial honesty motive. For this purpose, we restrict attention to rights structures $\Gamma = (S, h, \gamma)$, such that $S \subseteq X \times \mathcal{R}$. A few further definitions are in order:

Given a rights structure Γ, we define the *truth correspondence* $T^{\Gamma} : \mathcal{R} \rightrightarrows S$, such that, for any profile $R \in \mathcal{R}$, $T^{\Gamma}(R) = X \times \{R\}$. Now, given a rights structure Γ and a preference profile R, for any agent i we define $\succsim_i^R$ as a binary, reflexive and transitive relation on $S \times S$, as follows:

Definition 4.2.1. An agent $i \in N$ is *partially honest* if for all $s, t \in S$ and $R \in \mathcal{R}$, the following are true:

(i) If $h(s) I_i h(t)$, $s \in T^{\Gamma}(R)$ and $t \notin T^{\Gamma}(R)$, then $s \succ_i^R t$

(ii) Otherwise, $h(s) R_i h(t) \iff s \succsim_i^R t$.

An agent $i \in N$ is not partially honest if for all $s, t \in S$ and $R \in \mathcal{R}$, $h(s) R_i h(t) \iff s \succsim_i^R t$.

From now on, we make the following assumption:

Assumption: There exists at least one partially honest in agent in N.

Even though in this setting the social planner knows about the existence of a partially honest agent, she does not know who he is. Now, as before, we write $\succsim^R = (\succsim_1^R, ..., \succsim_n^R)$, for a profile of the agents' preferences on states, induced by R.

Definition 4.2.2. Given a rights structure $\Gamma = (S, h, \gamma)$ and a preference profile $R \in \mathcal{R}$, a state $s \in S$ is a γ-equilibrium in R, if for all $t \in S$ and $i \in N$ such that $\{i\} \in \gamma(s, t)$, we have $s \succsim_i^R t$.

Let $C(\Gamma, \succsim^R)$ be the set of γ-equilibrium states in R and $h \circ C(\Gamma, \succsim^R)$ the set of outcomes that correspond to the γ-equilibrium states. Then, a rights structure $\Gamma = (S, h, g)$ implements the social choice rule ϕ in γ-equilibrium, if for all $R \in \mathcal{R}$, $\phi(R) = h \circ C(\Gamma, \succsim^R)$.

The interpretation of our setting is as follows: The ethical concerns of the society are represented by a (fixed) social choice rule, which prescribes the outcomes that are considered acceptable for each possible preference profile of the agents. The social planner desires the realization of socially optimal outcomes, but the problem is that she does not know the actual preference profile of the agents. Therefore she designs a rights structure Γ such that, in the γ-equilibrium, socially optimal outcomes, and only those, are realized, for any possible preference profile. In this case we have an implementation of the social choice rule.

Additionally, in our setting, the rights structure is constructed such that each state consists of two parts: (i) an outcome and (ii) a preference profile (which of course has no direct relationship with the true preference profile). A broader interpretation due to Koray and Yildiz (2018) is that the preference profile part in the state represents the context or frame, with which each outcome is supported by.

To give an example of that, consider a bargaining situation, where for example outcome x is under consideration. Then, the social planner can design the bargaining protocol such that an agent might be entitled to propose outcome y from x, only if a certain context applies. For instance, an agent may be entitled to propose an outcome y when x is under consideration, only if they "claim" they prefer x to y. The claim is the context in our setting. This particular scheme resembles Maskin's canonical mechanism and we will utilize it in the proof of our theorem.

A partially honest agent then is one that prefers "truly" supporting contexts. Driven by a sense of honesty or consistency, a partially honest agent has preferences over the states and specifically on the part that corresponds to the preference profile. So, when a partially honest agent contemplates whether to deviate from a status quo to a new state, first she assesses whether the move is profitable with respect to the outcomes and, if and only if she indifferent between the outcomes, she strongly prefers to move to the new state when its preference profile part corresponds to the true preference profile, while this is not true for the status quo.

To explain the permissiveness of partial honesty in the implementation problem more intuitively, suppose first that there are no partially honest agents. Then, suppose that state (x, R') is an equilibrium in preference profile R. Now, if the rights structure implements the SCR ϕ, we must have that $x \in \phi(R')$. Consider now the existence of a partially honest agent. Then, given that the preference profile part of the state (x, R') is different than the true one, a partially honest agent may wish to deviate to another state. We are therefore not required to have $x \in \phi(R)$ in this case.

To clarify the above discussion, consider the following example. We have two possible outcomes $X = \{x, y\}$, two agents $N = \{1, 2\}$ and two possible preference profiles $\mathcal{R} = \{R, R'\}$ as follows:

R		R'	
1	2	1	2
x	y	x	xy
y	x	y	

Table 4.1: Example, Preferences

Suppose that the SCR ϕ is such that $\phi(R) = \{x, y\}$, while $\phi(R') = \{x\}$ and consider the following rights structure:

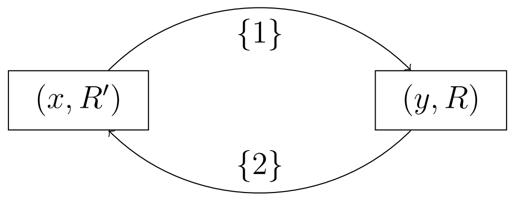

Figure 4.1: Example, rights structure

The state space is $S = \{(x, R'), (y, R)\}$. Let $h((x, R')) = x$ and $h((y, R)) = y$. Now notice that both (x, R') and (y, R) are γ-equilibria in preference profile R, as neither agent 1 wants to move to (y, R), nor 2 would like to move to (x, R'). Let us now focus on profile R'. Suppose that no agent is partially honest. First, see that (x, R') is a γ-equilibrium. However, (y, R) is also a γ-equilibrium in R' and implementation fails, as $y \notin \phi(R')$. Now suppose that 2 is partially honest. Then, since in R' he is indifferent between x and y, he would like to reveal the true state and move to (x, R'), thus eliminating (y, R) from the equilibria. Thus, implementation is restored.

4.3 Results

Before we proceed with our main result, we present the result of Korpela et al. (2018), who provide a complete characterization of the implementable social choice rules without any behavioural assumptions. The two relevant conditions are *Maskin-monotonicity* and *unanimity* and we state them below:

Definition 4.3.1. A SCR ϕ satisfies *Maskin-monotonicity* with respect to Y, if there exists a set $Y \subseteq X$ with $Y \supseteq \phi(\mathcal{R})$ such that, for all $R, R' \in \mathcal{R}$ and $x \in \phi(R)$,

$$[\text{ for all } i \in N,\ L_i(x, R) \cap Y \subseteq L_i(x, R')] \Rightarrow x \in \phi(R').$$

Definition 4.3.2. A SCR ϕ satisfies *unanimity* with respect to Y, if there exists a set $Y \subseteq X$ with $Y \supseteq \phi(\mathcal{R})$ such that, for all $R \in \mathcal{R}$ and $x \in Y$, if for all $i \in N$ we have $Y \subseteq L_i(x, R)$, then $x \in \phi(R)$.

Theorem 8 (Korpela et al. (2018))**.** A SCR ϕ is implementable if and only if it satisfies Maskin-monotonicity and unanimity.

In our result, we are able to dispose of Maskin-monotonicity as a sufficient condition. Instead we show that unanimity alone is sufficient:

Theorem 9. A SCR ϕ is implementable with partially honest agents if it satisfies unanimity.

Proof. Suppose that ϕ satisfies unanimity with respect to Y and consider the following rights structure:

- $S = \{(x, R) \in Y \times \mathcal{R}\}$.
- $h : S \to X$, such that for all $s = (x, R) \in S$, $h(s) = x$.
- $\gamma : S \times S \rightrightarrows N$, such that, for all $(s, t) \in S \times S$, where $s = (x, R)$:

  (i) If $x \in \phi(R)$, then for all $i \in N$ and $t \in S$, $\{i\} \in \gamma(s, t) \iff xR_ih(t)$.

  (ii) Otherwise, for all $i \in N$ and $t \in S$, $\{i\} \in \gamma(s, t)$.

We will show that Γ implements ϕ. We break the proof into two parts:

Part 1: For all $R \in \mathcal{R}$, $\phi(R) \subseteq h \circ C(\Gamma, \succeq^R)$:

Let $x \in \phi(R)$, for some $R \in \mathcal{R}$. Consider $s = (x, R)$. Then for all $t \in S$ with $h(t) \neq x$ and $i \in N$, $\{i\} \in \gamma(s, t)$ if and only if $xR_ih(t)$, thus $s \in C(\Gamma, \succeq^R)$ and clearly, $x \in h \circ C(\Gamma, \succeq^R)$.

Part 2: For all $R \in \mathcal{R}$, $h \circ C(\Gamma, R) \subseteq \phi(R)$:

We will prove this part in two steps. First, we will show that there cannot exist $s \in C(\Gamma, \succsim^R)$, such that $s = (x, R')$, where $R' \neq R$. Indeed, suppose otherwise. Then, any partially honest agent can move to $t = (x, R)$, since $t \in T^\Gamma(R)$ and $s \notin T^\Gamma(R)$. So, for all $s \in C(\Gamma, \succsim^R)$, $s = (y, R)$.

Second, we will show that there cannot exist $s \in C(\Gamma, \succsim^R)$, such that $s = (y, R)$ where $y \notin \phi(R)$. Suppose that this is the case. Then, any agent is entitled to change s to any state t. So, it must be that for all $i \in N$, $Y \subseteq L_i(y, R)$. This fulfills the premises of unanimity, so it must be that $y \in \phi(R)$ and we have a contradiction. So, we have that $s \in C(\Gamma, R)$ if and only if $s = (y, R)$, with $y \in \phi(R)$. This completes the proof.

□

4.4 Comments and conclusion

In the canonical rights structure constructions of Koray and Yildiz (2018) and Korpela et al. (2018), states are either of the form $s = (x, R)$ where $x \in \phi(R)$, or simply $s = x$. In the interpretation of Yildiz (2019), this means that there either every outcome is supported by the appropriate "context" or frame, or there is no context. In our canonical rights structure, we utlized inappropriate contexts to eliminate outcomes that are not socially optimal, as Dutta and Sen (2012) do in the case of Nash implementation with noncooperative game forms. Most importantly though, our permissive results are obtained with a very weak assumption on the number of partially honest agents, contrary to most of the literature in implementation with partial honesty, Dutta and Sen (2012) excluded, where in order to get the most positive results, extreme assumptions about the number of partially honest agents are made.

Our paper carries the positive results on noncooperative implementation with motives to the cooperative implementation framework. We feel that our results outline the merits of the cooperative approach in implementation theory. Nevertheless, our study was only one of the possible ways of incorporating motives into a cooperative game-theoretic framework and it would be interesting to explore other possible ways.

Chapter 5

Bibliography

Abreu, D. and Sen, A. (1990), 'Subgame perfect implementation: A necessary and almost sufficient condition', *Journal of Economic theory* **50**(2), 285–299.

Adachi, T. (2017), 'Nash implementation with honesty and ranges of honesty'. mimeo.

Aumann, R. J. (1960), 'Acceptable points in games of perfect information.', *Pacific J. Math.* **10**(2), 381–417.

Aumann, R. J. (1987), 'Game theory the new palgrave: A dictionary of economics. edited by john eatwell, murray milgate and peter newman'.

Bierbrauer, F. and Netzer, N. (2016), 'Mechanism design and intentions', *Journal of Economic Theory* **163**, 557–603.

Breton, M. L. and Sen, A. (1999), 'Separable preferences, strategyproofness, and decomposability', *Econometrica* **67**(3), 605–628.

Combe, J. (2017), 'Matching with ownership', *Available at SSRN 3071879* .

Corchón, L. C. (2015), 'The theory of implementation. what did we learn?', *Available at SSRN: https://ssrn.com/abstract=2690301* .

Corchón, L. C. and Herrero, C. (2004), 'A decent proposal', *Spanish Economic Review* **6**(2), 107–125.

Deb, R., Pattanaik, P. K. and Razzolini, L. (1997), 'Game forms, rights, and the efficiency of social outcomes', *Journal of Economic Theory* **72**(1), 74 – 95.

Doghmi, A. and Ziad, A. (2013), 'On partially honest nash implementation in private good economies with restricted domains: A sufficient condition', *The BE Journal of Theoretical Economics* **13**(1), 415–428.

Doğan, B. (2017), 'Eliciting the socially optimal allocation from responsible agents', *Journal of Mathematical Economics* **73**, 103 – 110.

Dutta, B. (2019), Complete information implementation and behavioral economics, *in* 'The Future of Economic Design', Springer, pp. 317–319.

Dutta, B. and Sen, A. (1991), 'Implementation under strong equilibrium: A complete characterization', *Journal of Mathematical Economics* **20**(1), 49–67.

Dutta, B. and Sen, A. (2012), 'Nash implementation with partially honest individuals', *Games and Economic Behavior* **74**(1), 154–169.

Eliaz, K. (2002), 'Fault tolerant implementation', *The Review of Economic Studies* **69**(3), 589–610.

Fishburn, P. C. (1975), 'Axioms for lexicographic preferences', *The Review of Economic Studies* **42**(3), 415–419.

Gärdenfors, P. (1981), 'Rights, games and social choice', *Nous* **15**(3), 341–356.

Gibbard, A. (1973), 'Manipulation of voting schemes: a general result', *Econometrica* **41**(4), 587–601.

Gibbard, A. (1974), 'A pareto-consistent libertarian claim', *Journal of Economic Theory* **7**(4), 388–410.

Glazer, J. and Rubinstein, A. (1998), 'Motives and implementation: On the design of mechanisms to elicit opinions', *Journal of Economic Theory* **79**(2), 157–173.

Greenberg, J. (1990), *The theory of social situations: an alternative game-theoretic approach*, Cambridge University Press.

Hagiwara, M. (2019), 'Double implementation without no-veto-power', *Mathematical Social Sciences* **101**, 124–130.

Hagiwara, M., Yamamura, H. and Yamato, T. (2017), 'Implementation with socially responsible agents', *Economic Theory Bulletin* pp. 1–8.

Hammond, P. (1996), Game forms versus social choice rules as models of rights, *in* 'Social choice re-examined', Springer, pp. 82–95.

Hayashi, T. and Lombardi, M. (2017), 'Implementation in partial equilibrium', *Journal of Economic Theory* **169**, 13–34.

Hayashi, T. and Lombardi, M. (2019), 'Constrained implementation', *Journal of Economic Theory* **183**, 546–567.

Hayek, F. A. (1980), *Individualism and economic order*, University of chicago Press.

Hurkens, S. and Kartik, N. (2009), 'Would i lie to you? on social preferences and lying aversion', *Experimental Economics* **12**(2), 180–192.

Hurwicz, L. (1960), 'Optimality and informational efficiency in resource allocation processes', *Mathematical methods in the social sciences* .

Hurwicz, L. (1972), 'On informationally decentralized systems', *Decision and organization: A volume in Honor of J. Marschak* .

Jackson, M. O. (1991), 'Bayesian implementation', *Econometrica: Journal of the Econometric Society* pp. 461–477.

Jackson, M. O. (1992), 'Implementation in undominated strategies: A look at bounded mechanisms', *The Review of Economic Studies* **59**(4), 757–775.

Jackson, M. O. (2001), 'A crash course in implementation theory', *Social choice and welfare* **18**(4), 655–708.

Jackson, M. O., Palfrey, T. R. and Srivastava, S. (1994), 'Undominated nash implementation in bounded mechanisms', *Games and Economic Behavior* **6**(3), 474–501.

Kartik, N. and Tercieux, O. (2012), 'Implementation with evidence', *Theoretical Economics* **7**(2), 323–355.

Kartik, N., Tercieux, O. and Holden, R. (2014), 'Simple mechanisms and preferences for honesty', *Games and Economic Behavior* **83**, 284–290.

Kimya, M. (2017), 'Nash implementation and tie-breaking rules', *Games and Economic Behavior* **102**, 138–146.

Koray, S. and Yildiz, K. (2018), 'Implementation via rights structures', *Journal of Economic Theory* **176**, 479–502.

Korpela, V. (2013), 'A simple sufficient condition for strong implementation', *Journal of Economic Theory* **148**(5), 2183–2193.

Korpela, V. (2014), 'Bayesian implementation with partially honest individuals', *Social Choice and Welfare* **43**(3), 647–658.

Korpela, V., Lombardi, M. and Vartiainen, H. (2018), 'Do coalitions matter in designing institutions?', *Available at SSRN 3305209* .

Lombardi, M. and Yoshihara, N. (2017), 'Natural implementation with semi-responsible agents in pure exchange economies', *International Journal of Game Theory* pp. 1–22.

Lombardi, M. and Yoshihara, N. (2018), 'Treading a fine line: (im)possibilities for nash implementation with partially-honest individuals', *Games and Economic Behavior* **111**, 203 – 216.

Lombardi, M. and Yoshihara, N. (2019), 'Partially-honest nash implementation: a full characterization', *Economic Theory* pp. 1–34.

Maskin, E. (1979), Implementation and strong nash equilibrium, *in* J. J. Laffont, ed., 'Aggregation and Revelation of Preferences', North Holland, pp. 433–440.

Maskin, E. (1999), 'Nash equilibrium and welfare optimality', *The Review of Economic Studies* **66**(1), 23–38.

Maskin, E. and Sjöström, T. (2002), 'Implementation theory', *Handbook of social Choice and Welfare* **1**, 237–288.

Matsushima, H. (2008), 'Role of honesty in full implementation', *Journal of Economic Theory* **139**(1), 353–359.

McQuillin, B. and Sugden, R. (2011), 'The representation of alienable and inalienable rights: games in transition function form', *Social Choice and Welfare* **37**(4), 683.

Mezzetti, C. and Renou, L. (2017), 'Repeated nash implementation', *Theoretical Economics* **12**(1), 249–285.

Miyagawa, E. (2002), 'Reduced-form implementation'.

Mookherjee, D. (2006), 'Decentralization, hierarchies, and incentives: A mechanism design perspective', *Journal of Economic Literature* **44**(2), 367–390.

Moore, J. and Repullo, R. (1990), 'Nash implementation: a full characterization', *Econometrica: Journal of the Econometric Society* pp. 1083–1099.

Moulin, H. and Peleg, B. (1982), 'Cores of effectivity functions and implementation theory', *Journal of Mathematical Economics* **10**(1), 115–145.

Mukherjee, S., Muto, N. and Ramaekers, E. (2017), 'Implementation in undominated strategies with partially honest agents', *Games and Economic Behavior* **104**, 613 – 631.

Nicolo, A., Sen, A. and Yadav, S. (2019), 'Matching with partners and projects', *Journal of Economic Theory* **184**, 104942.

Ortner, J. (2015), 'Direct implementation with minimally honest individuals', *Games and Economic Behavior* **90**, 1–16.

Palfrey, T. R. (2002), 'Implementation theory', *Handbook of game theory with economic applications* **3**, 2271–2326.

Peleg, B. (1997), 'Effectivity functions, game forms, games, and rights', *Social choice and welfare* **15**(1), 67–80.

Peleg, B. and Winter, E. (2002), 'Constitutional implementation', *Review of Economic Design* **7**(2), 187–204.

Roemer, J. E. (1988), 'Axiomatic bargaining theory on economic environments', *Journal of Economic Theory* **45**(1), 1–31.

Saporiti, A. (2014), 'Securely implementable social choice rules with partially honest agents', *Journal of Economic Theory* **154**, 216–228.

Satterthwaite, M. A. (1975), 'Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions', *Journal of economic theory* **10**(2), 187–217.

Savva, F. (2018), 'Strong implementation with partially honest individuals', *Journal of Mathematical Economics* **78**, 27–34.

Sen, A. (1970), 'The impossibility of a paretian liberal', *Journal of political economy* **78**(1), 152–157.

Sertel, M. R. (2002), Designing rights: Invisible hand theorems, covering and membership, Technical report, mimeo, 2001, Bogazici University.

Shin, S. and Suh, S.-C. (1996), 'A mechanism implementing the stable rule in marriage problems', *Economics Letters* **51**(2), 185–189.

Suh, S.-C. (1996), 'Implementation with coalition formation: A complete characterization', *Journal of Mathematical Economics* **26**(4), 409–428.

Suh, S.-C. (1997), 'Double implementation in nash and strong nash equilibria', *Social Choice and Welfare* **14**(3), 439–447.

Tadenuma, K. and Toda, M. (1998), 'Implementable stable solutions to pure matching problems', *Mathematical Social Sciences* **35**(2), 121–132.

Vartiainen, H. (2007*a*), 'Nash implementation and the bargaining problem', *Social Choice and Welfare* **29**(2), 333–351.

Vartiainen, H. (2007*b*), 'Subgame perfect implementation: a full characterization', *Journal of Economic Theory* **133**(1), 111–126.

Yildiz, K. (2019), Implementation with boundedly rational agents, *in* 'The Future of Economic Design', Springer, pp. 359–363.