AN IMPOSSIBILITY RESULT REGARDING BEHAVIORAL IMPLEMENTATION OF EFFICIENCY WITH TWO INDIVIDUALS

A Master's Thesis

by
ÖMER UÇKAÇ

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
December 2023

To Sıla

AN IMPOSSIBILITY RESULT REGARDING BEHAVIORAL IMPLEMENTATION OF EFFICIENCY WITH TWO INDIVIDUALS

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

ÖMER UÇKAÇ

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN ECONOMICS

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

December 2023

AN IMPOSSIBILITY RESULT REGARDING BEHAVIORAL IMPLEMENTATION OF EFFICIENCY WITH TWO INDIVIDUALS
By Ömer Uçkaç

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Nuh Aygün Dalkıran
Advisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Tarık Kara
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Serkan Küçükşenel
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

Refet S. Gürkaynak
Director

ABSTRACT

AN IMPOSSIBILITY RESULT REGARDING BEHAVIORAL IMPLEMENTATION OF EFFICIENCY WITH TWO INDIVIDUALS

Uçkaç, Ömer
M.A., Department of Economics
Supervisor: Assist. Prof. Dr. Nuh Aygün Dalkıran

December 2023

This thesis examines Nash implementation of behaviorally efficient social choice rules á la de Clippel (2014) with two individuals under the full behavioral domain, i.e., when individuals' choices do not satisfy the weak axiom of revealed preferences. We propose a new definition of a dictatorial social choice rule in the full behavioral domain and show that when there are at least four alternatives, a behaviorally efficient social choice rule á la de Clippel (2014) is implementable if and only if it is dictatorial according to our definition whenever there are only two individuals under consideration. Our result parallels the impossibility result of Maskin (1999), which says that in the full rational domain, a social choice rule that satisfies the Pareto property is implementable if and only if it is dictatorial whenever there are only two individuals in the society.

Keywords: Nash Implementation, Behavioral Implementation, Behavioral Efficiency, Dictatorship, Two Individuals

ÖZET

ETKİNLİĞİN İKİ BİREYLE DAVRANIŞSAL OLARAK UYGULANMASINA İLİŞKİN BİR İMKÂNSIZLIK SONUCU

Uçkaç, Ömer

Yüksek Lisans, İktisat Bölümü

Tez Danışmanı: Dr. Nuh Aygün Dalkıran

Aralık 2023

Bu tez, iki bireyli bir toplumda tam davranışsal alanda, yani bireylerin seçimlerinin açığa vurulan tercihlerin zayıf aksiyomunu sağlamadığı durumlarda, de Clippel (2014) tipi davranışsal etkin sosyal seçim kurallarının Nash uygulamasını incelemektedir. Bu tip alanlarda diktatöryallik için yeni bir tanım önerdikten sonra, en az dört alternatif olduğunda, iki bireyli bir toplum için de Clippel (2014) tipi davranışsal etkin bir sosyal seçim kuralının ancak ve ancak tanımımıza göre diktatöryal ise uygulanabilir olduğunu göstermekteyiz. Sonucumuz Maskin (1999)'in tam rasyonel alanda gösterdiği, iki bireyli toplumlarda Pareto özelliği olan sosyal seçim kurallarının ancak ve ancak diktatöryal ise uygulanabilirdir, imkansızlık sonucuyla paralellik göstermektedir.

Anahtar Kelimeler: Nash Uygulaması, Davranışsal Uygulama, Davranışsal Etkinlik, Diktatörlük, İki Birey

ACKNOWLEDGMENTS

I wish to express my gratitude to my advisor, Nuh Aygün Dalkıran, for his enduring patience and support.

I extend my thanks to the committee members, Tarık Kara and Serkan Küçükşenel, for dedicating their time to assess and evaluate this thesis.

I would also like to acknowledge the support and encouragement provided by my beloved Sıla, sisters Leyla and Sultan, and friends Samed and Kutay during this academic endeavor. Their understanding and encouragement have been crucial in navigating the challenges of graduate studies.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Economics accords significant importance to social planners, who, in economic models, are typically assumed to act in the best interest of society and are concerned with optimizing social welfare or achieving certain societal objectives. Social planners are endowed with the authority to make policy decisions and design mechanisms that influence the well-being of individuals in society. This is where mechanism design comes into play, scrutinizing individuals' choices and assessing the conditions that determine the feasibility of achieving desired outcomes. Implementation theory primarily aims to unveil the pertinent theoretical conditions governing the implementability of social choice rules (SCRs).

One of the seminal papers in the implementation theory literature is Maskin (1999)[circulated since 1977]. Even though Maskin (1999) is well-known for the necessity and sufficiency results that it presents for Nash implementation when the society consists of three or more individuals, this thesis is rather related to the impossibility result Maskin (1999) provides for the case of two individuals.

For the case of two individuals, Maskin (1999) identifies the following negative result: Whenever there are only two individuals in the society, an SCR that satisfies the Pareto property is implementable in Nash equilibrium in the full rational domain if and only if it is dictatorial. This result is interpreted as an impossibility result as it implies that implementing a non-dictatorial SCR in Nash equilibrium is impossible in a society with two individuals when one considers the full rational domain. Another interpretation of this result is that social planners must favor one of the individuals when they intend to implement an SCR in Nash equilibrium under the full rational domain. This does not seem reasonable from a social perspective. This is not surprising since there are plenty of impossibility results in social choice theory, the most popular being Arrow's impossibility result, which states that in the full rational domain, a non-dictatorial, Pareto-optimal voting scheme that satisfies independence of irrelevant alternatives does not exist.

Along the similar lines of the impossibility result of Maskin (1999), Hurwicz and Schmeidler (1978) shows that if a two-individual, Pareto-optimal, social choice rule is defined on the domain of all possible strict preferences is implementable, then it must be dictatorial.

de Clippel (2014) is the paper that dubbed the term *behavioral implementation* to refer to the implementation of social choice rules when individuals' choices fail to satisfy the key axiom behind rationality: Weak Axiom of Revealed Preferences (WARP). When individuals' choices fail WARP, it is not possible to model individuals as if their choices follow the maximization of a well-defined context-independent payoff function. de Clippel (2014) extends Maskin (1999)'s results on Nash Implementation to behavioral environments: He identifies that the existence of a collection of sets that is consistent with an SCR is necessary for Nash implementation of this SCR. Furthermore, he shows that an SCR that respects unanimity is Nash implementable when there is a collection of sets that satisfy a slightly stronger consistency requirement for this SCR whenever there are three or more individuals in the society. de Clippel does not provide any sufficiency results for the case of two individuals. However, he argues that an analysis à la Moore and Repullo (1990) would provide similar results for Nash implementation in behavioral domains when there are only two individuals.

de Clippel (2014) also provides an efficiency definition for behavioral domains as follows: An alternative is behaviorally efficient in some state of the world if, for each individual, there is a set from which the individual chooses this alternative, and the union of all individuals' sets constitutes the set of all alternatives. de Clippel compares and contrasts this notion of efficiency with the generalized Pareto efficiency defined by Bernheim and Rangel (2009) for behavioral domains that allow for non-rational choices. de Clippel (2014) shows that his behavioral efficiency strictly refines that of Bernheim and Rangel (2009), i.e., if an alternative is behaviorally efficient á la de Clippel (2014), then it is generalized Pareto efficient à la Bernheim and Rangel (2009), but the converse is not true in general.

In this thesis, we focus on Nash implementation in behavioral domains with two individuals and try to understand when the impossibility result of Maskin (1999) extends to behavioral domains, taking into account the behavioral efficiency notion defined by de Clippel (2014). In the rational domain, behavioral efficiency à la de Clippel (2014) boils down to Pareto efficiency. Therefore, if one considers a behavioral domain that contains all the choices associated with the full rational domain, then Maskin (1999)'s impossibility result for two-individual Nash implementation extends to this domain. However, if the planner is certain that individuals' choices do not satisfy WARP, then it is not clear when and how Maskin's negative result on Nash implementation with two individuals extends to behavioral domains.

To this end, we focus on the domain of choices where individuals' choices do not satisfy WARP. In particular, we assume that in every state of the world, there is at least one individual whose choices cannot be rationalized as maximization of context-independent payoff functions. We refer to the set of all choice profiles where this holds as the *full behavioral domain*.

We propose a definition of dictatorial social choice rules in the full behavioral domain. Then, we show that when there are at least four alternatives, a behaviorally efficient social choice rule á la de Clippel (2014) is Nash implementable if and only if it is dictatorial according to our definition whenever there are only two individuals under consideration. This parallels the impossibility result of Maskin (1999).

The rest of this thesis is organized as follows: Chapter 2 presents the literature review. Chapter 3 describes the model, sets up the preliminaries, and presents our main result; Chapter 4 concludes.

CHAPTER 2

LITERATURE REVIEW

In this section, we provide a brief literature review for implementation theory, focusing rather on the recent work on behavioral implementation. As we mentioned above, the first paper that comes to mind in implementation theory literature is the seminal work of Maskin (1999), which originated from Maskin's Ph.D. dissertation submitted in 1976. Some of the results in that study, including the impossibility result we consider in this thesis, were circulated since 1977. Maskin (1999) considers an environment of complete information under the full rational domain, i.e., an environment where the preferences of each individual can be represented with a rational preference relation that is complete and transitive and the preference profile of the society is common knowledge between the individuals in the society. Yet, the true preference profile is not known to the planner, who believes that any possible rational preference profile is possible for the society under consideration. Therefore, this is an environment where WARP holds for every individual in every state of the world.

Maskin identifies a necessary condition for the implementation of an SCR in Nash equilibrium, which is dubbed as Maskin-monotonicity. Moreover, he identifies that monotonicity together with no-veto power is sufficient for Nash implementation whenever there are three or more individuals in the society. As mentioned above, for the case of two individuals, Maskin (1999) identifies the following impossibility result: In a society with two individuals, a non-dictatorial SCR that satisfies the Pareto property cannot be implemented in Nash equilibrium in the full rational domain.

Moore and Repullo (1990) considers the problem of Nash implementation with three or more individuals and exactly two individuals under rationality. They show that the conditions they introduce, namely Condition μ and Condition $\mu 2$, are both necessary and sufficient for Nash implementation of an SCR for the case of three or more individuals and the case of two individuals, respectively. Sjöström (1991) also explores Nash implementability of SCRs and presents conditions M and $M2$ and shows that these conditions are equivalent to Moore and Repullo's Conditions μ and $\mu 2$, respectively.

Independently from Moore and Repullo (1990), Dutta and Sen (1991) analyzes the Nash implementation problem with two individuals in the rational domain and characterizes all Nash-implementable two-individual social choice rules. They use their characterization to identify domain restrictions to allow for Nash implementability of non-dictatorial social choice rules that satisfy the Pareto property. The key observation of Dutta and Sen (1991) is that Maskin-monotonicity coupled with no-veto power conditions can be used by the social planner to identify a whistle-blower from the message profile of others that is associated with truthful revelation when there are three or more individuals in the society; however, it is not possible to identify the whistle-blower from the other individual's message when there are only two individuals. In particular, Dutta and Sen (1991) introduces Condition β and demonstrates that this condition is both necessary and sufficient for an SCR to be implementable in the rational domain whenever there are only two individuals under consideration.

Following these works, there has been a huge literature on Nash Implementation under the rational domain. We refer the interested reader for more on Implementation under the rational domain to the following surveys: Jackson (2001), Palfrey (2002), Maskin and Sjöström (2002), and Serrano (2004).

We now turn to the rather recent literature on *behavioral implementation*, which considers the implementation problem under the behavioral domains that allow for non-rational choices, i.e., individuals' choices are allowed to fail WARP. To the best of our knowledge, Hurwicz (1986) is the first paper that considers the implementation problem for societies where the aggregated social preferences of rational individuals can lead to non-rational choices. Hurwicz refers to such societies as irrational societies.

Korpela (2012) points out that the behavioral equilibrium concept employed by Hurwicz (1986) is not adequate for studying the implementation problem without rationality axioms. After introducing the behavioral Nash equilibrium as a solution concept for behavioral domains, Korpela shows that the results of Moore and Repullo (1990) extend to behavioral domains when individuals' choices satisfy the independence of irrelevant alternatives axioms, which is also known as Sen (1971)'s α or Chernoff (1954)'s condition. With a slight abuse, we refer to the concept of behavioral Nash equilibrium introduced by Korpela simply as Nash equilibrium in this thesis.

Independently from Korpela (2012), de Clippel (2014) also considers the implementation problem when individuals' choices cannot be represented with context-independent preferences. de Clippel (2014) is the study that dubs the literature on implementation theory without rationality axioms as the behavioral implementation literature. As we mentioned before, de Clippel (2014) broadens Maskin (1999)'s seminal results on Nash implementation to behavioral domains and provides a notion of efficiency for behavioral domains. The efficiency notion de Clippel introduces is the maximal extension of Pareto efficiency that is implementable in Nash equilibrium in the (full) behavioral domain.

Barlo and Dalkıran (2009) provides an analysis of implementation in epsilon-Nash equilibrium, i.e., when individuals are satisficing so that they are satisfied by getting close to (but not necessarily achieving) their best responses. This study can be considered one of the first studies in behavioral implementation as it allows for preferences with intransitive indifference, violating the transitivity axiom of preferences, the key axiom behind rational preferences. In particular, Barlo and Dalkıran (2009) presents epsilon-monotonicity and epsilon-limited veto power and shows that epsilon-monotonicity is necessary for an SCR to be implementable in epsilon Nash equilibrium, whereas epsilon-monotonicity coupled with epsilon-limited veto power is sufficient for epsilon-Nash implementation when there are three or more individuals in the society. That study originates from Dalkıran (2006) and borrows the limited veto power property from Benoit and Ok (2008), which shows that the sufficiency result of Maskin (1999) can be generalized to cases with limited veto power.

More recently, Barlo and Dalkıran (2022) highlights the fact that using computational tools, one can compute all of the consistent collection of sets associated with an SCR—as defined by de Clippel (2014). This may help the planner design practical and simple mechanisms to implement the SCR under consideration. Furthermore, they provide a slight improvement in terms of the necessary result for behavioral implementation for the case of two individuals by introducing two-individual consistent collections of sets associated with an SCR.

Barlo and Dalkıran (2023a) investigates the problem of behavioral implementation under incomplete information from an interim perspective. They identify the existence of an interim consistent profile of sets of acts as a necessary condition for a social choice set to be implementable under incomplete information. Combining the existence of an interim consistent profile of sets of acts with a choice incompatibility condition, Barlo and Dalkıran obtain a sufficient condition for behavioral implementation under incomplete information when there are three or more individuals in the society. Furthermore, by splicing de Clippel (2014)'s behavioral efficiency notion with interim incentive Pareto efficiency introduced by Holmström and Myerson (1983), they introduce a new behavioral efficiency notion under incomplete information called behavioral interim incentive Pareto efficiency and investigate its behavioral implementability under incomplete information. In a companion paper, Barlo and Dalkıran (2023b), they consider the problem of robust behavioral implementation by taking into account not only individuals' interim choices over acts but also their ex-post choices over deterministic alternatives.

Hayashi et al. (2023) expands the exploration of behavioral implementation so that coalition formations shall also be taken into account. To this end, they introduce the concept of behavioral strong equilibrium and define behavioral strong implementation accordingly. As a byproduct, they also provide another behavioral efficiency notion, which is non-nested with the efficiency notions presented by Bernheim and Rangel (2009) and de Clippel (2014).

Barlo and Dalkıran have co-authored other papers on implementation theory both in the rational and behavioral domains where they take individuals' choices as primitives rather than preferences, highlighting the fact that this allows non-trivial generalizations of the results even in the rational domain as well as helps us extend these results to the behavioral domains: In Altun et al. (2023), they show that, in economic environments with at least three individuals, the planner may Nash implement a social goal by extracting only the essential information about how individuals' choices are associated with the payoff states whenever one of the individuals—whose identity is not known to the planner and the other individuals—is a sympathizer who is inclined toward the truthful revelation of the essential information. In Barlo and Dalkıran (2022), they analyze the implementation problem when there is missing choice data, i.e., when planners do not observe societies' complete choice data.

Another recent paper, Gavan and Penta (2023), introduces the concept of safe implementation where individuals' deviations can only lead to—exogenously given—acceptable outcomes, which can accommodate both a variety of behavioral considerations and robustness concerns similar to à la Eliaz (2002). A parallel approach is employed by Barlo et al. (2023) to introduce the concept of anonymous implementation where, for each individual, the planner is restricted to employing the same set of alternatives as punishments for unilateral deviations from equilibrium, i.e., in each equilibrium of the mechanism, all individuals end up with the same opportunity set.

Finally, we would like to draw attention to some other prominent results in the literature that point out impossibility results close to our main result, highlighting that only social choice rules that satisfy some properties must be dictatorial. Gibbard (1973) highlights the susceptibility of non-dictatorial voting schemes to manipulation, revealing that non-dictatorial voting schemes can occur but lose strategy-proofness due to manipulation. Expanding on this, Satterthwaite (1975) formulates the well-known impossibility result known as the Gibbard–Satterthwaite Theorem, establishing a fundamental connection between strategy-proofness and dictatorial outcomes. Both Gibbard (1973) and Satterthwaite (1975) offer complementary insights related to Arrow's impossibility result. Jackson and Srivastava (1996) highlights the aspects of the solution concepts that lead to impossibility results that have a similar structure to the one we obtain in this thesis, i.e., when it is the case that only dictatorial SCRs are implementable on a full domain of preferences.

CHAPTER 3

THE MODEL

In this chapter, we first present the preliminaries to set up our model. Following a detailed proof of the aforementioned impossibility result by Maskin (1999), we present our main result: It is not possible to behaviorally implement a non-dictatorial SCR which is behaviorally efficient à la de Clippel (2014) in the full behavioral domain whenever there are only two individuals under consideration.

3.1 **Preliminaries**

Let X be a non-empty set of alternatives and N={1,2,...,n} be the set of all individuals in a society. Let Θ be the set of all possible relevant states of the world. We denote the rational preferences of individual i at state θ by a complete and transitive binary relation $\succsim_i^\theta$ where for any $x, y \in X$, $x \succsim_i^\theta y$ means that alternative x is at least as good as alternative y for individual i at state θ. Similarly, for any individual i, if x is strictly better than y at state θ, we write $x \succ_i^\theta y$; and if individual i is indifferent between x and y at θ, we write $x \sim_i^\theta y$.

For any $S \subseteq$ X, $C_i(S, \theta)$ denotes the choice correspondence of individual i at state θ, which selects a non-empty subset of S, i.e., $C_i(S, \theta) \subseteq S$. When Θ consists only of all possible rational preference profiles of the society, we call it the full rational domain, whereas, if Θ includes all possible choice profiles on X that do not induce a rational preference for at least one of the individuals at each state of the world, then then we refer to it as the full behavioral domain.

In order to distinguish the extent of the domain, with a slight abuse of notation, we use $\succsim_i^\theta$ when we are on rational domains, and when we refer to behavioral domains, we employ choice correspondences $C_i(\cdot, \theta)$, as defined above.

Definition 1. Individual i's choices at state θ satisfy **the weak axiom of revealed preferences (WARP)** if for any $x, y \in X$ and $S, S' \subseteq X$, with $x, y \in S \cap S'$, $x \in C_i(S, \theta)$ and $y \in C_i(S', \theta)$ implies $x \in C_i(S', \theta)$.

WARP is also known as Houthakker (1950)'s Axiom, who generalized it from the original definition provided by Samuelson (1938).

An individual's choice correspondence satisfies WARP if and only if it satisfies the independence of irrelevant alternatives, also known as Sen (1971)'s α, and an expansion consistency axiom called Sen (1971)'s β.

Definition 2. A **social choice rule** (SCR) $f : \Theta \twoheadrightarrow X$ is a correspondence that selects a non-empty subset of X for any state of the world.

For any $\theta \in \Theta$, $f(\theta)$ specifies the desired outcomes for state θ from the perspective of the social planner. If $x \in f(\theta)$, we refer to x as an f-optimal outcome at θ.

Definition 3. The **Pareto SCR** $f^{PO} : \Theta \twoheadrightarrow X$ selects all (weakly) Pareto optimal alternatives for any given state θ, and is defined as: $f^{PO}(\theta) = \{x \in X \mid$ for all $y \in X$, there exists $i \in N$ such that $x \succsim_i^\theta y\}$.

Definition 4. A social choice rule (SCR) $f : \Theta \twoheadrightarrow X$ satisfies **the Pareto property** if $f(\theta) \subseteq f^{PO}(\theta)$ for all $\theta \in \Theta$.

Definition 5. Individual i is called **a dictator of SCR** $f : \Theta \twoheadrightarrow X$ if and only if $[\forall \theta \in \Theta, \forall x \in X, x \in f(\theta)$ if and only if $x \succsim_i^\theta y$ for all $y \in X]$.

Definition 6. If an SCR $f : \Theta \twoheadrightarrow X$ has a dictator, then f is called **dictatorial**.

Definition 7. A **mechanism** $\mu = (M, g)$ specifies a message space M_i for each individual $i \in N$ and an outcome function $g : M \to X$ which specifies an alternative for every message profile where $M = (M_i)_{i \in N}$.

To clarify further, M_i is the set of all messages (or equivalently strategies) available to individual i and $M = (M_1 \times M_2 \times \ldots \times M_n)$ is the set of all message profiles. We denote individual i's strategy (message) by m_i and $m_{-i} = (m_1, m_2, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n)$ denotes the strategies of all individuals other than individual i. A message profile is denoted by $m = (m_1, m_2, \ldots, m_n)$, which we also denote as $m = (m_i, m_{-i})$. Finally, $g : M \to X$ is the outcome function that selects an alternative for any given message profile of the individuals.

Definition 8. On rational domains, we define the set of all **Nash equilibrium outcomes of mechanism** μ **at state** θ by $NE^{\mu}(\theta) = \{x \in X \,|x = g(m_i^*, m_{-i}^*)$ such that $g(m_i^*, m_{-i}) \succsim_i^{\theta} g(m_i^{'}, m_{-i})$ for all $i \in N$ and for all $m_i^{'} \in M_i\}$.

On behavioral domains, we follow the setup introduced by de Clippel (2014):

Definition 9. In any mechanism $\mu = (M, g)$, the set of alternatives $O_i^{\mu}(m_{-i}) = \{g(m_i, m_{-i}) \mid m_i \in M_i \}$ is called **the opportunity set of individual** i under mechanism μ given the message profile m_{-i} of all individuals other than i.

In words, $O_i^{\mu}(m_{-i})$ is the set of all the alternatives that individual i can reach given all the others' strategy profile $m_{-i} \in M_{-i}$.

Definition 10. In behavioral domains, **the set of Nash equilibrium outcomes of mechanism** μ **at state** θ is given by $NE^{\mu}(\theta) = \{x \mid x = g(m_i^*, m_{-i}^*)$ with $g(m_i^*, m_{-i}^*) \in C_i(O_i(m_{-i}^*(\theta), \theta))$ for all $i \in N\}$.

In words, Nash equilibrium outcomes of mechanism μ at state θ are the alternatives that correspond to a (behavioral) Nash equilibrium strategy profile, i.e., $(m_i^*)_{i\in N}$ with $g(m_i^*, m_{-i}^*) \in C_i(O_i(m_{-i}^*(\theta), \theta))$ for all $i \in N$.

Definition 11. A mechanism $\mu = (M, g)$ **Nash implements** social choice rule $f : \Theta \to X$ whenever $f(\theta) = NE^{\mu}(\theta)$ for all $\theta \in \Theta$. If there is such a mechanism, then we say that SCR $f : \Theta \to X$ is **Nash implementable**.

Definition 12. An alternative is called **behaviorally efficient à la de Clippel (2014)** at θ if $x \in f^{eff}(\theta)$ where $f^{eff}(\theta) = \{x \in X \mid \exists(\Upsilon_i)_{i\in N}$ with $\Upsilon_i \subseteq X$ and $x \in C_i(\Upsilon_i, \theta)$ for all $i \in N$ and $X = \cup_{i\in N}\Upsilon_i$ $\}$.

If alternative x is behaviorally efficient à la de Clippel (2014) at θ, we refer to the corresponding profile of sets $(\Upsilon_i)_{i\in N}$ as a collection of **implicit opportunity sets associated with x at state θ.**

With a slight abuse of notation, we say that SCR $f : \Theta \to X$ is behaviorally efficient à la de Clippel (2014) whenever $f(\theta)$ is behaviorally efficient à la de Clippel (2014) for all $\theta \in \Theta$.

Definition 13. In behavioral domains, an individual i is a **dictator of SCR** $f : \Theta \twoheadrightarrow X$ if and only if $[\forall\theta \in \Theta, \forall x \in X, x \in f(\theta)$ if and only if $x \in C_i(X, \theta)]$.

We wish to emphasize that it is not clear how to define a dictator in behavioral domains. Therefore, there might be other reasonable dictatorship definitions different than the one we provide above.

3.2 Impossibility result of Maskin (1999) for the case of two individuals

Before presenting our main result, we extrapolate the impossibility result of Maskin (1999) for the case of two individuals, which is formally stated below.

Theorem 1: [*Maskin (1999)*] Let $n = 2$, i.e., there are only two individuals in the society. If SCR $f : \Theta \twoheadrightarrow X$ satisfies the Pareto property where Θ is the full rational domain, then f is Nash implementable if and only if it is dictatorial.

To make our contribution clearer, we reproduce a detailed proof of this theorem:

Proof. Suppose that SCR $f : \Theta \twoheadrightarrow X$ is dictatorial and satisfies the Pareto property, and without loss of generality, assume that Individual 1 is the dictator of SCR f. Therefore, $x \in f(\theta)$ if and only if $x \succsim_1^\theta y$ for all $y \in X$, i.e., f-optimal alternatives at θ are the top alternatives of Individual 1 at θ.

Consider the mechanism where Individual 1 makes an announcement of an alternative, and the announcement is implemented. Formally, $N = \{1, 2\}$, $\mu = (M_1, M_2, g)$ where $M_1 = M_2 = X$ and g : $M_1 \times M_2 \to X$ such that for all $(m_1, m_2) \in M_1 \times M_2$, $g(m_1, m_2) = m_1$.

In every state of the world, in every Nash equilibrium of this mechanism, Individual 1 chooses their top alternatives—there can be more than one top-alternatives due to indifference—. Otherwise, Individual 1 would have a unilateral profitable deviation. Moreover, Individual 2 cannot change the outcome via any unilateral deviation. Formally, if $x \in f(\theta)$, then $m^* = (m_1^*, m_2^*)$ with $m_1^* = x, m_2^* = y$ for any $y \in X$ would be a Nash equilibrium of μ at θ as $x \in f(\theta)$ implies x is top-ranked by Individual 1 at θ. Hence, $f(\theta) \subseteq NE^\mu(\theta)$. Furthermore, if $m^* = (m_1^*, m_2^*)$ is a Nash equilibrium of μ at θ, then m_1^* is one of the top-ranked alternatives of Individual 1 at θ as Individual 1's opportunity set is X for any message of Individual 2, and hence, otherwise, Individual 1 would have a unilateral profitable deviation, leading to a contradiction. Therefore, $g(m^*) = m_1^*$ and since Individual 1 is the dictator of SCR f and m_1^* is among the top-ranked alternatives of Individual 1 at θ, we must have $m_1^* \in f(\theta)$, i.e., $NE^\mu(\theta) \subseteq f(\theta)$, as well.

For the other direction of the proof, suppose that SCR $f : \Theta \to X$ is implementable by the mechanism $\mu = (M_1, M_2, g)$, where M_i is the message set of individual i for each $i = \{1, 2\}$ and the outcome function is $g : M_1 \times M_2 \to X$. Note that if the alternative set X has only one element, both individuals become dictators of SCR $f : \Theta \to X$ because both individuals' top alternative is f-optimal at every state of the world —as there is no other alternative. Hence, let us assume that X contains at least two alternatives.

For any $j \in \{1, 2\}$, let $\bar{m}_j$ be a message of individual j under mechanism $\mu = (M, g)$, and for all $i \neq j$, $T_i(\bar{m}_j)$ be the set of alternatives that individual i cannot induce as an outcome when individual j chooses message $\bar{m}_j$ in mechanism μ. Formally, $T_i(\bar{m}_j) := X \setminus O_i^{\mu}(m_j) = \{x \mid x \neq g(m_i, \bar{m}_j)$ for any $m_i \in M_i\}$.

Claim 1: For any $m_1 \in M_1$ and $m_2 \in M_2$, $T_1(m_2) \cap T_2(m_1) = \emptyset$.

As highlighted by Maskin (1999), Claim 1 implies that in any mechanism that implements SCR f, given any the strategy profile of the individuals, every alternative in X is reachable by a unilateral deviation by one of the individuals.

Proof of Claim 1: For a contradiction, suppose that for $m_1^{'} \in M_1$ and $m_2^{'} \in M_2$, $T_1(m_2^{'}) \cap T_2(m_1^{'})$ is not an empty set. So, let $x \in T_1(m_2^{'}) \cap T_2(m_1^{'})$. Also let $y = g(m_1^{'}, m_2^{'})$. As the $T_1(m_2^{'})$ (or $T_2(m_1^{'})$) is the set of alternatives that Individual 1 (or 2) cannot induce by altering strategies after taking $m_2^{'}$ (or $m_1^{'}$) fixed, $g(m_1^{'}, m_2^{'})$ is not included in $T_1(m_2^{'})$ (or $T_2(m_1^{'})$). Therefore, $x \neq y$. Furthermore, suppose that at state θ, we have $x \succ_1^{\theta} y \succ_1^{\theta} z$ and $x \succ_2^{\theta} y \succ_2^{\theta} z$ for all $z \in X \setminus \{x, y\}$. But, this implies that, $y = g(m_1^{'}, m_2^{'})$ is a Nash equilibrium of μ at the state θ as y is the second-best alternative for both individuals but none of them can reach their best alternative as $x \in T_1(m_2^{'}) \cap T_2(m_1^{'})$. Hence, we must have $y \in f(\theta)$ since $\mu = (M, g)$ Nash implements SCR f. Observe, on the other hand, y is not Pareto optimal since for all $i = \{1, 2\}$, we have that $x \succ_i^{\theta} y$. But this leads to a contradiction as SCR f satisfies the Pareto property. Therefore, Claim 1 holds. □

Claim 2: For all $x \in X$, if $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$, then there is $\hat{m}_1 \in M_1$ such that for all $m_2 \in M_2$, $g(\hat{m}_1, m_2) = x$. Similarly, for all $x \in X$, if $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$ then there is $\hat{m}_2 \in M_2$ such that for all $m_1 \in M_1$, $g(\hat{m}_2, m_1) = x$.

Claim 2 implies that one of the individuals can guarantee to get alternative x whenever this alternative is reachable for any given message of the other individual, i.e., if alternative x exists in every column or in every row of the game form, then there is a message of Individual 1 or Individual 2 that gives a constant column or a constant row that consists of only alternative x.

Proof of Claim 2: Let $x \in X$ with $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$. Consider state θ such that we have $x \succ_1^\theta y$ and $y \succ_2^\theta x$ for all $y \in X \setminus \{x\}$. That is, at state θ, x is the best alternative for Individual 1, whereas it is the worst for Individual 2. Then, x is Pareto efficient at state θ. Since SCR f is Nash implemented by mechanism μ, there is a Nash equilibrium of μ at θ, (m_1^*, m_2^*), such that $g(m_1^*, m_2^*) \in NE^\mu(\theta)$. The existence of such a Nash equilibrium follows from the fact that SCR f is Nash implementable by supposition. By construction, $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ means that $x \notin T_1(m_2)$ for any $m_2 \in M_2$. Hence, $x \notin T_1(m_2^*)$, i.e., x is not one of the alternatives that Individual 1 cannot reach given the message m_2^* of Individual 2. Equivalently, given m_2^*, Individual 1 can reach the outcome x. So, there is some $\breve{m}_1 \in M_1$, $g(\breve{m}_1, m_2^*) = x$. Because $g(m_1^*, m_2^*)$ is a Nash equilibrium, we must have $g(m_1^*, m_2^*) \succsim_1^\theta g(\breve{m}_1, m_2^*) = x$. Thus, either $\breve{m}_1 = m_1^*$ and $g(m_1^*, m_2^*) = x$ or $\breve{m}_1 \neq m_1^*$ and $g(\breve{m}_1, m_2^*) = g(m_1^*, m_2^*) = x$ since x is the top-alternative for Individual 1 at state θ. On the other hand, in order for Individual 2 to not have any incentive to deviate, there does not exist any message $\breve{m}_2 \in M_2$ such that $g(m_1^*, \breve{m}_2) = z \neq x$. This follows since otherwise $x = g(m_1^*, m_2^*) \not\succsim_2^\theta g(m_1^*, \breve{m}_2) = z$ would lead to a contradiction because, by construction, $y \succ_2^\theta x$ for all $y \in X \setminus \{x\}$ meaning that $z \succ_2^\theta x$ whenever $z \in X \setminus \{x\}$. Hence, it must be that for all $m_2 \in M_2$, $g(m_1^*, m_2) = x$, i.e., Individual 1 can guarantee inducing alternative x by sending the message m_1^* as claimed. Changing the roles of Individual 1 and Individual 2 will yield the same result of the existence of a strategy such that

Individual 2 can guarantee any alternative $x \in X$ whenever $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. Therefore, Claim 2 holds. $\square$

Claim 3: Either for all $x \in X$, $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or for all $x \in X$, $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$.

Proof of Claim 3: First, we show that $z \in \bigcup_{m_j \in M_j} T_i(m_j)$ implies that $z \notin \bigcup_{m_i \in M_i} T_j(m_i)$ for both $i = 1, 2$. Without loss of generality, suppose $z \in \bigcup_{m_2 \in M_2} T_1(m_2)$. Then, there is $m_2^{'} \in M_2$ such that $z \in T_1(m_2^{'})$. But then, there is no message $m_1^{'}$ of Individual 1, which implies that $z \in T_2(m_1^{'})$ because otherwise Claim 1 would be violated. Therefore, z cannot be in $T_2(m_1)$ for any $m_1 \in M_1$. Hence, $z \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. Thus, it must be that for any $x \in X$, either $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$.

Next, suppose that for some $a, b \in X$ with $a \neq b$, we have $a \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ and $b \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. It follows from Claim 2 that there is $\dot{m}_1 \in M_1$and $\dot{m}_2 \in M_2$ such that $g(\dot{m}_1, m_2) = a$ for all $m_2 \in M_2$ and $g(m_1, \dot{m}_2) = b$ for all $m_1 \in M_1$. But then, $g(\dot{m}_1, m_2) = a$ for all $m_2 \in M_2$, we must have that $g(\dot{m}_1, \dot{m}_2) = a$. Similarly, $g(m_1, \dot{m}_2) = b$ for all $m_1 \in M_1$ implies that $g(\dot{m}_1, \dot{m}_2) = b$. Since a and b are assumed to be different from each other, we obtain a contradiction. Therefore, for any $a, b \in X$ with $a \neq b$, we have either $a, b \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or $a, b \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. Hence, for all $x \in X$, either $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or for all $x \in X$, $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. $\square$

Therefore, it follows from the claims proved above, that either Individual 1 or Individual 2 can guarantee inducing all possible alternatives. However, this means that one of the individuals can always send a message to her top alternative at any state, and it will be a Nash equilibrium (no matter the other individual's message) at this state since the other individual cannot change the outcome. Since SCR f is implementable, this means that the top alternative of this individual must always coincide with f optimal alternatives, which implies that this individual is a dictator for SCR f, i.e., SCR f is dictatorial. ■

3.3 The relationship between SCR f^{PO} and SCR f^{eff}

In this section, we reaffirm the fact that SCR f^{eff} extends SCR f^{PO} to behavioral domains as pointed out by de Clippel (2014). That is, we show that these two SCRs coincide in rational domains.

Recall that when individuals' choices satisfy WARP, rational preferences can be represented by a choice correspondence, which can be defined as $C_i^\star(S,\theta) = \{x \in X \mid x \succsim_i^\theta y \text{ for all } y \in S\}$ for all $i \in N$, for all $\theta \in \Theta$ and for all $S \subseteq X$. Furthermore, under rational preferences, the lower contour sets are well-defined. We denote the lower contour set of x for individual i at state θ by $L_{\succsim_i^\theta}(x)$. Formally, $L_{\succsim_i^\theta}(x) = \{z \in X \mid x \succsim_i^\theta z\}$.

We now reproduce the aforementioned equivalence:

Claim 4: $f^{PO} = f^{eff}$ on any rational domain, i.e., whenever individuals' choices satisfy WARP.

Proof. Let $f^{PO} : \Theta \twoheadrightarrow X$ and $f^{eff} : \Theta \twoheadrightarrow X$ be SCRs defined as before, i.e., $f^{PO}(\theta) = \{x \in X \mid \text{for all } y \in X, \text{ there exists } i \in N \text{ such that } x \succsim_i^\theta y\}$ and $f^{eff}(\theta) = \{x \in X \mid \exists(\Upsilon_i)_{i\in N} \text{ with } \Upsilon_i \subseteq X \text{ and } x \in C_i(\Upsilon_i,\theta) \text{ for all } i \in N \text{ and } X = \cup_{i\in N}\Upsilon_i\}$.

<u>$f^{PO} \subseteq f^{eff}$</u>: Let $\theta \in \Theta$ and $x \in f^{PO}(\theta)$. As x is Pareto optimal at θ, by definition, for any $y' \in X$ there is some $i \in N$ such that $y' \in L_{\succsim_i^\theta}(x)$. Clearly, $x \in C_i^\star(L_{\succsim_i^\theta}(x),\theta)$ for any $i \in N$ as none of the options in the lower contour set of x is strictly better than x by definition. Since for all $y \in X$ there is some $j \in N$ such that $y \in L_{\succsim_j^\theta}(x)$, it must be that the union of all lower contour sets must yield the whole set, i.e., $\cup_{y\in X} = X \subseteq \cup_{j\in N} L_{\succsim_j^\theta}(x) \subseteq X$. Hence, we

must have $X = \cup_{j \in N} L_{\succsim_j^\theta}(x)$. Thus, by setting $\Upsilon_i = L_{\succsim_i^\theta}(x)$ for all $i \in N$ we obtain a collection of implicit opportunity, $(\Upsilon_i)_{i \in N}$ as in the definition of f^{eff}. Therefore, it must be that $x \in f^{eff}(\theta)$. So, $f^{PO}(\theta) \subseteq f^{eff}(\theta)$ for all θ whenever $C_i^\star(S, \theta) = \{x \in X \mid x \succsim_i^\theta y \text{ for all } y \in S\}$ for all $S \subseteq X$ for some rational preference $\succsim_i^\theta$, i.e., whenever individuals' choices satisfy WARP.

$\underline{f^{eff} \subseteq f^{PO}}$: Suppose individuals' choices satisfy WARP. Let $\delta \in \Theta$, $x' \in f^{eff}(\delta)$, and $\succsim_i^\delta$ represent the rational preferences of individual i at δ. Suppose that for the collection of implicit opportunity sets $(\Upsilon_i^{'})_{i \in N}$ we have $\Upsilon_i^{'} \subseteq X$ and $x' \in C_i^\star(\Upsilon_i^{'}, \delta)$ for all $i \in N$ and also $X = \cup_{i \in N} \Upsilon_i^{'}$ as required by the definition of SCR f^{eff}. Because $x' \in C_i^\star(\Upsilon_i^{'}, \delta)$, we have that $x' \succsim_i^\delta z'$ for all $z' \in \Upsilon_i^{'}$ and for all $i \in N$. Thus, it must be $\Upsilon_i^{'} \subseteq L_{\succsim_i^\delta}(x')$ by definition of the lower contour sets. Moreover, $\Upsilon_i^{'} \subseteq L_{\succsim_i^\delta}(x')$ together with $X = \cup_{i \in N} \Upsilon_i^{'}$ imply that $X = \cup_{i \in N} L_{\succsim_i^\delta}(x')$. The last result can also be stated as for all $z \in X$, there is some $i \in N$ such that $z \in L_{\succsim_i^\delta}(x')$ or equivalently, $x' \succsim_i^\delta z$. Therefore, x' is Pareto optimal at δ, i.e., $x' \in f^{PO}(\delta)$. Hence, $f^{eff}(\delta) \subseteq f^{PO}(\delta)$ as well.

Therefore, $f^{PO} = f^{eff}$ under rationality. ∎

3.4 Does Maskin (1999)'s impossibility result for Nash Implementation with two individuals extend to behavioral domains?

We present the main result of this thesis, Theorem 2: An extension of Maskin (1999)'s impossibility result for Nash implementation with two individuals in the full rational domain to the full behavioral domain.

Let Θ^B be the full behavioral domain, i.e., it contains each state of the world for each possible choice profile where the choice correspondence of each individual fails the weak axiom of revealed preferences (WARP).

Theorem 2: Let $n = 2$ and $|X| \geq 4$, i.e., there are only two individuals and at least four alternatives. If SCR $f : \Theta^B \twoheadrightarrow X$ is behaviorally efficient à la de Clippel (2014), then SCR f is Nash implementable if and only if it is dictatorial.

Proof. We follow similar steps as in the extended proof we provide for Maskin (1999)'s impossibility result for Nash implementation with two individuals in the full rational domain above.

We first show that if SCR f is dictatorial, then it is Nash implementable:

Suppose that SCR $f : \Theta^B \twoheadrightarrow X$ is dictatorial and without loss of generality, assume that Individual 1 is a dictator of SCR f. Consider the mechanism where Individual 1 chooses an alternative, and this alternative is implemented. Formally; $N = \{1, 2\}$, $\mu = (M_1, M_2, g)$ where $M_1 = M_2 = X$ and $g : M_1 \times M_2 \to X$ such that for all $(m_1, m_2) \in M_1 \times M_2$, $g(m_1, m_2) = m_1$. Notice that we have $O_1^{\mu}(m_2) = X$ for all $m_2 \in M_2$ and $O_2^{\mu}(m_1) = \{m_1\}$ for all $m_1 \in M_1$. Therefore, at any state $\theta \in \Theta^B$, the message profile $(m_1^{\star}, m_2)$ forms a Nash equilibrium for any $m_1^{\star} \in C_1(X, \theta) \subseteq M_1 = X$ and any $m_2 \in M_2$ because we have $m_1^{\star} \in C_1(X, \theta) \cap C_2(\{m_1^{\star}\}, \theta)$. In other words, $NE^{\mu}(\theta) = C_1(X, \theta) = f(\theta)$. Hence, μ Nash implements SCR f as desired.

Next, we show that if SCR f is behaviorally efficient à la de Clippel (2014) and Nash implementable, then it is dictatorial:

Suppose that SCR $f : \Theta^B \twoheadrightarrow X$ is Nash implemented by the mechanism $\mu = (M_1, M_2, g)$, where M_i is the message set for both $i = \{1, 2\}$ and the outcome function is $g : M_1 \times M_2 \to X$. Note that if $|X| \leq 2$, i.e., if there are two or fewer alternatives, independent of the choice structure, the non-empty choices induce rational preferences. So suppose $|X| \geq 3$, i.e., there are at least three alternatives. Even though the hypothesis of the theorem is $|X| \geq 4$, Claim 5 below holds when $|X| \geq 3$ as well.

Recall that for any given $\bar{m}_j \in M_j$ of individual j for all $j \in \{1, 2\}$, $T_i(\bar{m}_j) := \{x \mid x \neq g(m_i, \bar{m}_j) \text{ for any } m_i \in M_i\}$ is the set of alternatives that i cannot induce by changing her messages unilaterally when j sends the message $\bar{m}_j$.

Claim 5: For any $m_1 \in M_1$ and $m_2 \in M_2$, we have $T_1(m_2) \cap T_2(m_1) = \emptyset$.

Proof of Claim 5: Suppose for contradiction, for some $m_1^{'} \in M_1$ and some $m_2^{'} \in M_2$, $T_1(m_2^{'}) \cap T_2(m_1^{'})$ is not empty. Let $x_0 \in T_1(m_2^{'}) \cap T_2(m_1^{'})$ and let $y_0 = g(m_1^{'}, m_2^{'})$. Because $T_1(m_2^{'})$ (or $T_2(m_1^{'})$) is the set of alternatives that Individual 1 (or 2) cannot induce by altering her message given $m_2^{'}$ (or $m_1^{'}$), $g(m_1^{'}, m_2^{'})$ is not included in $T_1(m_2^{'})$ (or $T_2(m_1^{'})$), which implies that $x_0 \neq y_0$.

Suppose at state θ, we have the following choice profile:

$(i.)$ $C_1(\{z\}, \theta) = C_2(\{z\}, \theta) = \{z\}$ for all $z \in X$,

$(ii.)$ if $x_0, y_0 \in S \subseteq X$, $C_1(S, \theta) = C_2(S, \theta) = \{x_0\}$,

$(iii.)$ if $y_0 \in S \subseteq X \setminus \{x_0\}$, $C_1(S, \theta) = C_2(S, \theta) = \{y_0\}$,

$(iv.)$ if $x_0 \in S \subseteq X \setminus \{y_0\}$, $C_1(S, \theta) = C_2(S, \theta) = S \setminus \{x_0\}$,

$(v.)$ if $x_0, y_0 \notin S \subseteq X$, $C_1(S, \theta) = C_2(S, \theta) = S$.

Note that, by (ii), $C_1(X, \theta) = C_2(X, \theta) = \{x_0\}$ and by (iv), we have $C_1(X \setminus \{y_0\}, \theta) = C_2(X \setminus \{y_0\}, \theta) = X \setminus \{x_0, y_0\}$. Hence, both individuals' choices fail WARP at state θ as they choose x_0 from the set of all alternatives, X, but they do not choose it from $X \setminus \{y_0\}$, a proper subset of X. Therefore, $\theta \in \Theta^B$.

Observe that because $y_0 = g(m_1^{'}, m_2^{'})$, we have $y_0 \in [O_1(m_2^{'}) \cap O_2(m_1^{'})]$ and because $x_0 \in T_1(m_2^{'}) \cap T_2(m_1^{'})$, we have $x_0 \notin O_1(m_2^{'})$ and $x_0 \notin O_2(m_1^{'})$. Thus, each of $O_1(m_2^{'})$ and $O_2(m_1^{'})$ either belongs to item (i) or item (iii) in the itemized list of choices given above. This implies that $C_1(O_1(m_2^{'}), \theta) \cap C_2(O_2(m_1^{'}), \theta) =$

$\{y_0\}$, which means that (m_1', m_2') is a Nash equilibrium of μ at state θ, i.e., $y_0 \in NE^{\mu}(\theta)$. Because SCR f is Nash implemented by μ, this means that y_0 is also f-optimal at θ, i.e., $y_0 \in f(\theta)$. However, observe that y_0 cannot be behaviorally efficient à la de Clippel (2014) at state θ. To see why there is no collection of implicit opportunity sets (Υ_1, Υ_2) such that $y_0 \in C_1(\Upsilon_1, \theta) \cap C_2(\Upsilon_2, \theta)$ with $\Upsilon_1 \cup \Upsilon_2 = X$ observe that if $y_0 \in C_1(\Upsilon_1, \theta) \cap C_2(\Upsilon_2, \theta)$, then either $\Upsilon_i = \{y_0\}$ or $\Upsilon_i \subseteq X \setminus \{x_0\}$ for both $i \in \{1, 2\}$ because Υ_1 and Υ_2 shall either belong to item (i) or item (iii) in the listed choice categories. Hence, in all possible scenarios, $\Upsilon_1 \cup \Upsilon_2 \neq X$ as $x_0 \notin \Upsilon_i$ for both $i = \{1, 2\}$. This contradicts the supposition that SCR f is behaviorally efficient à la de Clippel (2014). Therefore, we conclude that Claim 5 holds. □

Claim 6: For all $x \in X$, if $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ then there is $\hat{m}_1 \in M_1$ such that for all $m_2 \in M_2$, $g(\hat{m}_1, m_2) = x$. Similarly, for all $x \in X$, if $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$ then there is $\hat{m}_2 \in M_2$ such that for all $m_1 \in M_1$, $g(\hat{m}_2, m_1) = x$.

Proof of Claim 6: Let $x_0, y_0, z_0 \in X$ be distinct alternatives, i.e., $x_0 \neq y_0$, $y_0 \neq z_0$, and $x_0 \neq z_0$. Without loss of generality, suppose that $x_0 \notin \bigcup_{m_2 \in M_2} T_1(m_2)$.

Consider state θ where we have the following choice structure:

$(i.)$ $C_1(\{t\}, \theta) = C_2(\{t\}, \theta) = \{t\}$ for all $t \in X$,

$(ii.)$ if $x_0, y_0 \in S \subseteq X$, $C_1(S, \theta) = \{x_0\}$ and $C_2(S, \theta) = \{y_0\}$

$(iii.)$ if $y_0 \in S \subseteq X \setminus \{x_0\}$, [if $z_0 \in S$, $C_1(S, \theta) = \{z_0\}$, otherwise $C_1(S, \theta) = S \setminus \{y_0\}$] and $C_2(S, \theta) = \{y_0\}$

$(iv.)$ if $x_0 \in S \subseteq X \setminus \{y_0\}$, $C_1(S, \theta) = \{x_0\}$ and $C_2(S, \theta) = S \setminus \{x_0, z_0\}$,

$(v.)$ if $x_0, y_0 \notin S \subseteq X$, $C_1(S, \theta) = S \setminus \{z_0\}$ and [if $z_0 \in S$, $C_2(S, \theta) = \{z_0\}$, otherwise $C_2(S, \theta) = S$].

Observe that, by (iii), we have $z_0 \in C_1(X \setminus \{x_0\}, \theta)$ and by (v), we have $z_0 \notin C_1(X \setminus \{x_0, y_0\})$ and hence Individual 1's choices at state θ fails WARP. Similarly, for any $t \in X \setminus \{x_0, y_0, z_0\}$, by (v), we have $t \in C_2(X \setminus \{y_0\}, \theta)$ but also $t \notin C_2(X \setminus \{x_0, y_0\}, \theta)$ since $C_2(X \setminus \{x_0, y_0\}, \theta) = \{z_0\}$. Therefore, Individual 2's choices at θ does not satisfy WARP as well.

Because SCR f is implementable by hypothesis, there exists $(m_1^\star, m_2^\star) \in (M_1, M_2)$ such that $g(m_1^*, m_2^*) \in NE^\mu(\theta)$ and $g(m_1^*, m_2^*) \in f(\theta)$. Since $x_0 \notin \bigcup_{m_2 \in M_2} T_1(m_2)$, we have $x_0 \notin T_1(m_2^*)$. In other words, Individual 1 can reach alternative x_0 by a unilateral deviation when Individual 2 sends message m_2^*. Formally, for some $m_1' \in M_1$, we have $g(m_1', m_2^*) = x_0$. Hence, $x_0 \in O_1^\mu(m_2^*)$. According to the choice structure at state θ, $O_1^\mu(m_2^*)$ belongs to one of the items (i), (ii) or (iv) as it contains x_0. In any case Individual 1 chooses x_0 from $O_1(m_2^*)$ as specified in (i), (ii) and (iv). Therefore, we either have $m_1' = m_1^*$ and $g(m_1^*, m_2^*) = x_0$ or $m_1' \neq m_1^*$ and $g(m_1', m_2^*) = g(m_1^*, m_2^*) = x_0$. On the other hand, observe that Individual 2 never chooses x_0 if she faces two or more alternatives. Because (m_1^*, m_2^*) is a Nash equilibrium, it must be that $O_2(m_1^*) = \{x_0\}$, i.e., setting $\hat{m}_1 = m_1^\star$, we get $g(\hat{m}_1, m_2) = x_0$ for all $m_2 \in M_2$ as desired. A similar argument follows by interchanging the roles of the individuals when $x_0 \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. Thus, Claim 6 holds. □

Claim 7: Either for all $x \in X$, $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or for all $x \in X$, $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$.

Proof of Claim 7: Suppose that for some $x \in X$, $x \in \bigcup_{m_2 \in M_2} T_1(m_2)$. Then, for some $m_2' \in M_2$, $x \in T_1(m_2')$. Hence, by Claim 5, we have $\forall m_1' \in M_1$, $x \notin T_2(m_1')$, which implies that, $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. It is clear that interchanging the roles of the individuals yields a similar result. Therefore, if an alternative cannot be induced via a unilateral deviation by one individual, then the other individual can induce it via a unilateral deviation, i.e., it must be that for any $x \in X$, either

$x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$.

Suppose that for some $y, z \in X$ with $y \neq z$, we have $y \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ and $z \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. Then, by Claim 6, we have a message profile $(m_1^{'}, m_2^{'})$ so that $g(m_1^{'}, m_2) = y$ for all $m_2 \in M_2$ and $g(m_1, m_2^{'}) = z$ for all $m_1 \in M_1$. However, $g(m_1^{'}, m_2) = y$ for all $m_2 \in M_2$ implies that $g(m_1^{'}, m_2^{'}) = y$ and $g(m_1, m_2^{'}) = z$ for all $m_1 \in M_1$ implies that $g(m_1^{'}, m_2^{'}) = z$. Thus, $g(m_1^{'}, m_2^{'}) = y = z$, a contradiction! Therefore, for any $y, z \in X$ with $y \neq z$, $y, z \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or $y, z \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. That is, if multiple alternatives can be unilaterally reached, only one of the individuals has the ability to reach all of these alternatives unilaterally. Hence, either for all $x \in X$, $x \notin \bigcup_{m_2 \in M_2} T_1(m_2)$ or for all $x \notin \bigcup_{m_1 \in M_1} T_2(m_1)$. □

Claims 5, 6, and 7 together imply that either Individual 1 or Individual 2 can obtain all possible alternatives unilaterally. This means that for some $i \in \{1, 2\}$, $O_i^{\mu}(m_j) = X$ for all $m_j \in M_j$. Then, any Nash equilibrium $(m_1^{\star}, m_2^{\star})$ of μ at any state θ induces an outcome such that $g(m_1^{\star}, m_2^{\star}) \in C_i(X, \theta)$. Since SCR f is implementable, we have $g(m_1^{\star}, m_2^{\star}) \in f(\theta)$ for any Nash equilibrium $(m_1^{\star}, m_2^{\star})$ at θ and $f(\theta) = NE^{\mu}(\theta)$ for all $\theta \in \Theta^B$. Therefore, $x \in f(\theta)$ if and only if $x \in C_i(X, \theta)$, which means that individual i is a dictator for SCR f, and hence, SCR f is dictatorial.

Finally, we point out why we need $|X| \geq 4$: At the beginning of the proof, we assume that the set of alternatives contains at least three alternatives to be in the full behavioral domain. If there are exactly three alternatives, i.e., $|X|$ =3, and for some $x_0, y_0 \in X$, $x_0 \in C_1(X, \theta)$ and $x_0 \notin C_2(S, \theta)$ for all $S \neq \{x_0\}$ whereas $y_0 \in C_2(X, \theta)$ and $y_0 \notin C_1(S, \theta)$ for all $S \neq \{y_0\}$, the choice structure used in the proof of Claim 6 implies that at least one individual's choices do not fail WARP. That is, $|X| \geq 4$ suffices to construct the choice structures we employ in the proof of Claim 6 under the full behavioral domain. ■

CHAPTER 4

CONCLUSION

In this thesis, we have focused on Nash Implementation of a behaviorally efficient SCR à la de Clippel (2014) in the full behavioral domain with two individuals. Following the impossibility result of Maskin (1999) for Nash Implementation with two individuals, which states that in the full rational domain, a Pareto efficient two-individual SCR is implementable if and only if it is dictatorial, we showed that this result extends to the full behavioral domain if one defines a dictatorial SCR appropriately in the context of behavioral domains. This extension follows from the fact that behavioral efficiency à la de Clippel (2014) extends Pareto efficiency to behavioral domains, i.e., as we have reaffirmed in this thesis, behavioral efficiency à la de Clippel (2014) coincides with Pareto efficiency in the rational domains.

REFERENCES

Altun, O. A., M. Barlo, and N. A. Dalkıran (2023). Implementation with a sympathizer. *Mathematical Social Sciences 121*, 36–49. 9

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy 58*(4), 328–346. 2, 10

Barlo, M. and N. A. Dalkıran (2022). Implementation with missing data. *Mimeo*. 9

Barlo, M. and N. A. Dalkıran (2023a). Behavioral implementation under incomplete information. *Journal of Economic Theory 213*. 8

Barlo, M. and N. A. Dalkıran (2023b). Robust behavioral implementation. *Mimeo*. 9

Barlo, M., N. A. Dalkıran, and M. Dönmez (2023). Anonymous implementation. *Mimeo*. 10

Barlo, M. and N. A. Dalkıran (2009). Epsilon-Nash implementation. *Economics Letters 102*(1), 36–38. 7, 8

Barlo, M. and N. A. Dalkıran (2022). Computational implementation. *Review of Economic Design 26*, 605–633. 8, 9

Benoit, J.-P. and E. A. Ok (2008). Nash implementation without no-veto power. *Games and Economic Behavior 64*(1), 51–67. 8

Bernheim, B. D. and A. Rangel (2009). Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics 124*(1), 51–104. 3, 9

Chernoff, H. (1954). Rational selection of decision functions. *Econometrica*, 422–443. 7

Dalkıran, N. A. (2006). On continuity of Maskin's implementation result. *M.A. Thesis*. Sabancı University. 8

de Clippel, G. (2014). Behavioral implementation. *American Economic Review 104*(10), 2975–3002. iii, iv, 2, 3, 4, 7, 8, 9, 11, 13, 14, 19, 21, 23, 26

Dutta, B. and A. Sen (1991). A necessary and sufficient condition for two-person nash implementation. *Review of Economic Studies 58*(1), 121–128. 6

Eliaz, K. (2002). Fault tolerant implementation. *Review of Economic Studies 69*(3), 589–610. 10

Gavan, M. J. and A. Penta (2023). Safe implementation. *Mimeo*. 10

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica 41*(4), 587–601. 10

Hayashi, T., R. Jain, V. Korpela, and M. Lombardi (2023). Behavioral strong implementation. *Economic Theory 76*, 1257–1287. 9

Holmström, B. and R. B. Myerson (1983). Efficient and durable decision rules with incomplete information. *Econometrica*, 1799–1819. 9

Houthakker, H. S. (1950). Revealed preference and the utility function. *Economica 17*(66), 159–174. 12

Hurwicz, L. (1986). On the implementation of social choice rules in irrational societies. *Social choice and public decision making: Essays in Honor of Kenneth J. Arrow 1*, 75–96. 7

Hurwicz, L. and D. Schmeidler (1978). Construction of outcome functions guaranteeing existence and pareto optimality of nash equilibria. *Econometrica*, 1447–1474. 2

Jackson, M. O. (2001). A crash course in implementation theory. *Social Choice and Welfare 18*, 655–708. 7

Jackson, M. O. and S. Srivastava (1996). A characterization of game-theoretic solutions which lead to impossibility theorems. *Review of Economic Studies 63*(1), 23–38. 10

Korpela, V. (2012). Implementation without rationality assumptions. *Theory and Decision 72*, 189–203. 7

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies 66*(3), 23–38. iii, iv, vi, 1, 2, 3, 4, 5, 7, 8, 11, 14, 16, 20, 21, 26

Maskin, E. and T. Sjöström (2002). Implementation theory. *Handbook of Social Choice and Welfare 1*, 237–288. 7

Moore, J. and R. Repullo (1990). Nash implementation: A full characterization. *Econometrica 58*(5), 1083–1099. 2, 6, 7

Palfrey, T. R. (2002). Implementation theory. *Handbook of Game Theory with Economic Applications 3*, 2271–2326. 7

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica 5*(17), 61–71. 12

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. 10

Sen, A. K. (1971). Choice functions and revealed preference. *Review of Economic Studies 38*(3), 307–317. 7, 12

Serrano, R. (2004). The theory of implementation of social choice rules. *SIAM Review 46*(3), 377–414. 7

Sjöström, T. (1991). On the necessary and sufficient conditions for nash implementation. *Social Choice and Welfare 8*(4), 333–340. 6