Behavioral Implementation under Incomplete Information*

Mehmet Barlo[†] Nuh Aygün Dalkıran[‡]

April 24, 2019

Abstract

We investigate implementation under incomplete information when individuals' choices need not be rational. Our results are complementary to de Clippel (2014) [*American Economic Review*, 104(10): 2975-3002], which investigates the same problem under complete information.

Keywords: Behavioral Implementation, Incomplete Information, Bounded Rationality, Ex-Post Implementation.

JEL Classification: D60, D79, D82, D90

*We would like to thank Ahmet Alkan, Dirk Bergemann, Eddie Dekel, Mehmet Ekmekci, Jeff Ely, Ehud Kalai, Semih Koray, Wojciech Olszewski, Alessandro Pavan, Yuval Salant, Marciano Siniscalchi, Asher Wolinsky, Kemal Yıldız, and the seminar participants at Northwestern University, Social Choice and Welfare 2018, ASSET 2018, and The Asian Meeting of the Econometric Society 2017 for enjoyable conversations and helpful comments. Ece Yegane provided excellent research assistance. Part of this work was completed while Nuh Aygün Dalkıran was visiting the Managerial Economics & Decision Sciences Department at Kellogg School of Management. He is thankful to The Fulbright Program for making this visit possible and Northwestern University | Kellogg School of Management for their hospitality. Any remaining errors are ours.

[†]Faculty of Arts and Social Sciences, Sabancı University; barlo@sabanciuniv.edu

[‡]Corresponding author, Department of Economics, Bilkent University; dalkiran@bilkent.edu.tr

Table of Contents

1 Introduction **1**

2 Related Literature **4**

3 Motivating Example **5**
- 3.1 The mechanism . . . 12
- 3.2 Ex-post equilibrium outcomes of the mechanism . . . 14
- 3.3 The revelation principle fails . . . 15

4 Notation and Definitions **17**

5 Three or More Individuals **19**

6 Two Individuals **25**

7 Simple Mechanisms **32**
- 7.1 Motivating example revisited . . . 33
- 7.2 Lower bounds on the number of messages . . . 34

8 Concluding Remarks **36**

A Mechanisms **37**
- A.1 Mechanism for the case with three or more individuals . . . 37
- A.2 Mechanism for the case with two individuals . . . 38

B Proofs **40**
- B.1 Proof of Claim 1 . . . 40
- B.2 Proof of Theorem 1 . . . 43
- B.3 Proofs of Propositions 1, 2, and 3 and Lemma 1 . . . 43
- B.4 Proof of Theorem 2 . . . 44
- B.5 Proof of Theorem 3 . . . 45
- B.6 Proof of Theorem 4 . . . 46
- B.7 Proof of Theorem 5 . . . 47
- B.8 Proof of Theorem 6 . . . 49
- B.9 Proof of Theorem 7 . . . 50

C An Illustration of Sufficiency with Two Individuals **50**

References **54**

List of Tables

1 Individual choices of Alice and Bob. 9
2 BR-optimal alternatives. 11
3 The social choice set F. 11
4 The mechanism μ for Alice and Bob. 12
5 The best responses and Nash equilibria of the mechanism at (γ_A, γ_B). 13
6 Nash equilibria and Nash equilibrium outcomes of the mechanism. 14
7 Ex-post equilibria and ex-post equilibrium outcomes of the mechanism. 15
8 The direct revelation mechanism g^d. 16
9 A two-individual example satisfying consistency-unanimity. 32
10 BR-optimal alternatives and SCF f. 32
11 A 3×2 mechanism proposal for Alice and Bob. 33
12 Another simplest mechanism for Alice and Bob. 34
13 The outcome function of the mechanism with three or more individuals. 37
14 The outcome function of the two individual mechanism. 39

1 Introduction

People have limited cognitive abilities and are prone to various behavioral biases; this is documented by ample evidence in the literature of marketing, psychology, and behavioral economics. Thus, it is not surprising that the behavior of individuals may not be consistent with the *standard axioms of rationality*.[1,2] What shall a planner do if he/she wants to implement a goal when the relevant information is distributed among "predictably irrational" individuals?

The present paper provides an analysis of the theory of implementation under incomplete information when individuals' choices do not necessarily comply with the standard axioms of rationality. That is, our paper is the incomplete information counterpart of de Clippel (2014), which provides an analysis for the case of complete information. Our results provide an important leap in behavioral implementation as information asymmetries are inescapable in many economic settings.

In particular, we analyze the problem of designing a mechanism under incomplete information when individuals are allowed to display any type of behavioral biases, such as falling for an attraction effect, displaying a status-quo bias or revealing cyclic preferences (as is the case when groups act as an individual), among many others. In doing so, we focus on full implementation and employ ex-post equilibrium (hereafter, abbreviated to EPE) as our main concept of equilibrium for the following reasons.

Full implementation of a predetermined social choice rule requires that the set of equilibrium outcomes of the associated mechanism fully coincide with the given social choice rule. On the other hand, *partial implementation* only requires that the prede-

[1] This is why the recent trend involving the use of behavioral insights in policy-making has been growing stronger, implying an increased interest in adapting economic models to allow behavioral biases. In particular, Thaler and Sunstein's New York Times best-seller book *Nudge* has been influential guiding real life policies. For instance, the Behavioral Insights Team, a.k.a. the *Nudge Unit*, has been established in 2010 in the United Kingdom. In the United States, President Obama released an executive order in 2015, emphasizing the importance of behavioral insights to deliver better policy outcomes at lower costs and at the same time encouraging executive departments and agencies to incorporate these insights into policy-making. Many countries and international institutions followed, there are now more than a dozen countries besides EU, OECD, UN agencies, and the World Bank that integrated behavioral insights into their operations (Afif, 2017). There is such a trend in the academic literature as well, e.g., Spiegler (2011) provides a synthesis of efforts to adapt models in industrial organization to bounded rationality.

[2] We say that individuals' choices satisfy the standard axioms of rationality whenever their choices obey the weak axiom of revealed preferences, which is formalized in Footnote 12. Besides *Nudge* (Thaler & Sunstein, 2008), two other New York Times best-seller books documenting various behavioral biases leading to failure of the standard axioms of rationality are *Predictably Irrational* (Ariely, 2009) and *Thinking, Fast and Slow* (Kahneman, 2011).

termined social choice rule be sustained by an equilibrium of the mechanism; hence, it allows for other equilibria associated with outcomes that are not aligned with the social goal at hand. Even though it is weaker than full implementation, partial implementation is rather widely used in the rational domain under incomplete information. In fact, the appeal of partial implementation depends heavily on *the revelation principle*, which implies, in the rational domain and under incomplete information, the following: if there exists a particular mechanism that partially implements a predetermined goal, then there exists a *direct revelation mechanism* that truthfully implements it.[3] The undesired equilibria are then often disregarded on the basis of the equilibrium with truthful revelation being the *salient* equilibrium. This is pointed out in Postlewaite and Schmeidler (1986) as follows:

> [*The partial* (*direct revelation*) *implementation*] does assure that the resulting outcome will be an equilibrium of some game; however, there may be others as well. This problem is sometimes dismissed with an argument that as long as truthful revelation is an equilibrium, it will somehow be the salient equilibrium even if there are other equilibria as well.

In our environment, we show that the revelation principle fails and the above argument loses its relevance, implying that partial implementation loses its practical appeal: One cannot restrict attention to direct revelation mechanisms without a loss of generality. This is why focusing on full implementation rather than partial implementation becomes crucial in our setup.

The concept of EPE, the notion of equilibrium of the current paper, is well-suited to our environment: it is belief-free, does not require any belief updating or any expectation considerations, and is robust to informational assumptions regarding the environment.

On the other hand, the concept of Bayesian Nash equilibrium is not suited to our setup. This is because the notion of Bayesian Nash equilibrium employs an aggregation of individuals' welfare in different states of the world using the associated probabilities. At the very least, this necessitates the need for complete and transitive preferences over the set of certain outcomes in order to obtain a utility representation. However, this is neither coherent nor consistent in a setting in which individuals' choices over certain outcomes (alternatively, degenerate acts) do not necessarily obey the standard axioms of rationality. Indeed, in our environment, individuals' choices over certain outcomes may not even be representable by well-defined preference relations.[4]

[3] A direct revelation mechanism is a game-form in which each individual's actions consist of a report about his/her own privately observed type.

[4] Please see Footnote 20 for further details.

We provide necessary as well as sufficient conditions for ex-post implementation when individuals' choices need not satisfy the standard axioms of rationality.

Our first result on necessity, Theorem 1, shows that if a mechanism ex-post implements a social choice set (SCS, hereafter), then the *opportunity sets* sustained by this mechanism form a collection of sets with two desirable properties.[5,6] The first of these implies an ex-post choice monotonicity condition (shown in Proposition 1), while the second implies a pseudo ex-post incentive compatibility (established in Proposition 2).[7]

An important implication of our result on necessity is that the revelation principle holds whenever individuals' choices satisfy the independence of irrelevant alternatives.

We provide sufficiency results for the case of three or more individuals and the case of two individuals, separately. With three or more individuals, we present two methods to strengthen our necessary conditions to deliver sufficiency: The first, presented in Theorem 2, involves *choice incompatibility*, a mild condition that requires some level of disagreement among the individuals at every state, and it performs a similar task as that done by the economic environment assumption in the rational domain.[8] Theorem 3 presents the second method in which we employ a combination of our necessary conditions and a choice counterpart of the no-veto-power property.

The mechanism we employ in the case of two individuals is novel as it features differences when compared with its counterparts in the rational domain. Furthermore, we provide two routes to strengthen the associated necessary conditions into sufficient conditions. These, presented in Theorems 5 and 6, require either some sort of choice

[5]The opportunity set of an individual consists of the alternatives that he/she can obtain by changing his/her messages, while those of the opponents remain the same.

[6]We refer to such family of sets as the collection of sets *consistent* with the given SCS under incomplete information. Each member of this collection of sets is associated with an individual and a social choice function (SCF) in the SCS and a type profile of the other individuals with the property that each such set is independent of the message (in the mechanism) chosen by the individual whom this set is associated with. Moreover the following hold: (1) Given any individual and any one of his/her particular types and any SCF in the SCS and any type profile for the other individuals, it must be that the individual's choices under the resulting type profile contain the alternative that corresponds to the outcome of the SCF for the same type profile; and (2) whenever there is a deception that leads to an outcome that is not compatible with the SCS, there exist an informant state (i.e., a type profile) and an informant individual such that this individual does not choose at the informant state the alternative generated by this deception from his/her set given that he/she is as in the informant state while others' types are identified via their deception from the informant state.

[7]It is useful to point out that the pseudo ex-post incentive compatibility is a result of item (1) of Footnote 6, while item (2) of the same footnote implies our ex-post choice monotonicity condition.

[8]The choice incompatibility required for sufficiency is satisfied when there exist a set of alternatives and a collection of choice sets described as in the necessity direction with the following two properties: (1) this set of alternatives contains the union of the collection of choice sets; and (2) for each alternative in this set, there exist at least two individuals who do not choose that alternative from this set.

incompatibility or some sort of choice unanimity.

Another contribution of our paper concerns the simplicity of the mechanisms needed for implementation, a topic of recent interest [see e.g., Li (2017) and Borgers and Li (2018)]. Naturally, simplicity becomes a bigger concern when dealing with individuals having cognitive limitations. In Theorem 7, we identify lower bounds on the number of messages that is required from a mechanism that ex-post implements a given SCS. Our result, therefore, provides a better understanding of the scope of a well-known criticism of the mechanism design literature, which involves the argument that, often, mechanisms employed are complicated and thus do not offer much practical appeal.

In Section 2, we discuss some related literature. Section 3 presents a motivating example illustrating the difficulties associated with the desired construction. In Section 4, we provide the notation and the definitions. Section 5 contains the necessity and sufficiency results concerning the case with three or more individuals, while Section 6 contains the same for the case with two individuals. In Section 7, we discuss simple mechanisms and present lower bounds on the number of messages needed, and Section 8 concludes. Meanwhile, the proofs are presented in the Appendix.

2 Related Literature

Besides de Clippel (2014), our paper is mostly related to Bergemann and Morris (2008), which analyzes ex-post implementation in the rational domain.[9] In a sense, our paper can be thought of as an envelope of de Clippel (2014) and Bergemann and Morris (2008). We extend de Clippel (2014)'s analysis to the case of incomplete information and Bergemann and Morris (2008)'s analysis to the case where individual choices' need not satisfy the standard axioms of rationality. A difference of note is that we provide a novel analysis for the case of two individuals in our setup, whereas Bergemann and Morris (2008) provides their analysis only for the case of three or more individuals. On the other hand, de Clippel (2014) discusses a modification of his sufficiency results in order to obtain sufficiency for the case of two individuals, and it involves the use of an argument similar to the one from Moore and Repullo (1990).

Another related paper is Jackson (1991), which analyzes Bayesian implementation for the case of three or more individuals in the rational domain. Jackson (1991) generalizes the analysis of Maskin (1999) (on Nash implementation under complete information) to

[9]Some of the other influential work on ex-post implementation and robust mechanism design in the rational domain include Bergemann and Morris (2005), Jehiel, Meyer-ter Vehn, Moldovanu, and Zame (2006), Jehiel, Meyer-ter Vehn, and Moldovanu (2008), and Bergemann and Morris (2009).

the case of incomplete information. In this sense, what Jackson (1991) is to the seminal work in Maskin (1999), our paper is to de Clippel (2014).[10]

Hurwicz (1986), Korpela (2012), and Ray (2018) have also investigated the problem of implementation under complete information when individual choices do not have to satisfy the standard axioms of rationality. Hurwicz (1986) considers choices that can be represented by a well-defined preference relation which does not have to be acyclic. On the other hand, Korpela (2012) shows that when individual choices fail rationality axioms, independence of irrelevant alternatives, also known as Sen's α, is key to obtaining the necessary and sufficient condition synonymous to that of Moore and Repullo (1990) (the so-called Condition μ) under complete information.

There have been other attempts at investigating the problem of implementation under complete information that allow for "non-rational" behavior of individuals. Eliaz (2002) provides an analysis of implementation when some of the individuals might be "faulty" and hence fail to act optimally. An earlier paper of ours, Barlo and Dalkiran (2009), provides an analysis of implementation for the case of epsilon-Nash equilibrium, i.e., when individuals are satisfied by getting close to (but not necessarily achieving) their best responses. Glazer and Rubinstein (2012) provides a mechanism design approach where the content and the framing of the mechanism affect individuals' ability to manipulate their information.[11]

In the rational domain, Ohashi (2012) provides sufficiency results for ex-post implementation with two individuals in an environment that is economic and has a bad outcome. Our sufficiency results for the case of two individuals differ with those of Ohashi (2012) in three dimensions: (i) we allow for non-economic environments, (ii) we do not require the existence of a bad outcome, and (iii) we allow individuals' choices to violate the standard axioms of rationality.

3 Motivating Example

The following example aims to display the intricacies concerning the design of a mechanism which implements a generalized welfare notion [strict generalized Pareto optimality due to Bernheim and Rangel (2009)] in EPE with two individuals whose

[10] Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1987) also provide analyses of full implementation under incomplete information allowing for different informational assumptions. Indeed, there is a large literature on implementation, and it would not be possible to mention many interesting work here. Instead, we refer the interested reader to surveys such as Maskin (1985), Moore (1992), Jackson (2001), Maskin and Sjöström (2002), Palfrey (2002), and Serrano (2004).

[11] Some of the other related work include Cabrales and Serrano (2011) and Saran (2011).

choices do not satisfy the standard axioms of rationality (alternatively, the weak axiom of revealed preferences abbreviated to WARP henceforth).[12] In fact, the individuals' choices involve three types of well-known behavioral biases: (1) attraction effect, (2) status-quo bias, and (3) Condorcet cycles. Below, we discuss these biases before going into the details of the example.

Attraction effect:

One of the commonly observed behavioral aspects implying a violation of the standard rationality assumptions is *the attraction effect*:[13] Decoy alternatives, alternatives that are known to be dominated by other alternatives, can cause preference reversals when they are introduced into the choice set. When an alternative (the decoy) is inferior to a particular alternative in terms of all relevant attributes yet at the same time is superior in some attributes and inferior in others than another alternative, an individual facing these alternatives altogether might be inclined to choose the non-dominated alternative. A relatively well-known real-life example makes it clearer: Consider the following options that were presented for subscription to *The Economist* Magazine at some point: Option (A) Internet-only subscription for \$59, Option (B) Print-only subscription for \$125, Option (C) Print-and-Internet subscription for \$125. One would immediately realize that Option B is dominated in all relevant attributes by Option C. At first, it might even seem puzzling to see an alternative like Option B (the decoy), which one would never expect to be chosen. Yet, it is common sense that anyone who is offered all three options would be inclined towards Option C.

Herne (1997) demonstrates how the presence of a decoy alternative causes the attraction effect in a policy-making context. One of her findings is in part the motivation behind our example: In September 1993, Finland took the decision to build a new nuclear power plant to a parliamentary vote. The majority of the opponents of nuclear power favored the alternative of decentralized solar power plants. Even though it was not on the table at all, the supporters of the nuclear power plant used coal power plants as a point of comparison to nuclear power plants. Nuclear power dominated coal as it

[12] Sen (1971) shows that a choice correspondence satisfies the WARP (and be represented by a complete and transitive preference relation) if and only if it satisfies independence of irrelevant alternatives (referred to as IIA or Sen's α) and an expansion consistency axiom (known as Sen's β). We refer to these as the *standard axioms of rationality*. Formally, we say that the individual choice correspondence $C : \mathcal{X} \to \mathcal{X}$ satisfies Sen's α if whenever $x \in S \subset T$ for some $S, T \in \mathcal{X}$, $x \in C(T)$ implies $x \in C(S)$. Meanwhile, we say that the individual choice correspondence $C : \mathcal{X} \to \mathcal{X}$ satisfies Sen's β if $x, y \in S \subset T$ for some $S, T \in \mathcal{X}$, and $x, y \in C(S)$ implies $x \in C(T)$ if and only if $y \in C(T)$.

[13] A seminal paper for the attraction effect is Huber, Payne, and Puto (1982). See also Ok, Ortoleva, and Riella (2015).

was more environment friendly and more reliable, at the time, in terms of stability and price. On the other hand, solar power was better for the environment when compared to both nuclear power and coal. However, the high costs of solar panels and intermittency made it less appealing than nuclear power and coal in terms of reliability. That is, coal was dominated by nuclear power in environment and reliability dimensions, but solar power dominated coal only in the dimension of environment. In this case, the supporters of nuclear power deliberately used coal as a decoy alternative in the sense that it was not intended to be implemented but was presented in the consideration set in order to increase the attractiveness of nuclear power. That is, coal was asymmetrically dominated by nuclear power in the presence of solar power.

Therefore, it was expected by the supporters of the nuclear power plant that *nuclear* would be chosen from the grand set $\{coal, nuclear, solar\}$, whereas *solar* would be chosen from the set $\{nuclear, solar\}$. Such a choice violates the WARP, in particular the IIA.[14]

Status-quo bias:

Another well-documented behavioral aspect that we observe in real life is *the status-quo bias*.[15] It is well-documented that when individuals face new alternatives to replace a status-quo they have a tendency to keep the status-quo unless it is fully dominated by one of the alternatives in all relevant attributes. A hypothetical example of many developing countries is as follows: Suppose the status-quo source of energy in a country is *coal* and the country is considering whether to switch to another type of energy. Suppose further that the options on the table are *nuclear* and *solar*. Since *nuclear* dominates *coal* (the status-quo) in the dimensions of environment and reliability, one might expect the country to choose *nuclear* from the grand set $\{coal, nuclear, solar\}$, whereas *coal* might be chosen from the set $\{coal, solar\}$ since *solar* does not dominate *coal* (the status-quo) in all relevant dimensions. Such a choice, by itself, does not violate the WARP. Yet there might not be a clear winner between *nuclear* and *solar* when staying with the status-quo is not an option, i.e., the choice from the set $\{nuclear, solar\}$ might be both *nuclear* and *solar*. Then, the WARP (in particular, Sen's β) would not hold.[16]

Groups as participants (Condorcet cycles):

It is well-known that when more than three options are voted on in a pairwise fashion by individuals, each of whom has "rational" preference orders, the preferences aggregated

[14] Please see Footnote 12 in order to see the associated formal definitions.

[15] A seminal paper for status-quo bias is Samuelson and Zeckhauser (1988), see also Kahneman, Knetsch, and Thaler (1991), Masatlioglu and Ok (2005), and Dean, Kıbrıs, and Masatlioglu (2017).

[16] The definition of this axiom is presented in Footnote 12.

may end up having Condorcet cycles.[17] In particular, one might observe *nuclear* chosen from the set $\{coal, nuclear\}$; *coal* chosen from the set $\{coal, solar\}$; and *solar* chosen from the set $\{nuclear, solar\}$. Such a choice violates the WARP, as whatever is chosen from the grand set $\{coal, nuclear, solar\}$ would lead to a violation of both the IIA and Sen's β.

Now we are ready to present the details of our example. Suppose that two individuals, whom we refer to as *Alice* and *Bob*, are to decide what type of energy to employ or jointly invest in, be it *coal* energy, *nuclear* energy, or *solar* energy.[18,19]

Suppose that the revealed preferences of Alice and Bob are not necessarily rational. That is, Alice and Bob's individual choices from different subsets of the grand set $X = \{coal, nuclear, solar\}$ may violate the WARP.

Let the set of all relevant states of the world regarding these individual choices be given by Θ. We assume that there is incomplete information regarding the true state of the world $\theta \in \Theta$. In particular, the true state of the world is distributed knowledge between Alice and Bob. That is, the set of all relevant states of the world, Θ, has a product structure, i.e., $\Theta = \Theta_A \times \Theta_B$. When the true state of the world is $\theta = (\theta_A, \theta_B)$, Alice is informed only of the θ_A component of the true state of the world, whereas Bob is informed only of the θ_B component of the true state of the world. Suppose that Alice and Bob have two possible types each, denoted by $\Theta_i = \{\rho_i, \gamma_i\}$ for $i \in \{A, B\}$. So the set of all possible states of the world is given by $\Theta = \{(\rho_A, \rho_B), (\rho_A, \gamma_B), (\gamma_A, \rho_B), (\gamma_A, \gamma_B)\}$.

The individual choices of Alice and Bob at state $\theta \in \Theta$ are described by the choice correspondences, $C_A^\theta : \mathcal{X} \to \mathcal{X}$, and $C_B^\theta : \mathcal{X} \to \mathcal{X}$, where $\mathcal{X}$ denotes the set of non-empty subsets of X and $C_i^\theta(S) \subseteq S$ for each $S \in \mathcal{X}$ and $i \in \{A, B\}$. Table 1 pinpoints the specific choices to be used in our example with the convention that c stands for *coal*, n for *nuclear* power, and s for *solar* energy.

Let us elaborate on the individual choices of Alice and Bob at each state:

At state (ρ_A, ρ_B), Alice's choices can be rationalized by the preference relation $n \succ_A c \sim_A s$, and Bob's choices can be rationalized by the preference relation $s \succ_B n \succ_B c$.

The identical choices of Alice and Bob at (ρ_A, γ_B) can be explained by the attraction

[17] Hurwicz (1986) investigates the problem of implementation when individuals represent groups of rational agents. On the other hand, cyclic or intransitive preferences may also arise when individuals are regular human beings (Tversky, 1969).

[18] Alice and Bob can also be interpreted as region A and region B within the same legislation, such as two states in the U.S. or two countries in the E.U.

[19] In his Nobel Prize Lecture "Mechanism Design: How to Implement Social Goals" (December 8, 2007), Eric Maskin provides an example in which an energy authority "is charged with choosing the type of energy to be used by Alice and Bob" in the rational domain under complete information.

S	$C_A^{(\rho_A,\rho_B)}$	$C_B^{(\rho_A,\rho_B)}$	$C_A^{(\rho_A,\gamma_B)}$	$C_B^{(\rho_A,\gamma_B)}$	$C_A^{(\gamma_A,\rho_B)}$	$C_B^{(\gamma_A,\rho_B)}$	$C_A^{(\gamma_A,\gamma_B)}$	$C_B^{(\gamma_A,\gamma_B)}$
$\{c,n,s\}$	$\{n\}$	$\{s\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{c\}$	$\{c,s\}$	$\{n,s\}$
$\{c,n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{c\}$	$\{n\}$	$\{c\}$
$\{c,s\}$	$\{c,s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{c\}$	$\{c\}$	$\{c\}$	$\{s\}$
$\{n,s\}$	$\{n\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{n,s\}$	$\{n,s\}$	$\{s\}$	$\{s\}$

Table 1: Individual choices of Alice and Bob.

effect, as discussed above. That is, *coal* can be thought of as the decoy alternative where the relevant attributes are environment and reliability. As in the Finland power plant example, Alice and Bob choose *nuclear* from the grand set $\{coal, nuclear, solar\}$ and *solar* from the set $\{nuclear, solar\}$. They also choose *nuclear* from the set $\{coal, nuclear\}$. This means Alice and Bob individually choose *nuclear* whenever it is presented with *coal*, the decoy option. Yet, whenever *coal* is not available they choose *solar* over *nuclear*. These also show that at state (ρ_A, γ_B), their individual choices cannot be rationalized by a complete and transitive preference relation, as they violate the IIA.[20] We would like to emphasize that we allow individual choices to be *interdependent*: between (ρ_A, ρ_B) and (ρ_A, γ_B), Alice's private information (type) does not change; yet, the choice behavior of Alice is not identical at these states.[21]

On the other hand, at state (γ_A, ρ_B), Bob's choices can be rationalized by the preference relation $c \succ_B s \sim_B n$, whereas Alice's choices feature a status-quo bias where the status-quo is *coal*. Similar to the hypothetical status-quo bias example discussed above, Alice chooses *nuclear* from the grand set $\{coal, nuclear, solar\}$ and *coal* from the set $\{coal, solar\}$. Yet, when faced with the set $\{nuclear, solar\}$ Alice is indifferent and chooses both, possibly because *nuclear* and *solar* do not dominate each other in all

[20]In fact, there is no *well-defined preference relation* representing these choices. We wish to define what we mean by a choice correspondence being represented by a well-defined preference relation. To that regard, for any given individual choice correspondence $C : \mathcal{X} \to \mathcal{X}$, let $\succeq^C$ be the induced preference relation and be defined by: $x \succeq^C y$ if and only if there exists $S \in \mathcal{X}$ with $x, y \in S$ and $x \in C(S)$. On the other hand, given a preference relation $\succeq$ on X, the induced normal choice correspondence $C^{\succeq} : \mathcal{X} \to \mathcal{X}$ is defined by $C^{\succeq}(S) = \{x \in S : x \succeq y \text{ for all } y \in S\}$ for $S \in \mathcal{X}$. We say that the individual choice correspondence $C : \mathcal{X} \to \mathcal{X}$ is represented by a well-defined preference relation $\succeq^C$ if C equals $C^{\succeq^C}$. Further, Theorem 9 of Sen (1971) in the current setting says that a choice correspondence can be represented with a well-defined preference relation (which is not necessarily transitive) if and only if the choice correspondence satisfies Sen's α and γ. While Sen's α is defined in Footnote 12, a choice correspondence $C : \mathcal{X} \to \mathcal{X}$ satisfies Sen's γ if $x \in C(S) \cap C(T)$ for some $S, T \in \mathcal{X}$ implies $x \in C(S \cup T)$. It can easily be verified that at state (ρ_A, γ_B) the individuals' choices satisfy neither Sen's α nor γ.

[21]We note that even though Alice does not know Bob's private information (type), she knows the set of all possible types for Bob. Therefore, Alice might consider what she were to choose contingent upon each possible type of Bob. This is especially relevant when the information in the hands of Bob is relevant for Alice's choices as in the case of a common value auction.

relevant dimensions. This also proves that Alice's choices cannot be rationalized by a complete and transitive preference relation as they violate Sen's β.[22]

Finally, at state (γ_A, γ_B), neither of the individual choices can be rationalized by a complete and transitive preference relation because the individual choices of Alice and Bob violate the IIA and Sen's β.[23] Furthermore, Alice's choices lead to a Condorcet cycle as discussed above. In particular, Alice chooses *nuclear* from the set $\{coal, nuclear\}$, *coal* from the set $\{coal, solar\}$, and *solar* from the set $\{nuclear, solar\}$. Such a pattern may arise when Alice makes her choices by consulting *a group of individuals*, such as pairwise voting with her parents, or a parliamentary vote.

Next comes the social choice notion, the generalized welfare criterion developed by Bernheim and Rangel (2009).[24] This welfare criterion provides a choice theoretic foundation for behavioral welfare economics as it is directly based on individual choices.

Following Bernheim and Rangel (2009), we say that an alternative x is *strictly unambiguously chosen over* another alternative z, if z is never chosen whenever x is available. On the other hand, an alternative x is *weakly unambiguously chosen over* another alternative z, if whenever they are both available, z is never chosen *unless* x is chosen as well. These deliver an intuitive way of extending the notion of Pareto efficiency beyond the rational domain:

An alternative x is a *strict generalized Pareto optimum* if there does not exist any other alternative y, such that y is weakly unambiguously chosen over x for every individual, and y is strictly unambiguously chosen over x for some individual(s). We refer to a strict generalized Pareto optimum alternative as a *BR-optimal* outcome.

The social planner who faces the individual choices of Alice and Bob does not know the true state of the world but cares about their welfare according to the welfare notion of Bernheim and Rangel (2009). Thus, the planner aims to provide Alice and Bob a state contingent allocation which is BR-optimal at every state.[25]

When individual choices can be rationalized with a complete and transitive pref-

[22]Even though Sen's β fails, Alice's individual choice correspondence can be represented by a well-defined (but intransitive) preference relation as both Sen's α and γ hold. See Footnotes 12 and 20 for the definitions.

[23]At this state, neither of the individual's choices can be represented by a well-defined preference relation as the IIA is violated for both of the individuals. For more, please see Footnotes 12 and 20.

[24]Another paper that provides a welfare analysis that is in line with non-rational choices is Rubinstein and Salant (2011).

[25]BR-optimal alternatives are defined under certainty. There is not an easy way to generalize this notion of efficiency to the case of uncertainty as the individual choices of Alice and Bob violate the standard rationality axioms and hence the expected utility hypothesis. So, the goal of the social planner can be thought of as obtaining an ex-post strict generalized Pareto optimal state contingent allocation.

erence relation, the BR-optimal outcomes are the same as the standard strict Pareto optimal outcomes. Therefore, at (ρ_A, ρ_B), the BR-optimal outcomes are *nuclear* and *solar*. On the other hand, at (ρ_A, γ_B), since *coal* is never chosen (except when it is offered as a singleton), it is easy to see that the BR-optimal outcomes are *nuclear* and *solar*.[26] At (γ_A, ρ_B), *coal* is strictly unambiguously chosen over *solar* by both Alice and Bob.[27] Hence, *solar* is not BR-optimal at (γ_A, ρ_B). Even though *coal* is also strictly unambiguously chosen over *nuclear* by Bob at (γ_A, ρ_B), *coal* is not weakly unambiguously chosen over *nuclear* by Alice, since Alice chooses *nuclear* from the set $\{coal, nuclear\}$ at (γ_A, ρ_B). Thus, the BR-optimal outcomes at (γ_A, ρ_B) are *coal* and *nuclear*. Finally, at (γ_A, γ_B), *solar* is strictly unambiguously chosen over *nuclear* by Alice and it is weakly unambiguously chosen over *nuclear* by Bob.[28] So, *nuclear* is not BR-optimal at (γ_A, γ_B). Both *coal* and *solar* are BR-optimal at (γ_A, γ_B), since from the set $\{coal, solar\}$ Alice chooses *coal* and Bob *solar* at (γ_A, γ_B).

The BR-optimal alternatives contingent on the states are summarized in Table 2. As in Palfrey and Srivastava (1987), a social choice set (SCS) refers to a selection of

State	(ρ_A, ρ_B)	(ρ_A, γ_B)	(γ_A, ρ_B)	(γ_A, γ_B)
BR-optimal alternatives	$\{n, s\}$	$\{n, s\}$	$\{c, n\}$	$\{c, s\}$

Table 2: BR-optimal alternatives.

state contingent allocations. In fact, the BR-optimal alternatives form a collection of state contingent allocations and we refer to it as the *BR-optimal SCS*. In what follows, the planner aims to implement a particular mutually exhaustive selection from the BR-optimal SCS, which is given by $F = \{f, f'\}$, described in Table 3: F is mutually

State	(ρ_A, ρ_B)	(ρ_A, γ_B)	(γ_A, ρ_B)	(γ_A, γ_B)
f	n	n	n	s
f'	s	s	c	c

Table 3: The social choice set F.

exhaustive as $\{f(\theta)\} \cup \{f'(\theta)\}$ equals the set of BR-optimal outcomes at every $\theta \in \Theta$.[29]

[26] We also note that at (ρ_A, γ_B), even though *nuclear* is strictly unambiguously chosen over *solar* by Alice at (ρ_A, γ_B), *nuclear* is not weakly unambiguously chosen over *solar* by Bob since Bob chooses *solar* from the set $\{nuclear, solar\}$ at (ρ_A, γ_B).

[27] Whenever *coal* is present, *solar* is never chosen by both Alice and Bob at (γ_A, ρ_B).

[28] At (γ_A, γ_B), Alice never chooses *nuclear* when *solar* is available; Bob chooses *nuclear* only if he chooses *solar* as well.

[29] We note that there is no particular reason other than simplicity for choosing the mutually exhaustive selection $\{f, f'\}$ as the SCS F. In general, the design of a mechanism would depend on the particular SCS under consideration.

3.1 The mechanism

A mechanism makes Alice and Bob send individual messages to the social planner and describes the outcome to be implemented as a function of these messages. We consider the following mechanism: messages available to Alice and Bob are given by $M_A = \{U, M, D\}$ and $M_B = \{L, M, R\}$, respectively; the outcome function that maps messages to alternatives is represented by $g : M \to X$ and is described in Table 4. This

		Bob		
		L	M	R
Alice	U	n	c	n
	M	c	s	c
	D	n	s	s

Table 4: The mechanism μ for Alice and Bob.

mechanism is denoted by $\mu = (M, g)$, with $M = M_A \times M_B$ and $g : M \to X$.

In what follows, we show that this mechanism implements the aforementioned goal of the social planner in EPE. But first, we turn to the Nash equilibrium (NE) outcomes of $\mu = (M, g)$.

de Clippel (2014) points out an intuitive and straightforward extension of the notion of NE involving individuals' choices that cannot be rationalized by a complete and transitive preference relation. The intuition is as follows: for each individual, the equilibrium outcome should be among the chosen within the set of alternatives he/she can generate by unilateral deviations. This intuition is aligned with the opportunity criterion of Sugden (2004) in that the set of alternatives an individual is free to choose from, i.e., the opportunity set of an individual, is determined in a mechanism by the messages of the other individuals.

We follow de Clippel (2014) and denote the opportunity sets of Alice and Bob in our mechanism by $O^{\mu}_A(m_B) := \{g(m_A, m_B) | m_A \in M_A\}$ and $O^{\mu}_B(m_A) := \{g(m_A, m_B) | m_B \in M_B\}$, respectively. We are now ready to give a formal definition of an NE of our mechanism at a given state of the world.

We say that $m^* = (m^*_A, m^*_B)$ is a *Nash equilibrium* of the mechanism $\mu = (M, g)$ at θ if $g(m^*) \in C^{\theta}_A(O^{\mu}_A(m^*_B))$ and $g(m^*) \in C^{\theta}_B(O^{\mu}_B(m^*_A))$. Whenever m^* is an NE of μ at θ, we refer to $g(m^*)$ as a *Nash equilibrium outcome* of μ at θ.

Let us exemplify by identifying the NE of our mechanism at state (γ_A, γ_B). Please refer to Table 1 for the individuals' choices at (γ_A, γ_B).

If Alice sends the message U, Bob can unilaterally generate the set $\{c, n\}$ under the mechanism μ, i.e., $O^{\mu}_{B}(U) = \{c, n\}$. Bob chooses c from the set $\{c, n\}$ at (γ_A, γ_B), which implies that Bob finds it optimal to send the message M. The best response action M chosen by Bob against Alice's message U is depicted in Table 5 by a superscript B in cell (U, M). Similarly, when Alice sends the message M, Bob can unilaterally generate the set $\{c, s\}$ under the mechanism μ, i.e., $O^{\mu}_{B}(M) = \{c, s\}$, and Bob chooses s from the set $\{c, s\}$ at (γ_A, γ_B). Thus, Bob finds it optimal to send the message M against Alice's action M. Finally, if Alice sends the message D, Bob can unilaterally generate the set $\{n, s\}$ under the mechanism μ, i.e., $O^{\mu}_{B}(D) = \{n, s\}$. Bob chooses s from the set $\{c, n\}$ at (γ_A, γ_B); hence, both M and R are the best responses for Bob.

On the other hand, when Bob sends the message L, Alice can unilaterally generate the set $\{c, n\}$ under the mechanism μ, i.e., $O^{\mu}_{A}(L) = \{c, n\}$. Alice chooses n from the set $\{c, n\}$ at (γ_A, γ_B). Therefore, her best responses to Bob sending message L consist of U and D, which are indicated in Table 5 by a superscript A in cells (U, L) and (D, L). If Bob sends the message M, Alice can unilaterally generate the set $\{c, s\}$ under the mechanism μ, i.e., $O^{\mu}_{A}(M) = \{c, s\}$. Alice chooses c from the set $\{c, s\}$ at (γ_A, γ_B). Finally, when Bob sends the message R, Alice can unilaterally generate the set $\{c, n, s\}$ under the mechanism μ, i.e., $O^{\mu}_{A}(R) = \{c, n, s\}$. Alice chooses both c and s from the set $\{c, n, s\}$ at (γ_A, γ_B).

The resulting best responses are summarized in Table 5, which shows that the NE of the mechanism μ at state (γ_A, γ_B) are the message profiles (U, M) and (D, R). Hence, the corresponding NE outcomes at (γ_A, γ_B) are c and s.

	L	M	R
U	${}^{A}n$	${}^{A}\boxed{c}^{B}$	n
M	c	s^{B}	${}^{A}c$
D	${}^{A}n$	s^{B}	${}^{A}\boxed{s}^{B}$

Table 5: The best responses and Nash equilibria of the mechanism at (γ_A, γ_B).

Repeating this exercise, one can easily show that the NE and NE outcomes of our mechanism at other states of the world are as presented in Table 6 (where NE message profiles are depicted using circles in the corresponding cells).

Going over Tables 2 and 6 reveals that the set of BR-optimal outcomes and the set of NE outcomes of our mechanism coincide at every state of the world. Therefore, if the true state of the world were common knowledge between Alice and Bob, our mechanism would be (fully) implementing the BR-optimal outcomes in NE at every state.

State: (ρ_A, ρ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	s	s

NE outcomes: $\{n, s\}$

State: (ρ_A, γ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	s	s

NE outcomes: $\{n, s\}$

State: (γ_A, ρ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	s	s

NE outcomes: $\{c, n\}$

State: (γ_A, γ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	s	s

NE outcomes: $\{c, s\}$

Table 6: Nash equilibria and Nash equilibrium outcomes of the mechanism.

3.2 Ex-post equilibrium outcomes of the mechanism

With incomplete information, the state of the world is distributed knowledge between Alice and Bob as they can only observe their own types before sending their messages. Thus, their plans of actions (strategies) can depend only on their own types and not the whole state of the world.

Therefore, under incomplete information, the strategies of Alice and Bob in the mechanism μ should be measurable with respect to their private information: a strategy for Alice and Bob in μ is a function $\sigma_i : \Theta_i \to M_i$ for $i \in \{A, B\}$.

There is not a clear way of defining a Bayesian Nash Equilibrium of the mechanism μ in our example. The main reason is not because we have not specified any beliefs over Θ, but because the individual choices of Alice and Bob violate standard rationality axioms concerning certain outcomes: at states (ρ_A, γ_B) and (γ_A, γ_B), Alice and Bob do not have a well-defined preference relation on X, while at state (γ_A, ρ_B), Alice's choices can be represented with a well-defined but intransitive preference relation on X.[30] Thus, Alice and Bob cannot be modeled as (expected) utility maximizers.

Similarly, as there is no clear way to evaluate an individual's well-being with mixed strategies in our setup, we restrict our attention to pure strategies in the mechanism μ.

Fortunately, a pure strategy EPE of our mechanism is belief-free and does not require any expectation considerations. We provide the definition of an EPE of our mechanism:

We say that the strategy profile $\sigma^* = (\sigma_A^*, \sigma_B^*)$ is an *ex-post equilibrium of the mechanism* $\mu = (M, g)$ if for all $\theta \in \Theta$, $g(\sigma^*(\theta)) \in C_A^\theta(O_A^\mu(\sigma_B^*(\theta_B)))$ and $g(\sigma^*(\theta)) \in C_B^\theta(O_B^\mu(\sigma_A^*(\theta_A)))$. In words, an EPE requires that the strategies of Alice and Bob induce an NE of the mechanism μ at every state of the world and that they are measurable with respect to their private information.

The following shows that there are three EPE of our mechanism, two of which are

[30] Please see Footnotes 12, 20, 22, and 23.

equivalent in terms of the outcomes they generate:

Claim 1. *The strategy profiles $\sigma'^* = (\sigma'^*_A, \sigma'^*_B)$, $\sigma''^* = (\sigma''^*_A, \sigma''^*_B)$, and $\sigma'''^* = (\sigma''^*_A, \sigma'''^*_B)$ described below are the only EPE of the mechanism $\mu = (M, g)$, where the outcomes generated under σ''^* and σ'''^* are equivalent (i.e., $g(\sigma''^*(\theta)) = g(\sigma'''^*(\theta))$ for each $\theta \in \Theta$).*

$$\begin{array}{lllll} \sigma'^* & : & \sigma'^*_A(\rho_A) = U \quad \sigma'^*_A(\gamma_A) = D & \text{and} & \sigma'^*_B(\rho_B) = L \quad \sigma'^*_B(\gamma_B) = R, \\ \sigma''^* & : & \sigma''^*_A(\rho_A) = D \quad \sigma''^*_A(\gamma_A) = U & \text{and} & \sigma''^*_B(\rho_B) = M \quad \sigma''^*_B(\gamma_B) = M, \\ \sigma'''^* & : & \sigma'''^*_A(\rho_A) = M \quad \sigma'''^*_A(\gamma_A) = U & \text{and} & \sigma'''^*_B(\rho_B) = M \quad \sigma'''^*_B(\gamma_B) = M. \end{array}$$

Below we focus on two EPE, σ'^* and σ''^*, as σ''^* and σ'''^* correspond to the same outcomes under μ. In Table 7, we summarize the EPE outcomes of μ under σ'^* and σ''^* (where message profiles corresponding to σ'^* are depicted using circles while those associated with σ''^* are indicated using squares in the corresponding cells).

State: (ρ_A, ρ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	$\boxed{s}$	s

EPE outcomes: $\{n, s\}$

State: (ρ_A, γ_B)

	L	M	R
U	n	c	n
M	c	s	c
D	n	$\boxed{s}$	s

EPE outcomes: $\{n, s\}$

State: (γ_A, ρ_B)

	L	M	R
U	n	$\boxed{c}$	n
M	c	s	c
D	n	s	s

EPE outcomes: $\{c, n\}$

State: (γ_A, γ_B)

	L	M	R
U	n	$\boxed{c}$	n
M	c	s	c
D	n	s	s

EPE outcomes: $\{c, s\}$

Table 7: Ex-post equilibria and ex-post equilibrium outcomes of the mechanism.

Tables 2 and 7 show that the set of BR-optimal outcomes and the set of EPE outcomes of μ coincide. Referring to Table 3 which describes the SCS F, we also observe that $g(\sigma'^*(\theta)) = f(\theta)$ for each $\theta \in \Theta$, and $g(\sigma''^*(\theta)) = f'(\theta)$ for each $\theta \in \Theta$ (while $g(\sigma'''^*(\theta)) = f'(\theta)$ for all $\theta \in \Theta$ holds trivially). That is, (1) each social choice function (SCF) in the SCS induces the same outcomes under a particular EPE of the mechanism μ; and (2) for each EPE of the mechanism μ, there exists a particular SCF in the SCS that induces the same outcomes state by state. Thus, μ (fully) ex-post implements the SCS F. Indeed, in Section 7.1 we show that the mechanism μ is the "simplest mechanism" ex-post implementing the SCS F.

3.3 The revelation principle fails

Now we show that the revelation principle (for partial implementation) fails in our example. Recall that the revelation principle implies if there exists a mechanism that partially (ex-post) implements a particular SCF, then there exists a direct revelation mechanism that truthfully partially (ex-post) implements the same SCF.

Consider the SCF f given in Table 3, and notice that $f(\theta)$ is a BR-optimal alternative at every $\theta \in \Theta$. Because σ'^* is an EPE of μ with $g(\sigma'^*(\theta)) = f(\theta)$ for all $\theta \in \Theta$, the mechanism μ *partially* ex-post implements the SCF f. However, the corresponding direct revelation mechanism, $g^d : \Theta \to X$, given in Table 8, fails to partially ex-post implement f truthfully as truthful revelation is not an EPE of g^d.[31] When the state is

		Bob	
		ρ_B	γ_B
Alice	ρ_A	n	n
	γ_A	n	s

Table 8: The direct revelation mechanism g^d.

(ρ_A, γ_B), reporting truthfully delivers n (circled in Table 8). But, the opportunity set of Alice at state (ρ_A, γ_B) is given by $\{n, s\}$ and $n \notin C_A^{(\rho_A,\gamma_B)}(\{n, s\}) = \{s\}$.[32]

Therefore, the revelation principle may fail with individuals whose choices do not satisfy the standard axioms of rationality. Hence, the argument saying that as long as truthful revelation is an equilibrium it will be the salient one, even when there are other equilibria, no longer holds. This, in turn, means partial implementation loses its practical appeal when individuals' choices do not comply with the standard axioms of rationality. In Section 5, we show that the key axiom which guarantees the revelation principle is the IIA.

To sum up, our motivating example displays that it is possible to have full ex-post implementation under incomplete information when individuals' choices do not satisfy standard rationality axioms. This is one of the most reasonable ways for behavioral implementation under incomplete information since (i) it is belief free and thus demanding Bayesian considerations are not required, and (ii) the revelation principle might fail and, hence, partial implementation through direct revelation mechanisms loses its practical appeal. Below, we provide necessary as well as sufficient conditions for behavioral (ex-post) implementation under incomplete information. To move forward, we turn to notation and definitions for the general setup.

[31] A direct revelation mechanism is one where the message sets equal the type spaces of individuals. This is why it is enough to specify only an outcome function $g^d : \Theta \to X$ to describe a direct revelation mechanism.

[32] Indeed, misreporting her type as γ_A (and obtaining s) is Alice's best response at state (ρ_A, γ_B) when Bob reports truthfully γ_B.

4 Notation and Definitions

Consider a set of individuals, denoted by $N = \{1, \ldots, n\}$, who have to select an alternative from a non-empty set of alternatives X. Let Θ denote the set of all relevant states of the world regarding the choices of the individuals from (the subsets of) the set of alternatives X. We assume that there is incomplete information among the individuals regarding the true state of the world, and that the true state of the world is distributed knowledge. That is, Θ has a product structure, i.e., $\Theta = \times_{i \in N} \Theta_i$ where $\theta_i \in \Theta_i$ denotes the private information (type) of individual $i \in N$ at state $\theta = (\theta_1, \ldots, \theta_n) \in \Theta$. We also assume that the choice behavior of individual i at state θ is described by the individual choice correspondence $C_i^\theta : \mathcal{X} \to \mathcal{X}$, such that the feasibility requirement of $C_i^\theta(S) \subseteq S$ for all $S \in \mathcal{X}$ holds where $\mathcal{X}$ denotes the set of all non-empty subsets of X. Therefore, the environment we are interested in can be summarized by the tuple $\langle N, X, \Theta, (C_i^\theta)_{i \in N, \theta \in \Theta} \rangle$. We assume that the environment, $\langle N, X, \Theta, (C_i^\theta)_{i \in N, \theta \in \Theta} \rangle$, is common knowledge among the individuals, and that it is known to the designer. We also note that our setup allows (but does not depend on) individual choices to be interdependent. That is, individuals are allowed to choose differently when their own type is fixed but others' are different.[33]

An SCF is a function $f : \Theta \to X$ that specifies a socially optimal alternative—as evaluated by the planner—for each possible state of the world. In other words, f can be viewed as a state contingent allocation. As there may be many socially optimal state contingent allocations that a designer wishes to consider simultaneously, we focus on social choice sets (SCS) rather than SCFs. An SCS, denoted by F, is simply a set of SCFs, i.e., $F \subset \{f | f : \Theta \to X\}$.[34] $\mathcal{F}$ denotes the set of non-empty SCSs.

We denote a mechanism by $\mu = (M, g)$ where M_i denotes the non-empty set of messages available to individual i with $M = \times_{i \in N} M_i$, and $g : M \to X$ describes the outcome function that specifies the alternative to be selected for each message profile.

As in de Clippel (2014), we define the opportunity set of an individual under a mechanism as the set of alternatives that he/she can generate by unilateral deviations given the messages of the other individuals: The *opportunity set* of individual i under $\mu = (M, g)$ for each $m_{-i} \in M_{-i}$ is given by $O_i^\mu(m_{-i}) = \{g(m_i, m_{-i}) \in X : m_i \in M_i\}$. Consequently, an NE of a mechanism at a particular state of the world is defined as

[33] Please see also Footnote 21.

[34] We note that it is customary to denote a social choice rule as an SCS rather than a social choice correspondence under incomplete information. To that regard, we refer to Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1987), Jackson (1991) and Bergemann and Morris (2008).

follows:[35] A message profile m^* is a *Nash equilibrium* of $\mu = (M, g)$ at θ if $g(m^*) \in C_i^{\theta}(O_i^{\mu}(m_{-i}^*))$ for all $i \in N$. The intuition is that for each individual, the NE outcome should be among the chosen alternatives within the set of alternatives the individual can generate by unilateral deviations.

Next, we turn to the case of incomplete information. We note that the mechanism μ in our environment induces an incomplete information game-form. Hence, a strategy for individual i, a contingent plan of actions, has to specify an action for each possible type of i. We denote a strategy of an individual i by $\sigma_i : \Theta_i \to M_i$, under the mechanism $\mu = (M, g)$. As mentioned before, there is no clear way to adopt a Bayesian formulation and employ mixed strategies in our setup because individuals' choices may violate the standard axioms of rationality. So, we restrict attention to pure EPE.

Definition 1. *A strategy profile $\sigma^* : \Theta \to M$ is an* **ex-post equilibrium** *of $\mu = (M, g)$ if for each $\theta \in \Theta$, we have $g(\sigma^*(\theta)) \in C_i^{\theta}(O_i^{\mu}(\sigma_{-i}^*(\theta_{-i})))$ for all $i \in N$.*

In words, an EPE requires that the outcomes generated by the mechanism be NE at every state of the world, while individuals' strategies have to be measurable with respect to only their own types. This delivers the notion of ex-post implementability:

Definition 2. *We say that an SCS $F \in \mathcal{F}$ is* **ex-post implementable** *if there exists a mechanism $\mu = (M, g)$ such that:*

(i) For every $f \in F$, there exists an EPE σ^* of $\mu = (M, g)$ that satisfies $f = g \circ \sigma^*$, i.e., $f(\theta) = g(\sigma^*(\theta))$ for all $\theta \in \Theta$; and

(ii) For every EPE σ^* of $\mu = (M, g)$, there exists $f \in F$ such that $g \circ \sigma^* = f$, i.e., $g(\sigma^*(\theta)) = f(\theta)$ for all $\theta \in \Theta$.

Given an SCS, ex-post implementability demands the existence of a mechanism such that (i) every SCF in the SCS must be sustained by an EPE strategy profile, and (ii) every EPE strategy profile of the mechanism must correspond to an SCF in the SCS. Hence, this is full ex-post implementation. We refer to an SCF f as being *partially ex-post implementable* whenever condition (i) in Definition 2 holds.

Any mechanism that ex-post implements an SCS should take into consideration the private information of the individuals. However, individuals may misreport their private information. This is why we turn to the concept of deception. We denote a deception by

[35] Korpela (2012) refers to this concept as behavioral Nash equilibrium while we follow de Clippel (2014) and designate this notion by NE.

individual i as $\alpha_i : \Theta_i \to \Theta_i$. The interpretation is that $\alpha_i(\theta_i)$ is individual i's reported type. Therefore, $\alpha(\theta) := (\alpha_1(\theta_1), \alpha_2(\theta_2), \ldots, \alpha_n(\theta_n))$ is a profile of reported types, which might be deceptive. We are now ready to move forward with the necessary conditions for ex-post implementation.

5 Three or More Individuals

In this section, we investigate necessary as well as sufficient conditions for ex-post implementation when there are three or more individuals.

We employ the following notion of consistency of a collection of sets under incomplete information in both necessity and sufficiency. As the name evidently suggests, our notion can be viewed as an incomplete information version of the notion of consistency of de Clippel (2014). When the meaning is clear, we refer to this notion simply as *consistency*.

Definition 3. *We say that a non-empty collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) | i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\} \subset \mathcal{X}$ is* **consistent with the SCS** *$F \in \mathcal{F}$* **under incomplete information** *if for every SCF $f \in F$, we have*

(i) *for all $i \in N$, $f(\theta'_i, \theta_{-i}) \in C_i^{(\theta'_i, \theta_{-i})}(S_i(f, \theta_{-i}))$ for each $\theta'_i \in \Theta_i$, and*

(ii) *for any deception profile α with $f \circ \alpha \notin F$, there exists $\theta^* \in \Theta$ and $i^* \in N$ such that $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*})))$.*

A collection of sets $\mathbb{S}$ satisfying consistency with an SCS F under incomplete information obeys the property that $S_i(f, \theta_{-i})$ does not depend on θ_i for all $i \in N$ and $f \in F$ and $\theta_{-i} \in \Theta_{-i}$ and the following hold: (1) Given any $i \in N$ and any $f \in F$ and any $\theta_{-i} \in \Theta_{-i}$, it must be that i's choices when he/she is of type θ'_i at state (θ'_i, θ_{-i}) contains $f(\theta'_i, \theta_{-i})$ for all $\theta'_i \in \Theta_i$; and (2) given any $f \in F$, whenever there is a deception profile α that leads to an outcome not compatible with the SCS, i.e., $f \circ \alpha \notin F$, there exists an informant state θ^* and an informant individual i^* such that i^* does not choose at state θ^* the alternative $f(\alpha(\theta^*))$ (the alternative generated by this deception) from $S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*}))$.

Our first result establishes that consistency with an SCS is a necessary condition for that SCS to be ex-post implementable, and this result is not restricted to the case of three or more individuals.

Theorem 1. *If an SCS $F \in \mathcal{F}$ is ex-post implementable, then there exists a non-empty collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) | i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ consistent with F under incomplete information.*

If a mechanism μ ex-post implements an SCS F, then Theorem 1 establishes that the opportunity sets obtained from the mechanism form a non-empty collection of sets consistent with F. Moreover, our necessary condition implies analogs of the necessary conditions in the rational domain à la Bergemann and Morris (2008): an ex-post-choice monotonicity condition and a quasi-ex-post choice incentive compatibility condition. In what follows, we prove that both of these conditions are necessary for ex-post implementation.

Definition 4. *An SCS $F \in \mathcal{F}$ is* **ex-post choice monotonic** *if, for every SCF $f \in F$ and deception profile α with $f \circ \alpha \notin F$, there is a state $\theta^* \in \Theta$ and an individual $i^* \in N$ and a non-empty set of alternatives $S^* \in \mathcal{X}$ such that*

(i) $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S^*)$, *and*

(ii) $f((\theta'_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*}))) \in C_{i^*}^{(\theta'_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*}))}(S^*)$ *for all* $\theta'_{i^*} \in \Theta_{i^*}$.

Proposition 1. *If there exists a non-empty collection of sets consistent with an SCS $F \in \mathcal{F}$ under incomplete information, then F is ex-post choice monotonic.*

Ex-post choice monotonicity requires that when there is a deception leading to an outcome not compatible with the state contingent allocations allowed by the SCS, there exists an informant state and an informant whistle-blower for this state and an informant reward set for this whistle-blower such that (*i*) the whistle-blower does not choose the outcome arising due to going along with the deception from the reward set at the informant state; and (*ii*) the whistle-blower does not falsely accuse the other individuals of deceiving when the outcome is compatible with the SCS at hand.

Definition 5. *An SCS $F \in \mathcal{F}$ is* **quasi-ex-post choice incentive compatible** *if, for every SCF $f \in F$ and state $\theta \in \Theta$ and individual $i \in N$, there exists a non-empty subset of alternatives $S \in \mathcal{X}$ such that*

(i) $f(\theta) \in C_i^{\theta}(S)$, *and*

(ii) $S \supseteq \{f(\theta'_i, \theta_{-i}) | \theta'_i \in \Theta_i\}$.

Proposition 2. *If there exists a non-empty collection of sets consistent with an SCS $F \in \mathcal{F}$ under incomplete information, then F is quasi-ex-post choice incentive compatible.*

The set $\{f(\theta'_i, \theta_{-i}) : \theta'_i \in \Theta_i\} \in \mathcal{X}$ specifies the *set of alternatives achievable by individual i under an SCF f given others' type profile θ_{-i}*. Quasi-ex-post choice incentive compatibility of an SCS F demands that for every SCF $f \in F$ and for every state $\theta \in \Theta$ and for every individual $i \in N$, there exists a set S from which i chooses $f(\theta)$ at θ while S contains all the alternatives achievable by i under f given θ_{-i}.

Quasi-ex-post choice incentive compatibility describes a necessary condition for partial ex-post implementation of each $f \in F$. However, as we have shown in section 3.3, the revelation principle does not have to hold in our setup. In fact, when the containment relation in (ii) of quasi-ex-post choice incentive compatibility holds strictly, the revelation principle may fail. Consequently, Lemma 1 below identifies a straightforward *necessary and sufficient condition* for the revelation principle when individuals' choices do not satisfy the standard axioms of rationality.

Lemma 1. *An SCF f is partially truthfully (ex-post) implementable in a direct mechanism if and only if for every $\theta \in \Theta$, $i \in N$, we have $f(\theta) \in C_i^{\theta}(\{f(\theta_i', \theta_{-i}) | \theta_i' \in \Theta_i\})$.*

In general, the condition provided in Lemma 1 neither implies nor is implied by the quasi-ex-post choice incentive compatibility condition. Yet, it is easy to see that if the IIA holds, then quasi-ex-post choice incentive compatibility implies the revelation principle.[36]

Proposition 3. *If individual choices satisfy the IIA, then quasi-ex-post choice incentive compatibility implies the revelation principle.*

In summary, these establish that if a mechanism μ partially ex-post implements an SCF f and individuals' choices satisfy the IIA axiom, then there exists a direct revelation mechanism g^d which partially implements the same SCF f in truthful EPE.

Put differently, the revelation principle holds whenever individuals' choices satisfy the IIA.[37]

The failure of the revelation principle when individuals' choices do not satisfy the standard axioms of rationality leads us to search for indirect mechanisms, even for partial implementation. In this context, our results identifying (indirect) mechanisms for full implementation are also useful as full implementation implies partial implementation.

Next, we identify sufficient conditions for ex-post implementation when there are at least three individuals. The sufficient conditions for the case of two individuals is analyzed separately in Section 6.

[36] Please see Footnote 12 for the definition of the IIA (Sen's α).

[37] Saran (2011) studies conditions for revelation principle to hold when individuals have menu-dependent preferences over interim Anscombe-Aumann acts.

Ex-post implementation of an SCS F is not feasible when there is no collection of sets consistent with F. Therefore, the planner should start the design by identifying such collections and then explore additional requirements to be imposed on these collections for sufficiency. Below, we present such new conditions:[38]

Definition 6. *We say that a non-empty set of alternatives $S \in \mathcal{X}$ satisfies the* **choice incompatible pair property** *at state θ if for each alternative $x \in S$ there exist individuals $i, j \in N$ such that $x \notin C_i^\theta(S)$ and $x \notin C_j^\theta(S)$.*

This condition implies some level of disagreement among individuals regarding the socially optimal alternatives at a given state of the world. In words, a set satisfies the choice incompatible pair property at a state, if for each alternative in this set there exists a pair of individuals who do not choose this alternative from this set at that state. Then, any alternative in this set can be chosen by at most $n-2$ individuals at that particular state.[39]

The choice incompatible pair property plays an important role in Theorem 2: This property coupled with consistency are sufficient for ex-post implementation.

Theorem 2. *Let $n \geq 3$. If $F \in \mathcal{F}$ is an SCS for which there exist*

(i) a collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) : i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ consistent with F under incomplete information, and

(ii) a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}} S \subseteq \bar{X}$ which satisfies the choice incompatible pair property at every state $\theta \in \Theta$,

then F is ex-post implementable.

In words, Theorem 2 establishes the following when there are three or more individuals who are not in perfect agreement concerning the socially optimal alternatives: If (i) there exists a collection of sets $\mathbb{S}$ consistent with a non-empty SCS F under incomplete information, and (ii) there exists a set of alternatives $\bar{X}$ which contains every alternative that appears in $\mathbb{S}$ and satisfies the choice incompatible pair property at every state of the world, then F is ex-post implementable.[40]

[38] We note that there is plenty of room for other sufficient conditions since we do not restrict individual choices by requiring universal choice axioms. However, it seems neither easy nor practical to close the gap between the necessary and sufficient conditions.

[39] The choice incompatible pair property is similar to the *economic environment* assumption in the rational domain. Yet, it is *weaker* in our setup since it is now defined on a particular set.

[40] Non-emptiness of $\mathbb{S}$ and $\bar{X}$ follow from $F \in \mathcal{F}$.

Theorem 2 identifies conditions that make sure that all EPE of the mechanism described in Section A.1 falls under Rule 1 at every state of the world. Below, we provide another set of sufficient conditions by employing the same mechanism, but this time allowing for EPE to arise under Rule 2 and Rule 3 as well. To do so, we turn to the counterpart of the no-veto power property in our environment.

Definition 7. *We say that an SCF f satisfies the* **choice no-veto-power property** *on a non-empty set of alternatives $S \in \mathcal{X}$ at state $\theta \in \Theta$ if $x \in C_i^{\theta}(S)$ for all $i \in N \setminus \{j\}$ for some $j \in N$ implies $f(\theta) = x$.*

The choice-no-veto power property on a set, at a particular state, requires that if an alternative is chosen from this set by at least $n-1$ individuals at the particular state, then this alternative must be f-*optimal* at this particular state.

Our second sufficiency result for the case of three or more individuals employs a combination of consistency and the choice no-veto-power property. Below, we present this sufficiency condition followed by the result.[41]

Definition 8. *An SCS $F \in \mathcal{F}$ satisfies the* **consistency-no-veto property** *whenever there exist*

(i) a collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) : i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ such that for all $f \in F$ and for all $i \in N$, $f(\theta'_i, \theta_{-i}) \in C_i^{(\theta'_i, \theta_{-i})}(S_i(f, \theta_{-i}))$ for each $\theta'_i \in \Theta_i$,

(ii) and a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}} S \subseteq \bar{X}$

such that for any collection of product sets of states $\{\bar{\Theta}_f\}_{f \in F}$ with $\bar{\Theta} = \bigcup_{f \in F} \bar{\Theta}_f \subset \Theta$, there exists $f^* \in F$ such that

(iii) f^* satisfies choice no-veto-power property on $\bar{X}$ at every $\theta \in \Theta \setminus \bar{\Theta}$, and

(iv) if for any $f \in F$ and any deception profile α, $f(\alpha(\theta)) \neq f^*(\theta)$ for some $\theta \in \bar{\Theta}_f$, then there exists $i^* \in N$ and $\theta^* \in \bar{\Theta}_f$ such that $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*})))$.

Theorem 3. *Let $n \geq 3$. If an SCS $F \in \mathcal{F}$ satisfies the consistency-no-veto property, then F is ex-post implementable.*

Given a non-empty SCS F, the *consistency-no-veto* property, in words, requires the existence of a collection of sets $\mathbb{S}$ with the property that $S_i(f, \theta_{-i})$ does not depend on θ_i for all $i \in N$ and $f \in F$ and $\theta_{-i} \in \Theta_{-i}$ and a set of alternatives $\bar{X}$ which contains every alternative that appears in $\mathbb{S}$ such that the following hold:

[41] The set $\bar{\Theta} \subseteq \Theta$ is a product set whenever $\bar{\Theta} = \times_{i \in N} \bar{\Theta}_i$ where $\bar{\Theta}_i \subseteq \Theta_i$ with the convention that $\bar{\Theta} = \emptyset$ whenever $\bar{\Theta}_i = \emptyset$ for some $i \in N$.

— Given any $i \in N$ and any $f \in F$ and any $\theta_{-i} \in \Theta_{-i}$, it must be that i's choices when he/she is of type θ_i' at state (θ_i', θ_{-i}) contains $f(\theta_i', \theta_{-i})$ for all $\theta_i' \in \Theta_i$; and

— for any collection of product sets of states $\{\bar{\Theta}_f\}_{f \in F}$ with $\bar{\Theta} = \bigcup_{f \in F} \bar{\Theta}_f \subset \Theta$, there is an SCF f^* in F such that

- if $\theta \in \Theta \setminus \bar{\Theta}$, then f^* obeys the choice no-veto-power property on $\bar{X}$ at θ, and
- if a deception profile α and an SCF $f \in F$ lead to an outcome different than $f^*(\theta)$ for some $\theta \in \bar{\Theta}_f$, then there exists a whistle-blower $i^* \in N$ and an informant state θ^* such that i^* does not choose at θ^* the alternative $f(\alpha(\theta^*))$ (the alternative generated by this deception at θ^*) from $S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*}))$.

Evidently, the consistency-no-veto property is analogous to the monotonicity-no-veto condition of Jackson (1991) and the ex post monotonicity no veto property of Bergemann and Morris (2008). Moreover, our findings are parallel with these papers in the following sense: Jackson (1991) considers a rational domain with expected utility maximizing individuals and establishes that monotonicity-no-veto and incentive compatibility and a condition called closure are sufficient for the Bayesian implementation of SCSs. Meanwhile, Bergemann and Morris (2008) providing sufficiency conditions for ex-post implementation in the rational domain employs ex-post monotonicity no veto condition and ex-post incentive compatibility, both of which are "ex-post analogs of the Bayesian implementation" conditions. In our setting, the closure condition is trivially satisfied as in Bergemann and Morris (2008); by repeating the same arguments presented in the proof of Proposition 2, one can easily show that quasi-ex-post choice incentive compatibility is implied by (i) of the consistency-no-veto property.

A due remark concerns the cases when attention is restricted to the behavioral ex-post implementation of an SCF. Then, the hypothesis of Theorem 3 simplifies to deliver the following analog of Theorem 3 of Bergemann and Morris (2008):

Corollary 1. *Let $n \geq 3$. An SCF $f : \Theta \to X$ is ex-post implementable whenever there exists a collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) : i \in N, \theta_{-i} \in \Theta_{-i}\}$ such that and for all individuals $i \in N$, $f(\theta_i', \theta_{-i}) \in C_i^{(\theta_i', \theta_{-i})}(S_i(f, \theta_{-i}))$ for each $\theta_i' \in \Theta_i$, and there exists a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}} S \subseteq \bar{X}$ such that for any product set of states $\bar{\Theta} \subset \Theta$,*

(i) f satisfies choice no-veto-power property on $\bar{X}$ at every $\theta \in \Theta \setminus \bar{\Theta}$, and

(ii) for any deception profile α with $f(\alpha(\theta)) \neq f(\theta)$ for some $\theta \in \bar{\Theta}$, there exists $i^* \in N$ and $\theta^* \in \bar{\Theta}$ such that $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*})))$.

6 Two Individuals

We now turn to the case of two individuals. First, we slightly improve the necessary conditions for the case of two individuals, and then provide sufficient conditions.

In what follows, Theorem 4 establishes that the notion of *two-individual consistency under incomplete information* is necessary. This concept is defined below and improves the notion of consistency—specified in Definition 3—slightly for the case of two individuals. We refer to it as *two-individual consistency* when the meaning is clear.

Definition 9. *We say that collections of sets* $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in F, \theta_2 \in \Theta_2\} \subset \mathcal{X}$ *and* $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\} \subset \mathcal{X}$ *are* **two-individual consistent with the SCS** $F \in \mathcal{F}$ **under incomplete information** *if*

(i) *for all* $f \in F$, $f(\theta_1', \theta_2) \in C_1^{(\theta_1', \theta_2)}(S_1(f, \theta_2))$ *for each* $\theta_1' \in \Theta_1$,

(ii) *for all* $f \in F$, $f(\theta_1, \theta_2') \in C_2^{(\theta_1, \theta_2')}(S_2(f, \theta_1))$ *for each* $\theta_2' \in \Theta_2$,

(iii) *for all* $f, f' \in F$, $S_1(f, \theta_2) \cap S_2(f', \theta_1) \neq \emptyset$ *for each* $\theta_1 \in \Theta_1$ *and* $\theta_2 \in \Theta_2$,

(iv) *for all* $f \in F$, *if* $f \circ \alpha \notin F$, *then there exists* $\theta^* = (\theta_1^*, \theta_2^*) \in \Theta$ *such that either* $f(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(f, \alpha_2(\theta_2^*)))$ *or* $f(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(f, \alpha_1(\theta_1^*)))$.

We wish to emphasize that (i) and (ii) of two-individual consistency is implied by (i) of consistency while (ii) of consistency implies (iv) of two-individual consistency. That is why the novel condition of two-individual consistency is (iii): For any given pairs of SCFs in the SCS, the two collections of sets must be such that each set associated with individual 1 has a common alternative with each set associated with individual 2.[42]

Theorem 4. *Let* $n = 2$. *If an SCS* $F \in \mathcal{F}$ *is ex-post implementable, then there exist collections of sets* $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in F, \theta_2 \in \Theta_2\}$ *and* $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\}$ *that are two-individual consistent with* F *under incomplete information.*

Now, we display the practical implications of Theorem 4 pertaining to our motivating example. We remind the reader that the individual choices of Alice and Bob and the SCS under consideration are given in Table 1 and Table 3, respectively.

[42]Item (iii) of two-individual consistency, is similar in spirit to part $(i) - (a)$ of Condition β of Dutta and Sen (1991), a paper presenting a necessary and sufficient condition for Nash implementation with two individuals under complete information in the rational domain.

By employing two-individual consistency, below we investigate the collections of sets $\mathbb{S}_A = \{S_A(f, \rho_B), S_A(f, \gamma_B), S_A(f', \rho_B), S_A(f', \gamma_B)\}$ for Alice and $\mathbb{S}_B = \{S_B(f, \rho_A), S_B(f, \gamma_A), S_B(f', \rho_A), S_B(f', \gamma_A)\}$ for Bob.

In particular, using (i) and (ii) of two-individual consistency, we narrow down the candidates for each of these sets as follows. Let us start with Alice:

$\underline{S_A(f, \rho_B)}$: $f(\rho_A, \rho_B) = n$ and $f(\gamma_A, \rho_B) = n$ imply $n \in C_A^{(\rho_A,\rho_B)}(S_A(f, \rho_B))$ and $n \in C_A^{(\gamma_A,\rho_B)}(S_A(f, \rho_B))$. There are four such sets: $\{c, n, s\}, \{c, n\}, \{n, s\}, \{n\}$.

$\underline{S_A(f, \gamma_B)}$: $f(\rho_A, \gamma_B) = n$ and $f(\gamma_A, \gamma_B) = s$ imply $n \in C_A^{(\rho_A,\gamma_B)}(S_A(f, \gamma_B))$ and $s \in C_A^{(\gamma_A,\gamma_B)}(S_A(f, \rho_B))$. There is only one such set: $\{c, n, s\}$.

$\underline{S_A(f', \rho_B)}$: $f'(\rho_A, \rho_B) = s$ and $f'(\gamma_A, \rho_B) = c$ imply $s \in C_A^{(\rho_A,\rho_B)}(S_A(f', \rho_B))$ and $c \in C_A^{(\gamma_A,\rho_B)}(S_A(f', \rho_B))$. There is only one such set as well: $\{c, s\}$.

$\underline{S_A(f', \gamma_B)}$: $f'(\rho_A, \gamma_B) = s$ and $f'(\gamma_A, \gamma_B) = c$ imply $s \in C_A^{(\rho_A,\gamma_B)}(S_A(f', \rho_B))$ and $c \in C_A^{(\gamma_A,\gamma_B)}(S_A(f', \rho_B))$. There is again only one such set: $\{c, s\}$.

Next comes Bob:

$\underline{S_B(f, \rho_A)}$: $f(\rho_A, \rho_B) = n$ and $f(\rho_A, \gamma_B) = n$ imply $n \in C_B^{(\rho_A,\rho_B)}(S_B(f, \rho_A))$ and $n \in C_B^{(\rho_A,\gamma_B)}(S_B(f, \rho_A))$. There are two such sets: $\{c, n\}$ and $\{n\}$.

$\underline{S_B(f, \gamma_A)}$: $f(\gamma_A, \rho_B) = n$ and $f(\gamma_A, \gamma_B) = s$ imply $n \in C_B^{(\gamma_A,\rho_B)}(S_B(f, \gamma_A))$ and $s \in C_B^{(\gamma_A,\gamma_B)}(S_B(f, \gamma_A))$. There is again only one such set: $\{n, s\}$.

$\underline{S_B(f', \rho_A)}$: $f'(\rho_A, \rho_B) = s$ and $f'(\rho_A, \gamma_B) = s$ imply $s \in C_B^{(\rho_A,\rho_B)}(S_B(f', \rho_A))$ and $s \in C_B^{(\rho_A,\gamma_B)}(S_B(f', \rho_A))$. There are three such sets $\{c, s\}$ and $\{n, s\}$ and $\{s\}$.

$\underline{S_B(f', \gamma_A)}$: $f'(\gamma_A, \rho_B) = c$ and $f'(\gamma_A, \gamma_B) = c$ imply $c \in C_B^{(\gamma_A,\rho_B)}(S_B(f', \gamma_A))$ and $c \in C_B^{(\gamma_A,\gamma_B)}(S_B(f', \gamma_A))$. There are two such sets $\{c, n\}$ and $\{c\}$.

Therefore, we conclude that $S_A(f, \gamma_B) = \{c, n, s\}$, $S_A(f', \rho_B) = \{c, s\}$, $S_A(f', \gamma_B) = \{c, s\}$, and $S_B(f, \gamma_A) = \{n, s\}$.

Furthermore, condition (iii) of two-individual consistency implies that $S_A(f, \theta_B) \cap S_B(f', \theta_A) \neq \emptyset$ and $S_A(f', \theta_B) \cap S_B(f, \theta_A) \neq \emptyset$ for each $\theta_A \in \{\rho_A, \gamma_A\}$ and $\theta_B \in \{\rho_B, \gamma_B\}$. Thus, $S_A(f', \rho_B) \cap S_B(f, \rho_A) \neq \emptyset$, and this implies $S_B(f, \rho_A) = \{c, n\}$.

These uniquely identify 5 out of 8 of the two-individual consistent collections of sets of Alice and Bob. In particular, we must have $\mathbb{S}_A = \{S_A(f, \rho_B), \{c, n, s\}, \{c, s\}\}$ and $\mathbb{S}_B = \{\{c, n\}, \{n, s\}, S_B(f', \rho_A), S_B(f', \gamma_A)\}$. It is possible to narrow down $\mathbb{S}_A$ and $\mathbb{S}_B$ further by employing condition (iv) of two-individual consistency. Yet, this would be tedious since there are many deceptions to consider.[43] However, the mechanism given in Table

[43]There are 15 possible deceptions where either Alice or Bob misrepresents their types. All 15 of them lead to $f \circ \alpha \neq f$ and 12 of them lead to $f' \circ \alpha \neq f'$. It is useful to note that deceptions are non-cooperative and hence measurable only with respect to private information. Thus, we cannot have $\alpha(\rho_A, \rho_B) = (\gamma_A, \rho_B)$ and $\alpha(\rho_A, \gamma_B) = (\rho_A, \rho_B)$ where Alice lies about her type when her type is ρ_A

4 implementing F in our motivating example implies the following collections of sets are two-individual consistent with F: $\mathbb{S}_A$ given by $S_A(f,\rho_B)=\{c,n\}$, $S_A(f,\gamma_B)=\{c,n,s\}$, $S_A(f',\rho_B)=\{c,s\}$, $S_A(f',\gamma_B)=\{c,s\}$ and $\mathbb{S}_B$ by $S_B(f,\rho_A)=\{c,n\}$, $S_B(f,\gamma_A)=\{n,s\}$, $S_B(f',\rho_A)=\{c,s\}$, $S_B(f',\gamma_A)=\{c,n\}$. In Section 7.1, we show that this is not the unique pair of two-individual consistent collections of sets with F.[44]

We proceed with sufficiency conditions for the case of two individuals. Below, we introduce the additional properties needed to transform the necessary conditions into sufficient ones concerning ex-post implementation with two individuals.

The first of these concerns choice incompatibility for the case of two individuals.

Definition 10. *Given an SCS $F\in\mathcal{F}$, we say that F involves* **choice incompatibility** *among a non-empty set of alternatives $S\in\mathcal{X}$ and non-empty collections of sets $\mathbb{S}_1:=\{S_1(f,\theta_2)|f\in F,\theta_2\in\Theta_2\}\subset\mathcal{X}$ and $\mathbb{S}_2:=\{S_2(f,\theta_1)|f\in F,\theta_1\in\Theta_1\}\subset\mathcal{X}$ at state θ if*

(i) $x\in C_i^\theta(S)$ implies $x\notin C_j^\theta(S)$, $i\neq j$; and

(ii) for any $T\in\mathbb{S}_i$, $x\in C_i^\theta(T)$ implies $x\notin C_j^\theta(S)$, $i=1,2$ and $i\neq j$; and

(iii) for any deception profile α and any $f,f'\in F$ with $f\neq f'$, $x\in C_i^\theta(S_i(f,\alpha_j(\theta_j)))$ implies $x\notin C_j^\theta(S_j(f',\alpha_i(\theta_i)))$, $i=1,2$ and $i\neq j$.

Choice incompatibility conditions require that there is sufficiently strong disagreement between the two individuals at a given state. Indeed, the intuition behind choice incompatibility is as follows: A non-empty SCS F involves choice incompatibility among a non-empty set of alternatives S and non-empty collections of sets $\mathbb{S}_1$ and $\mathbb{S}_2$ at state θ means that the individual choices at θ are not aligned when (i) both individuals make choices separately from S; and (ii) one individual, i, is making a choice from a set in $\mathbb{S}_i$ and the other individual, j, is making a choice from S where $i,j=1,2$ with $i\neq j$; and (iii) individual i makes a choice from a set in $\mathbb{S}_i$ that is associated with a particular SCF f and the other individual, j, makes a choice from a set in $\mathbb{S}_j$ which is associated with a different SCF $f'\neq f$ while $f,f'\in F$ and $i,j=1,2$ with $i\neq j$.[45]

---

and Bob's type is ρ_B but not when her type is ρ_A and Bob's type is γ_B.

[44]We identify another mechanism ex-post implementing F on page 34 that implies another pair of two-individual consistent collections of sets with the only difference being $S_B(f',\rho_A)=\{c\}$.

[45]Choice incompatibility has some relations with part (iv) of Condition $\mu 2$ of Moore and Repullo (1990) and part $(i)-(b)$ of Condition β of Dutta and Sen (1991); both among the necessary and sufficient conditions for Nash implementation in the rational domain under complete information. These require the existence of a common alternative x in the choice sets of the two individuals where the choice set of the first individual is associated with a preference profile and alternative pair (R,a), while that of the second with (R',b) such that x being maximal with respect to some preference profile R'' for both

At this stage, we wish to emphasize that we handle cases in which individuals' choices are aligned later in the section, when we start discussing choice unanimity.

Theorem 5, below, shows that two-individual consistency coupled with choice incompatibility are sufficient for ex-post implementation.

Theorem 5. *Let $n = 2$. If $F \in \mathcal{F}$ is an SCS for which there exist*

(*i*) *collections of sets $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in F, \theta_2 \in \Theta_2\}$ and $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\}$ which are two-individual consistent with F under incomplete information, and*

(*ii*) *a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}_1 \cup \mathbb{S}_2} S \subseteq \bar{X}$ such that F involves choice incompatibility among $\bar{X}$ and $\mathbb{S}_1$ and $\mathbb{S}_2$ at every $\theta \in \Theta$,*

then F is ex-post implementable.

In words, when there are two individuals, Theorem 5 implies that if (i) there exist individual specific collections of sets $\mathbb{S}_1$ and $\mathbb{S}_2$ that are two-individual consistent with F under incomplete information, and (ii) there exists a set of alternatives $\bar{X}$ which contains every alternative that appears in $\mathbb{S}_1$ and $\mathbb{S}_2$ and the afore discussed choice incompatibility among $\bar{X}$ and $\mathbb{S}_1$ and $\mathbb{S}_2$ hold at every state of the world, then F is ex-post implementable. That is, Theorem 5 demands sufficiently "strong" disagreement between the two individuals for sufficiency of ex-post implementation.[46]

A common thread between Theorems 2 and 5 is that their hypotheses enable us to sustain all EPE of the associated mechanisms under Rule 1 at every state of the world. As was done in the case of three or more individuals, we provide another set of sufficient conditions by employing the same mechanism presented in Section A.2, but allowing EPE to arise under other rules as well. Because no-veto power is "hopelessly strong" with two individuals (Moore & Repullo, 1990), we turn to the concept of choice unanimity.

Definition 11. *We say that an SCS $F \in \mathcal{F}$ respects* **choice unanimity** *on a non-empty set of alternatives $S \in \mathcal{X}$ and non-empty collections of sets $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in$*

of the individuals from these choice sets implies x being a member of the social choice correspondence at R''. On the other hand, choice incompatibility (akin to the economic environment assumption of the rational domain with incomplete information) does not allow any alternative to be ranked first by both individuals even when the two individuals' choices are represented by complete and transitive preferences. Thus, choice incompatibility brings about a requirement that is similar in spirit to part (iv) of Condition $\mu 2$ and part $(i)-(b)$ of Condition β.

[46] Ohashi (2012) presents sufficient conditions for ex-post implementation with two individuals in the rational domain. Unlike ours, his sufficient conditions require the existence of a *bad outcome*: an alternative that is strictly worse than any other in the union of the ranges of the SCFs in the SCS.

$F, \theta_2 \in \Theta_2\} \subset \mathcal{X}$ and $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\} \subset \mathcal{X}$ at state θ if there exists $f^* \in F$ such that

(i) $x \in C_1^\theta(S) \cap C_2^\theta(S)$ implies $f^*(\theta) = x$; and

(ii) for any $T \in \mathbb{S}_i$, $x \in C_i^\theta(T) \cap C_j^\theta(S)$ implies $f^*(\theta) = x$, for $i, j = 1, 2$ with $i \neq j$; and

(iii) for any deception profile α and $f, f' \in F$ with $f \neq f'$, $x \in C_i^\theta(S_i(f, \alpha_j(\theta_j))) \cap C_j^\theta(S_j(f', \alpha_i(\theta_i)))$ implies $f^*(\theta) = x$.

The general intuition behind choice unanimity of an SCS F with two individuals is that if the choices of the individuals (from some particular sets) agree at a given state, then the SCS must respect this: the chosen alternatives must be achievable with one of SCFs in the SCS at that state. It is a mild condition as it allows the SCS to accommodate SCFs that are not restricted to deliver commonly agreed upon alternatives.

Our second sufficiency result for the case of two individuals makes use of the following combination of consistency and choice unanimity:

Definition 12. *Let $n = 2$. An SCS $F \in \mathcal{F}$ satisfies the* **consistency-unanimity property** *whenever there exist collections of sets $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in F, \theta_2 \in \Theta_2\} \subset \mathcal{X}$ and $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\} \subset \mathcal{X}$ such that*

(i) for all $f \in F$, $f(\theta_1', \theta_2) \in C_1^{(\theta_1', \theta_2)}(S_1(f, \theta_2))$ for each $\theta_1' \in \Theta_1$, and

(ii) for all $f \in F$, $f(\theta_1, \theta_2') \in C_2^{(\theta_1, \theta_2')}(S_2(f, \theta_1))$ for each $\theta_2' \in \Theta_2$, and

(iii) for all $f, f' \in F$, $S_1(f, \theta_2) \cap S_2(f', \theta_1) \neq \emptyset$ for each $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$,

and there is a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}_1 \cup \mathbb{S}_2} S \subseteq \bar{X}$ such that for any collection of product sets $\{\bar{\Theta}_f\}_{f \in F}$ with $\bar{\Theta} = \bigcup_{f \in F} \bar{\Theta}_f \subset \Theta$,

(iv) F respects choice unanimity on $\bar{X}$ and $\mathbb{S}_1$ and $\mathbb{S}_2$ at every $\theta \in \Theta \setminus \bar{\Theta}$, and

(v) for all $f \in F$ and deception profile α, if $f(\alpha(\theta)) \neq f^*(\theta)$ for some $\theta \in \bar{\Theta}_f$ where f^* is the SCF that satisfies (i)–(iii) of choice unanimity, then there exists $\theta^* \in \bar{\Theta}_f$ such that either $f(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(f, \alpha_2(\theta_2^*)))$ or $f(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(f, \alpha_1(\theta_1^*)))$.

Consistency-unanimity, in words, requires the following:[47] Given a non-empty SCS F, there exist collections of sets $\mathbb{S}_i$ with the property that $S_i(f, \theta_j)$ does not depend on

[47] Consistency-unanimity entails choice unanimity conditions that are similar in spirit to (iv) of Condition $\mu 2$ and $(i) - (b)$ of Condition β. For more on them, please see Footnote 45.

θ_i for all $i, j = 1, 2$ with $i \neq j$ and $f \in F$ and $\theta_j \in \Theta_j$ and a set of alternatives $\bar{X}$ which contains every alternative that appears in $\mathbb{S}_1 \cup \mathbb{S}_2$ such that the following hold:

— Given any $i \in \{1, 2\}$ and any $f \in F$ and any $\theta_j \in \Theta_j$, it must be that i's choice when he/she is of type θ'_i at state (θ'_i, θ_j) contains $f(\theta'_i, \theta_j)$ for all $\theta'_i \in \Theta_i$, with $j = 1, 2$ and $i \neq j$; and

— any set from $\mathbb{S}_1$ must have a common element with any set from $\mathbb{S}_2$; and

— for any collection of product sets of states $\{\bar{\Theta}_f\}_{f\ \in F}$ with $\bar{\Theta} = \bigcup_{f \in F} \bar{\Theta}_f \subset \Theta$, there is an SCF f^* in F such that

- F respects choice unanimity on $\bar{X}$ and $\mathbb{S}_1$ and $\mathbb{S}_2$ whenever $\theta \in \Theta \setminus \bar{\Theta}$; and
- for any deception profile α and SCF $f \in F$ that lead to an outcome different than $f^*(\theta)$ for some $\theta \in \bar{\Theta}_f$ where f^* is the SCF that satisfies the choice unanimity conditions (i)–(iii), there exists a whistle-blower $i^* \in \{1, 2\}$ and an informant state $\theta^* \in \bar{\Theta}_f$ such that i^* does not choose at θ^* the alternative $f(\alpha(\theta^*))$ from $S_{i^*}(f, \alpha_j(\theta^*_j))$ where $j \in \{1, 2\}$ and $j \neq i^*$.

Below we establish that the consistency-unanimity property is sufficient for ex-post implementation with two individuals. Indeed, it is a novel two-individual condition which draws its motivation from three or more individuals sufficiency conditions, consistency-no-veto of the current paper, monotonicity-no-veto of Jackson (1991), and ex post monotonicity no veto of Bergemann and Morris (2008).

Theorem 6. *Let $n = 2$. If an SCS $F \in \mathcal{F}$ satisfies the consistency-unanimity property, then F is ex-post implementable.*

We wish to emphasize that when $F = \{f\}$, i.e., the case in which the planner is seeking to implement an SCF in EPE, then, (iii) of choice unanimity holds vacuously.[48] This simplifies the hypothesis of Theorem 6 and delivers the following:

Corollary 2. *Let $n = 2$. An SCF $f : \Theta \to X$ is ex-post implementable whenever there are collections of sets $\mathbb{S}_i := \{S_i(f, \theta_j) | \theta_j \in \Theta_j\} \subset \mathcal{X}$ with $i, j = 1, 2$ and $i \neq j$ such that*

(i) $f(\theta'_1, \theta_2) \in C_1^{(\theta'_1, \theta_2)}(S_1(f, \theta_2))$ for each $\theta'_1 \in \Theta_1$, and $f(\theta_1, \theta'_2) \in C_2^{(\theta_1, \theta'_2)}(S_2(f, \theta_1))$ for each $\theta'_2 \in \Theta_2$, and $S_1(f, \theta_2) \cap S_2(f, \theta_1) \neq \emptyset$ for each $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$,

[48] We note that when $F = \{f\}$, Rule 3 of the mechanism presented in Section A.2 becomes redundant.

and there is a set of alternatives $\bar{X} \subseteq X$ *with* $\bigcup_{S \in \mathbb{S}_1 \cup \mathbb{S}_2} S \subseteq \bar{X}$ *such that for any product set* $\bar{\Theta} \subseteq \Theta$,

(ii) $x \in C_1^{\theta}(\bar{X}) \cap C_2^{\theta}(\bar{X})$ *implies* $f(\theta) = x$, $x \in C_1^{\theta}(T) \cap C_2^{\theta}(\bar{X})$ *with* $T \in \mathbb{S}_1$ *implies* $f(\theta) = x$, *and* $x \in C_1^{\theta}(\bar{X}) \cap C_2^{\theta}(T')$ *with* $T' \in \mathbb{S}_2$ *implies* $f(\theta) = x$, *for each* $\theta \in \Theta \setminus \bar{\Theta}$, *and*

(iii) *for any deception profile* α, *if* $f(\alpha(\theta)) \neq f(\theta)$ *for some* $\theta \in \bar{\Theta}$, *then there exists* $\theta^* \in \bar{\Theta}$ *such that either* $f(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(f, \alpha_2(\theta_2^*)))$ *or* $f(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(f, \alpha_1(\theta_1^*)))$.

In what follows, we briefly elaborate on the relation of our motivating example to the sufficiency results we provide for the case of two individuals. The individual choices of Alice and Bob specified in Table 1 and collections of sets $\mathbb{S}_A$ and $\mathbb{S}_B$ satisfying neither consistency-no-veto nor consistency-unanimity (see Definitions 8 and 12, respectively) establish that our sufficiency conditions for the case of two individuals are not necessary in general, as expected. To see why, consider the individual choices of Alice from the set $S_A(f, \rho_B) = \{c, n\}$, and of Bob from the set $\bar{X} = \{c, n, s\}$ at state (γ_A, γ_B).[49] Alice chooses n from $S_A(f, \rho_B)$, whereas Bob chooses both n and s from the set $\bar{X}$ at (γ_A, γ_B). As n is chosen by both, choice incompatibility fails at (γ_A, γ_B). Thus, we should turn to choice unanimity to be able to employ the mechanism of Section A.2 to deliver sufficiency. However, there is no SCF $f^* \in F$ in the SCS F such that $f^*(\gamma_A, \gamma_B) = n$. Therefore, choice unanimity as well as choice incompatibility on $(\bar{X}, \mathbb{S}_A)$ fail and hence we can employ neither Theorem 5 nor Theorem 6.

To demonstrate the relevance and applicability of our sufficiency results, below, we show how one of them, Corollary 2, can be employed on an example that is inspired from Masatlioglu and Ok (2014).[50]

Suppose that the states of the world regarding the individual choices of Alice and Bob are given by $\Theta = \{(\Diamond, \Diamond), (\Diamond, c), (c, \Diamond), (c, c)\}$. That is, $\Theta_A = \Theta_B = \{\Diamond, c\}$, where type $\Diamond$ stands for *not having a status-quo* and type c stands for *status-quo being coal.* We consider the individual choices of Alice and Bob from (the subsets) of $X = \{c, n, s\}$ as specified in Table 9.

[49] By following the same line of arguments made in the discussions immediately following Theorem 4, one can show that (*i*)–(*iii*) of both consistency-no-veto and consistency-unanimity determine five members of such collections of sets of Alice and Bob: $\mathbb{S}_A$ must be such that $S_A(f, \gamma_B) = \{c, n, s\}$, $S_A(f', \rho_B) = \{c, s\}$, and $S_A(f', \gamma_B) = \{c, s\}$; and $\mathbb{S}_B$ must be such that $S_B(f, \rho_A) = \{c, n\}$, and $S_B(f, \gamma_A) = \{n, s\}$. Consequently, $\bar{X} = \{c, n, s\}$ because $\bar{X} \subseteq X = \{c, n, s\}$.

[50] Masatlioglu and Ok (2014) presents a "model of individual decision making when the endowment of an agent provides a reference point that may influence her choices".

S	$C_A^{(\Diamond,\Diamond)}$	$C_B^{(\Diamond,\Diamond)}$	$C_A^{(\Diamond,c)}$	$C_B^{(\Diamond,c)}$	$C_A^{(c,\Diamond)}$	$C_B^{(c,\Diamond)}$	$C_A^{(c,c)}$	$C_B^{(c,c)}$
$\{c,n,s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{n\}$	$\{n\}$	$\{s\}$	$\{n\}$	$\{n\}$
$\{c,n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$	$\{n\}$
$\{c,s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{c\}$	$\{c\}$	$\{s\}$	$\{c\}$	$\{c\}$
$\{n,s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{s\}$	$\{n\}$	$\{n\}$

Table 9: A two-individual example satisfying consistency-unanimity.

A social planner wants to ex-post implement the SCF f, a particular selection from the BR-optimal outcomes, described in Table 10: The social planner breaks the tie in favor of s whenever n and s are both BR-optimal.

State	$(\Diamond,\Diamond)$	$(\Diamond,c)$	$(c,\Diamond)$	(c,c)
BR-optimal	$\{s\}$	$\{n,s\}$	$\{n,s\}$	$\{n\}$
f	$\{s\}$	$\{s\}$	$\{s\}$	$\{n\}$

Table 10: BR-optimal alternatives and SCF f.

Indeed, the planer can employ the mechanism described in Section A.2 to implement the SCF f in ex-post equilibrium: In Section C of the Appendix, we show that the collections $\mathbb{S}_A := \{S_A(f,\Diamond), S_A(f,c)\}$ and $\mathbb{S}_B := \{S_B(f,\Diamond), S_B(f,c)\}$ with $S_A(f,\Diamond) = \{n,s\}$, $S_A(f,c) = \{c,n,s\}$ and $S_B(f,\Diamond) = \{n,s\}$, $S_B(f,c) = \{c,n,s\}$ satisfy conditions (i), (ii), and (iii) of Corollary 2.[51]

7 Simple Mechanisms

There has been a recent interest in simple mechanisms in the mechanism design literature.[52] Indeed, dealing with individuals having limited cognitive abilities increases the relevance and importance of the simplicity of mechanisms. In this section, we elaborate on the simplicity of mechanisms that can be used for ex-post implementation when individuals' choices do not necessarily satisfy the standard axioms of rationality.

In our view, given that the individuals under consideration are susceptible to behavioral biases, a designer should choose a consistent collection of sets that gives rise to the *simplest* mechanism, one in which the number of messages is as low as possible. The following analysis provides insights regarding simplicity in this regard.

[51]We note that Rule 3 of the mechanism described in Section A.2 becomes redundant when ex-post implementation of an SCF is desired.

[52]See for example, Li (2017) and Borgers and Li (2018).

7.1 Motivating example revisited

We start with the mechanism we employ in our motivating example presented in Section 3.1 in Table 4, which ex-post implements the SCS described in Table 3 for the individual choices of Alice and Bob as specified in Table 1. In particular, we discuss why there does not exist any simpler mechanism that implements that SCS in EPE.

To see why, consider the discussion on page 25 and recall that two-individual consistency (see Definition 9) pins down 5 of 8 members of such collections of sets of Alice and Bob: $\mathbb{S}_A$ must be such that $S_A(f, \gamma_B) = \{c, n, s\}$, $S_A(f', \rho_B) = \{c, s\}$, and $S_A(f', \gamma_B) = \{c, s\}$; and $\mathbb{S}_B$ must be such that $S_B(f, \rho_A) = \{c, n\}$, and $S_B(f, \gamma_A) = \{n, s\}$.

Fortunately, it is possible to see that our mechanism, presented in Table 4, is one of the simplest mechanisms that ex-post implements F without any need to further narrow down $\mathbb{S}_A$ and $\mathbb{S}_B$ by employing condition (iv) of two-individual consistency.

As $S_A(f, \gamma_B)$ must be $\{c, n, s\}$, Alice must have at least three messages to be able to generate this opportunity set in any mechanism that ex-post implements F. Furthermore, Bob must have at least two messages: one for Alice to be able to generate $\{c, n, s\}$ and another for Alice to be able to generate $\{c, s\}$ since we must have $S_A(f', \rho_B) = S_A(f', \gamma_B) = \{c, s\}$. Therefore, the best we can hope for is three messages for Alice and two messages for Bob.

Below, we explain why we need at least one more message. Suppose that there exists a mechanism that ex-post implements the SCS F where Alice has three messages and Bob has two messages. This means we must have $S_A(f, \rho_B) = \{c, n, s\}$ and both $S_B(f', \rho_A)$ and $S_B(f', \gamma_A)$ must be either $\{c, n\}$ or $\{n, s\}$. Therefore, the collections $\mathbb{S}_A = \{\{c, n, s\}, \{c, s\}\}$ and $\mathbb{S}_B = \{\{c, n\}, \{n, s\}, S_B(f', \rho_A), S_B(f', \gamma_A)\}$ hint to us that the mechanism should look like the game form given in Table 11. In this mechanism,

		Bob	
		$\{c, n, s\}$	$\{c, s\}$
Alice	$\{c, n\}$	x	c
	$\{n, s\}$	y	s
	{t, z}	z	t

Table 11: A 3×2 mechanism proposal for Alice and Bob.

the messages are labeled with the opportunity sets that the other individual should be able to generate. This is because any message in a mechanism can be thought of as an opportunity set generated for the other individual. For example, if Bob sends the message on the left, then Alice should be able to generate the set $\{c, n, s\}$. Thus,

$\{\mathrm{x}, \mathrm{y}, \mathrm{z}\} = \{c, n, s\}$ and, hence, $\mathrm{x} \neq \mathrm{y} \neq \mathrm{z}$. If Bob sends the message on the right, then Alice should be able to obtain $\{c, s\}$, the other set in $\mathbb{S}_A$. On the other hand, if Alice sends the message on top, then Bob should be able to generate $\{c, n\}$, and if Alice sends the message in the middle, Bob should be able to generate $\{n, s\}$. Furthermore, each outcome specified in the mechanism must be in both of the sustained opportunity sets of each individual. In particular, if Alice sends $\{c, n\}$ (sustaining the opportunity set $\{c, n\} \in \mathbb{S}_B$ for Bob) and Bob sends $\{c, s\}$ (sustaining the opportunity set $\{c, s\} \in \mathbb{S}_A$ for Alice), the outcome must be c as $\{c, n\} \cap \{c, s\} = \{c\}$. So, for Bob to be able to generate $\{c, n\}$, the outcome must be n whenever Alice sends $\{c, n\}$ and Bob $\{c, n, s\}$. Similarly, the outcome must equal s if Alice sends $\{n, s\}$ and Bob $\{c, s\}$ and hence the outcome must equal n whenever Alice sends $\{n, s\}$ and Bob $\{c, n, s\}$. Thus, we must have $\mathrm{x} = n = \mathrm{y}$, a contradiction to $\mathrm{x} \neq \mathrm{y}$. Thus, the simplest mechanism cannot have three messages for Alice and two for Bob. It must have at least one more message. This makes the mechanism given in Table 4 one of the simplest mechanisms that ex-post implements the SCS F described in Table 3 . We note this observation as a corollary:

Corollary 3. *Given the individual choices of Alice and Bob in Table 1, any mechanism that ex-post implements the SCS F described in Table 3 must have at least three messages for Alice and the total number of messages for both players must be at least six. In this regard, there does not exist any simpler mechanism than the one given in Table 4.*

We note that the mechanism given in Table 4 is not the unique simplest mechanism that works for our motivating example: the mechanism given in Table 12 also ex-post implements the SCS F for the individual choices as specified in Table 1.[53]

		Bob		
		L	M	R
Alice	U	n	c	n
	M	c	c	c
	D	n	s	s

Table 12: Another simplest mechanism for Alice and Bob.

7.2 Lower bounds on the number of messages

In the proof of Theorem 1, the collection of sets $\mathbb{S} = \{S_i(f, \theta_{-i}) | f \in F, i \in N, \theta_{-i} \in \Theta_{-i}\}$ consistent with the SCS F is constructed from the mechanism that ex-post implements F. When there are multiple such mechanisms, there could be possibly different

[53]The only difference from the mechanism defined in Table 4 is $S_B(f', \rho_A) = \{c\}$, instead of $\{c, s\}$.

collections of sets consistent with the same SCS. How many sets there are in a collection, and how small these sets are, turn out to be important when designing simple mechanisms.

The observations made in Section 7.1 lead us to lower bounds on the number of messages needed for behavioral implementation under incomplete information. In fact, we observe that collections $\{\mathbb{S}_i\}_{i\in N}$ that satisfy two-individual consistency formalized in Definition 9 for the case of two individuals (and consistency specified in Definition 3 for the case of three or more individuals) provide the key for obtaining these lower bounds.

Let $\{\mathbb{S}_\gamma\}_{\gamma\in\Gamma}$ be the set of all the collections of sets that satisfy consistency (or two-individual consistency for the case of two individuals) represented by $\mathbb{S}_\gamma = \{\mathbb{S}_i^\gamma\}_{i\in N}$ for each $\gamma \in \Gamma$ with $\mathbb{S}_i^\gamma = \{S_i^\gamma(f, \theta_{-i}) | f \in F, \theta_{-i} \in \Theta_{-i}\}$. Clearly, the goal of the planner is to pick up one of these collections and design a mechanism that ex-post implements F.

In our final result, below, we provide lower bounds for the simplicity of mechanisms that are employed for ex-post implementation:

Theorem 7. *In any mechanism that ex-post implements the SCS $F \in \mathcal{F}$,*

(i) *the minimum number of messages required for individual i is $\min_{\gamma\in\Gamma} \max_{S\in\mathbb{S}_i^\gamma} \#S$,*

(ii) *the minimum number of message profiles required for the individuals other than i is $\min_{\gamma\in\Gamma} \#\mathbb{S}_i^\gamma$, and*

(iii) *the minimum number of total messages required for all individuals is* $\max\left\{\min_{\gamma\in\Gamma} \max_{i\in N}[\#\mathbb{S}_i^\gamma + \max_{S\in\mathbb{S}_i^\gamma} \#S], \min_{\gamma\in\Gamma} \sum_{i\in N} \max_{S\in\mathbb{S}_i^\gamma} \#S\right\}$.

The intuition behind Theorem 7 is simple: If the collection $\mathbb{S}^\gamma$ happens to be the collection of opportunity sets generated by the mechanism that ex-post implements F, then individual i is able to generate any set in $\mathbb{S}_i^\gamma$. Therefore, individual i must have at least as many messages as the cardinality of the maximal set in $\mathbb{S}_i^\gamma$, which implies (i).

At the same time, for each different set in the collection $\mathbb{S}_i^\gamma$, there must exist a particular message profile of the individuals other than i that should allow individual i to generate this particular set, which implies (ii).

Therefore, if the collection $\mathbb{S}^\gamma$ happens to be the collection of opportunity sets generated by the mechanism that ex-post implements F, then the total number of messages in this mechanism must be at least as much as $\max_{i\in N}[\#\mathbb{S}_i^\gamma + \max_{S\in\mathbb{S}_i^\gamma} \#S]$. On the other hand, by (i), the total number of messages required in this mechanism for all the individuals must be also more than $\sum_{i\in N} \max_{S\in\mathbb{S}_i^\gamma} \#S$ for the particular collection $\mathbb{S}_i^\gamma$.

Combining together, the total number of messages must exceed both $\min_{\gamma\in\Gamma} \max_{i\in N}$ $[\#\mathbb{S}_i^\gamma + \max_{S\in\mathbb{S}_i^\gamma} \#S]$ and $\min_{\gamma\in\Gamma} \sum_{i\in N} \max_{S\in\mathbb{S}_i^\gamma} \#S$, which implies ($iii$).

8 Concluding Remarks

We investigate the problem of implementation under incomplete information when individuals' choices need not satisfy the standard axioms of rationality.

The focus is on full implementation in ex-post equilibrium because (i) the revelation principle fails for partial implementation and, hence, one cannot restrict attention to direct revelation mechanisms without a loss of generality; and (ii) the concept of ex-post equilibrium is belief-free, does not require any expectation considerations or any belief updating, and is robust to informational assumptions regarding the environment, which makes it well suited when individuals' choices fail the standard axioms of rationality.

We provide necessary as well as sufficient conditions for the case of three or more individuals and for the case of two individuals separately. Even though, the mechanisms offered in the proofs of our sufficiency theorems can be criticized for not being sufficiently simple, our necessary conditions provide us with hints regarding the limits of simplicity for behavioral implementation under incomplete information.

An interesting direction for future research would be to analyze whether practical and simple mechanisms are available for specific types of behavioral biases. We hope that our results pave the way for contributions in this direction.

A Mechanisms

A.1 Mechanism for the case with three or more individuals

The mechanism we construct for the case with three or more individuals has a standard structure and makes use of the following observations: (i) the outcome should be $f(\theta)$ when there is unanimous agreement between the individuals over $f \in F$ and the true state is θ; (ii) under such a unanimous agreement each individual j should be able to generate unilaterally the set $S_j(f, \theta_{-j})$, i.e., when all other individuals (all $i \neq j$) have unanimously decided on the particular SCF $f \in F$ and sending messages as if their types are $\theta_{-j} \in \Theta_{-j}$, j should be able to generate $S_j(f, \theta_{-j})$; (iii) whenever there is an attempt to deceive the designer so that an outcome not compatible with the SCS is to be implemented, a whistle-blower should be able to alert the designer; (iv) undesirable EPE should be eliminated according to some procedure, e.g., by a modulo game or an integer game.[54]

Consider an SCS $F \in \mathcal{F}$ for which the collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) : i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ and $\bar{X}$ are as specified in Theorem 2 or Theorem 3. For any $i \in N$, $f \in F$, $\theta_{-i} \in \Theta_{-i}$, let $\bar{x}(i, f, \theta_{-i})$ be an arbitrary alternative in $S_i(f, \theta_{-i})$.[55]

The mechanism $\mu = (M, g)$ is defined as follows: The message space of each individual $i \in N$ is $M_i = F \times \Theta_i \times \bar{X} \times N$, while a generic message is denoted by $m_i = (f, \theta_i, x_i, k_i)$, and the outcome function $g : M \to X$ is as specified in Table 13.

Rule 1 :	$g(m) = f(\theta)$	if $m_i = (f, \theta_i, \cdot, \cdot)$ for all $i \in N$,
Rule 2 :	$g(m) = \begin{cases} x_j & \text{if } x_j \in S_j(f, \theta_{-j}), \\ \bar{x}(j, f, \theta_{-j}) & \text{otherwise.} \end{cases}$	if $m_i = (f, \theta_i, \cdot, \cdot)$ for all $i \in N \setminus \{j\}$ and $m_j = (\tilde{f}, \tilde{\theta}_j, x_j, \cdot)$ with $\tilde{f} \neq f$,
Rule 3 :	$g(m) = x_j$ where $j = \sum_i k_i$ (mod n)	otherwise.

Table 13: The outcome function of the mechanism with three or more individuals.

In words, each individual is required to send a message that specifies an SCF $f \in F$, a type for himself $\theta_i \in \Theta_i$, an alternative x_i in $\bar{X}$, and a number $k_i \in N = \{1, 2, \ldots, n\}$.[56]

[54] We note that the mechanism we construct has a standard structure in the sense that it is similar to those that have been widely used for sufficiency proofs in the implementation literature. See for example, Repullo (1987), Saijo (1988), Moore and Repullo (1990), Jackson (1991), Danilov (1992), Maskin (1999), Bergemann and Morris (2008), de Clippel (2014), Koray and Yildiz (2018), among many others. Similar mechanisms are sometimes referred to as augmented mechanisms, integer mechanisms or canonical mechanisms. Indeed, we are puzzled why there is not a consensus over a specific name in the literature.

[55] Such an $\bar{x}(i, f, \theta_{-i})$ exists simply because $S_i(f, \theta_{-i})$ must be non-empty due to (i) of consistency.

[56] We would like to emphasize that since $\bar{X}$ contains $S_i(f, \theta_{-i})$ for each $i \in N$, $f \in F$, $\theta_{-i} \in \Theta{-i}$,

Rule 1 indicates that if there is unanimity among the individuals' messages regarding the SCF to be implemented, then the outcome is determined according to this SCF and the reported type profile in the messages. Rule 2 indicates that if there is agreement between all the individuals but one regarding the SCF $f \in F$ in their messages, then the outcome is determined according to the alternative proposed by the odd-man-out, j, only if this alternative is in $S_j(f, \theta_{-j})$, otherwise the outcome is $\bar{x}(j, f, \theta_{-j})$ which is in $S_j(f, \theta_{-j})$ as well. That is, as desired, when all the other individuals (all $i \neq j$) have unanimously decided on the particular SCF $f \in F$ and sending messages as if their types are $\theta_{-j} \in \Theta_{-j}$, the odd-man-out j is able to generate unilaterally $S_j(f, \theta_{-j})$—and nothing else since $\bar{x}(i, f, \theta_{-i}) \in S_j(f, \theta_{-j})$ as well. Finally, Rule 3 applies when both Rule 1 and Rule 2 fail, then the outcome is determined only according to the reported numbers (k_i's) and the outcome x_j is implemented where j is the individual $\sum_i k_i$ modulo n. Rule 3 makes sure that there are no undesirable EPE of the mechanism.[57]

A.2 Mechanism for the case with two individuals

The mechanism we design for the case of two individuals relies on the following observations: (i) the outcome should be $f(\theta)$ when there is agreement between the two individuals over $f \in F$ and the true state is θ; (ii) each individual i should be able to generate unilaterally the set $S_i(f, \theta_j)$ when the other individual $j \neq i$ intends a particular SCF $f \in F$ and sends a message as if his/her type is $\theta_j \in \Theta_j$; (iii) whenever there is an attempt to deceive the designer so that an undesired outcome is to be implemented, a whistle-blower should be able to alert the designer; (iv) undesirable EPE should be eliminated according to a procedure, e.g., a modulo game or an integer game.

Consider any $F \in \mathcal{F}$ for which $\mathbb{S}_1 := \{S_1(f, \theta_2) | f \in F, \theta_2 \in \Theta_2\}$, $\mathbb{S}_2 := \{S_2(f, \theta_1) | f \in F, \theta_1 \in \Theta_1\}$, and $\bar{X}$ are as specified in Theorem 5 or Theorem 6.

For any $i, j \in \{1, 2\}$ with $i \neq j$, $f \in F$, $\theta_j \in \Theta_j$, let $\bar{x}(i, f, \theta_j)$ be an arbitrary alternative in $S_i(f, \theta_j)$.[58]

For any $f, f' \in F$, $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$, let $\bar{x}(f, f', \theta_1, \theta_2)$ be an arbitrary alternative in $S_1(f, \theta_2) \cap S_2(f', \theta_1)$. Such an alternative $\bar{x}(f, f', \theta_1, \theta_2) \in \bar{X}$ exists since $S_1(f, \theta_2) \cap S_2(f', \theta_1)$ is non-empty for each $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, by (iii) of two-individual consistency (see Definition 9).

---

any alternative that is not in $\bar{X}$ is non-essential for the design problem. Note that such an alternative is not going to be in the opportunity set of any individuals and it is never to be implemented.

[57] We note that we need at least three individuals for our mechanism to be well defined. Otherwise, Rule 2 in the mechanism becomes ambiguous.

[58] Such an $\bar{x}(i, f, \theta_j)$ exists as $S_i(f, \theta_j)$ is non-empty due to (i) and (ii) of two-individual consistency.

The mechanism we employ is denoted by $\mu = (M, g)$ where the message space of individual i is $M_i = \{0,1\} \times F \times \Theta_i \times \bar{X} \times \{0,1\}$ and the outcome function equals $g : M \to X$. A generic message is denoted by $m_i = (n_i, f_i, \theta_i, x_i, k_i)$. That is, each individual message's first entry and last entry is required to be either 0 or 1, and each individual i is required to send a message that specifies an SCF $f \in F$, a type $\theta_i \in \Theta_i$, an alternative $x_i \in \bar{X}$.[59] The outcome function g is specified in Table 14.

Rule 1 :	$g(m) = f(\theta)$	if $m_i = (0, f, \theta_i, \cdot, \cdot)$ for both $i \in \{1, 2\}$,
Rule 2.1 :	$g(m) = \begin{cases} x_1 & \text{if } x_1 \in S_1(f_2, \theta_2) \\ \bar{x}(1, f_2, \theta_2) & \text{otherwise,} \end{cases}$	if $m_1 = (1, f_1, \theta_1, x_1, \cdot)$ and $m_2 = (0, f_2, \theta_2, x_2, \cdot)$,
Rule 2.2 :	$g(m) = \begin{cases} x_2 & \text{if } x_2 \in S_2(f_1, \theta_1) \\ \bar{x}(2, f_1, \theta_1) & \text{otherwise,} \end{cases}$	if $m_1 = (0, f_1, \theta_1, x_1, \cdot)$ and $m_2 = (1, f_2, \theta_2, x_2, \cdot)$,
Rule 3 :	$g(m) = \bar{x}(f_1, f_2, \theta_1, \theta_2)$	if $m_1 = (0, f_1, \theta_1, x_1, k_1)$ and $m_2 = (0, f_2, \theta_2, x_2, k_2)$ with $f_1 \neq f_2$,
Rule 4 :	$g(m) = x_j$	if $m_1 = (1, f_1, \theta_1, x_1, k_1)$ and $m_2 = (1, f_2, \theta_2, x_2, k_2)$ with $j = \begin{cases} 1 & \text{if } k_1 + k_2 \text{ is odd,} \\ 2 & \text{if } k_1 + k_2 \text{ is even.} \end{cases}$

Table 14: The outcome function of the two individual mechanism.

Rule 1 indicates that if the first entries of both individuals' messages are 0 and there is agreement between the two individuals' messages regarding the SCF, then the outcome is determined according to this SCF and the reported type profile in the messages.

Rule 2.1 and Rule 2.2 indicate that if the first entry of the individual messages do not coincide, then the outcome is the alternative proposed by individual i whose message's first entry is 1 whenever this alternative is in $S_i(f_j, \theta_j)$ where $j \neq i$ is the individual whose message's first entry is 0. Otherwise, the outcome is $\bar{x}(i, f_j, \theta_j)$, also in $S_i(f_j, \theta_j)$.

Rule 3.1 and Rule 3.2 indicate that whenever the first entries of both individuals' messages are 0 but there is no agreement between the individuals' messages regarding the SCF, i.e., $f_1 \neq f_2$, the outcome is $\bar{x}(f_1, f_2, \theta_1, \theta_2) \in S_1(f_2, \theta_2) \cap S_2(f_1, \theta_1)$, which is non-empty due to (iii) of two-individual consistency.

[59] In our mechanism, the first entry of the message of individual i, $n_i \in \{0, 1\}$ with $i = 1, 2$, parallels with the "flag" or "no flag" choice featured in the mechanism of Dutta and Sen (1991).

Rule 4 indicates that if the first entries of both individuals' messages are 1, then the outcome is determined according to the sum of the last entries of the messages. If this sum is odd, then the outcome is the alternative proposed by individual 1 and if this sum is even, the outcome is the alternative proposed by individual 2.

When ex-post implementation of an SCF is desired, i.e., $\#F = 1$, then (iii) of choice incompatibility holds vacuously while Rule 3 of our mechanism becomes redundant. This simplifies our proofs by eliminating discussions and arguments about Rule 3.

B Proofs

B.1 Proof of Claim 1

We identify all EPE of $\mu = (M, g)$ by a case by case analysis on what Alice plays when her type is ρ_A. Let σ^* be an ex-post equilibrium of $\mu = (M, g)$.

Case 1. If $\sigma_A^*(\rho_A) = U$: Then, $O_B^\mu(\sigma_A^*(\rho_A)) = \{c, n\}$. At (ρ_A, ρ_B) and (ρ_A, γ_B), Bob chooses n from the set $\{c, n\}$. Thus, $\sigma_B^*(\rho_B)$ and $\sigma_B^*(\gamma_B)$ must be either L or R.

Subcase 1.1. If $\sigma_B^*(\rho_B) = L$ and $\sigma_B^*(\gamma_B) = L$: Then, $g(\sigma^*(\rho_A, \rho_B)) = n = g(\sigma^*(\rho_A, \gamma_B))$. We have $O_A^\mu(\sigma_B^*(\rho_B)) = O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, n\}$. At (γ_A, ρ_B), Alice chooses n from $\{c, n\}$ and hence $\sigma_A^*(\gamma_A)$ must be either U or D. But, at (γ_A, γ_B), Alice chooses c from $\{c, n\}$ which implies $\sigma_A^*(\gamma_A)$ must be M, a contradiction. Thus, we cannot have $\sigma_B^*(\rho_B) = L$ and $\sigma_B^*(\gamma_B) = L$.

Subcase 1.2. If $\sigma_B^*(\rho_B) = L$ and $\sigma_B^*(\gamma_B) = R$: Then, $g(\sigma^*(\rho_A, \rho_B)) = n = g(\sigma^*(\rho_A, \gamma_B))$. We have $O_A^\mu(\sigma_B^*(\rho_B)) = \{c, n\}$ and $O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, n, s\}$. At (γ_A, ρ_B), Alice chooses n from $\{c, n\}$, which implies $\sigma_A^*(\gamma_A)$ must be either U or D. At (γ_A, γ_B), Alice chooses c and s from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be M or D. Therefore, we must have $\sigma_A^*(\gamma_A) = D$.

Indeed the following observations imply that our first EPE is σ'^* such that $\sigma_A'^*(\rho_A) = U$, $\sigma_A'^*(\gamma_A) = D$, and $\sigma_B'^*(\rho_B) = L$, $\sigma_B'^*(\gamma_B) = R$

$$
\begin{aligned}
\text{At } (\rho_A, \rho_B) \quad : \quad & n \in C_A^{(\rho_A,\rho_B)}(\{c, n\}) \implies g(\sigma'^*(\rho_A, \rho_B)) \in C_A^{(\rho_A,\rho_B)}(O_A^\mu(\sigma_B'^*(\rho_B))), \\
& n \in C_B^{(\rho_A,\rho_B)}(\{c, n\}) \implies g(\sigma'^*(\rho_A, \rho_B)) \in C_B^{(\rho_A,\rho_B)}(O_B^\mu(\sigma_A'^*(\rho_A))). \\
\text{At } (\rho_A, \gamma_B) \quad : \quad & n \in C_A^{(\rho_A,\gamma_B)}(\{c, n, s\}) \implies g(\sigma'^*(\rho_A, \gamma_B)) \in C_A^{(\rho_A,\gamma_B)}(O_A^\mu(\sigma_B'^*(\gamma_B))), \\
& n \in C_B^{(\rho_A,\gamma_B)}(\{c, n\}) \implies g(\sigma'^*(\rho_A, \gamma_B)) \in C_B^{(\rho_A,\gamma_B)}(O_B^\mu(\sigma_A'^*(\rho_A))). \\
\text{At } (\gamma_A, \rho_B) \quad : \quad & n \in C_A^{(\gamma_A,\rho_B)}(\{c, n\}) \implies g(\sigma'^*(\gamma_A, \rho_B)) \in C_A^{(\gamma_A,\rho_B)}(O_A^\mu(\sigma_B'^*(\rho_B))), \\
& n \in C_B^{(\gamma_A,\rho_B)}(\{n, s\}) \implies g(\sigma'^*(\gamma_A, \rho_B)) \in C_B^{(\gamma_A,\rho_B)}(O_B^\mu(\sigma_A'^*(\gamma_A))).
\end{aligned}
$$

$$\begin{aligned}\text{At } (\gamma_A,\gamma_B) \;:\;\; & s \in C_A^{(\gamma_A,\gamma_B)}(\{c,n,s\}) \implies g(\sigma'^*(\gamma_A,\gamma_B)) \in C_A^{(\gamma_A,\gamma_B)}(O_A^\mu(\sigma_B'^*(\gamma_B))),\\ & s \in C_B^{(\gamma_A,\gamma_B)}(\{n,s\}) \implies g(\sigma'^*(\gamma_A,\gamma_B)) \in C_B^{(\gamma_A,\gamma_B)}(O_B^\mu(\sigma_A'^*(\gamma_A))).\end{aligned}$$

Subcase 1.3. If $\sigma_B^*(\rho_B) = R$ and $\sigma_B^*(\gamma_B) = L$: Then, $g(\sigma^*(\rho_A,\rho_B)) = n = g(\sigma^*(\rho_A,\gamma_B))$. We have $O_A^\mu(\sigma_B^*(\rho_B)) = \{c,n,s\}$ and $O_A^\mu(\sigma_B^*(\gamma_B)) = \{c,n\}$. At (γ_A,ρ_B), Alice chooses n from $\{c,n,s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. On the other hand, at (γ_A,γ_B), Alice chooses n from $\{c,n\}$, which implies $\sigma_A^*(\gamma_A)$ must be U or D. Therefore, we must have $\sigma_A^*(\gamma_A) = U$. This implies $O_B^\mu(\sigma_A^*(\gamma_A)) = \{c,n\}$. But, at (γ_A,ρ_B), Bob chooses c from $\{c,n\}$ even though it would be $g(\sigma^*(\gamma_A,\rho_B)) = n$. Hence, we cannot have $\sigma_B^*(\rho_B) = R$ and $\sigma_B^*(\gamma_B) = L$.

Subcase 1.4. If $\sigma_B^*(\rho_B) = R$ and $\sigma_B^*(\gamma_B) = R$: Then, $g(\sigma^*(\rho_A,\rho_B)) = n = g(\sigma^*(\rho_A,\gamma_B))$. We have $O_A^\mu(\sigma_B^*(\rho_B)) = O_A^\mu(\sigma_B^*(\gamma_B)) = \{c,n,s\}$. At (γ_A,ρ_B), Alice chooses n from $\{c,n,s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. But, at (γ_A,γ_B), Alice chooses c and s from $\{c,n,s\}$, which implies $\sigma_A^*(\gamma_A)$ must be either M or D, a contradiction. Therefore, we cannot have $\sigma_B^*(\rho_B) = R$ and $\sigma_B^*(\gamma_B) = R$.

Case 2. If $\sigma_A^*(\rho_A) = M$: Then, $O_B^\mu(\sigma_A^*(\rho_A)) = \{c,s\}$. At (ρ_A,ρ_B) and (ρ_A,γ_B), Bob chooses s from the set $\{c,s\}$. Therefore, $\sigma_B^*(\rho_B)$ and $\sigma_B^*(\gamma_B)$ must both be M. Then, $O_A^\mu(\sigma_B^*(\rho_B)) = O_A^\mu(\sigma_B^*(\gamma_B)) = \{c,s\}$. At (γ_A,ρ_B) and (γ_A,γ_B) Alice chooses c from the set $\{c,s\}$, which implies it must be that $\sigma_A^*(\rho_A) = U$.

Indeed the following observations imply that our second EPE is σ'''^* such that $\sigma_A'''^*(\rho_A) = M$, $\sigma_A'''^*(\gamma_A) = U$, and $\sigma_B'''^*(\rho_B) = M$, $\sigma_B'''^*(\gamma_B) = M$

$$\begin{aligned}\text{At } (\rho_A,\rho_B) \;:\;\; & s \in C_A^{(\rho_A,\rho_B)}(\{c,s\}) \implies g(\sigma'''^*(\rho_A,\rho_B)) \in C_A^{(\rho_A,\rho_B)}(O_A^\mu(\sigma_B'''^*(\rho_B))),\\ & s \in C_B^{(\rho_A,\rho_B)}(\{c,s\}) \implies g(\sigma'''^*(\rho_A,\rho_B)) \in C_B^{(\rho_A,\rho_B)}(O_B^\mu(\sigma_A'''^*(\rho_A))).\\ \text{At } (\rho_A,\gamma_B) \;:\;\; & s \in C_A^{(\rho_A,\gamma_B)}(\{c,s\}) \implies g(\sigma'''^*(\rho_A,\gamma_B)) \in C_A^{(\rho_A,\gamma_B)}(O_A^\mu(\sigma_B'''^*(\gamma_B))),\\ & s \in C_B^{(\rho_A,\gamma_B)}(\{c,s\}) \implies g(\sigma'''^*(\rho_A,\gamma_B)) \in C_B^{(\rho_A,\gamma_B)}(O_B^\mu(\sigma_A'''^*(\rho_A))).\\ \text{At } (\gamma_A,\rho_B) \;:\;\; & c \in C_A^{(\gamma_A,\rho_B)}(\{c,s\}) \implies g(\sigma'''^*(\gamma_A,\rho_B)) \in C_A^{(\gamma_A,\rho_B)}(O_A^\mu(\sigma_B'''^*(\rho_B))),\\ & c \in C_B^{(\gamma_A,\rho_B)}(\{c,n\}) \implies g(\sigma'''^*(\gamma_A,\rho_B)) \in C_B^{(\gamma_A,\rho_B)}(O_B^\mu(\sigma_A'''^*(\gamma_A))).\\ \text{At } (\gamma_A,\gamma_B) \;:\;\; & c \in C_A^{(\gamma_A,\gamma_B)}(\{c,s\}) \implies g(\sigma'''^*(\gamma_A,\gamma_B)) \in C_A^{(\gamma_A,\gamma_B)}(O_A^\mu(\sigma_B'''^*(\gamma_B))),\\ & c \in C_B^{(\gamma_A,\gamma_B)}(\{c,n\}) \implies g(\sigma'''^*(\gamma_A,\gamma_B)) \in C_B^{(\gamma_A,\gamma_B)}(O_B^\mu(\sigma_A'''^*(\gamma_A))).\end{aligned}$$

Case 3. If $\sigma_A^*(\rho_A) = D$: Then, $O_B^\mu(\sigma_A^*(\rho_A)) = \{n,s\}$. At (ρ_A,ρ_B) and (ρ_A,γ_B), Bob chooses s from the set $\{n,s\}$. Therefore, $\sigma_B^*(\rho_B)$ and $\sigma_B^*(\gamma_B)$ must be either M or R.

Subcase 3.1. If $\sigma_B^*(\rho_B) = M$ and $\sigma_B^*(\gamma_B) = M$: So, $g(\sigma^*(\rho_A,\rho_B)) = s = g(\sigma^*(\rho_A,\gamma_B))$.

We have $O^\mu_A(\sigma^*_B(\rho_B)) = O^\mu_A(\sigma^*_B(\gamma_B)) = \{c, s\}$. At (γ_A, ρ_B) and (γ_A, γ_B), Alice chooses c from $\{c, s\}$, which implies it must be $\sigma^*_A(\gamma_A) = U$.

Indeed the following observations imply that our third EPE is σ''^* such that $\sigma''^*_A(\rho_A) = D$, $\sigma''^*_A(\gamma_A) = U$, and $\sigma''^*_B(\rho_B) = M$, $\sigma''^*_B(\gamma_B) = M$.

$$\begin{aligned}
\text{At } (\rho_A, \rho_B) \quad : \quad & s \in C_A^{(\rho_A,\rho_B)}(\{c, s\}) \implies g(\sigma'''^*(\rho_A, \rho_B)) \in C_A^{(\rho_A,\rho_B)}(O^\mu_A(\sigma''^*_B(\rho_B))), \\
& s \in C_B^{(\rho_A,\rho_B)}(\{n, s\}) \implies g(\sigma''^*(\rho_A, \rho_B)) \in C_B^{(\rho_A,\rho_B)}(O^\mu_B(\sigma''^*_A(\rho_A))). \\
\text{At } (\rho_A, \gamma_B) \quad : \quad & s \in C_A^{(\rho_A,\gamma_B)}(\{c, s\}) \implies g(\sigma''^*(\rho_A, \gamma_B)) \in C_A^{(\rho_A,\gamma_B)}(O^\mu_A(\sigma''^*_B(\gamma_B))), \\
& s \in C_B^{(\rho_A,\gamma_B)}(\{n, s\}) \implies g(\sigma''^*(\rho_A, \gamma_B)) \in C_B^{(\rho_A,\gamma_B)}(O^\mu_B(\sigma''^*_A(\rho_A))). \\
\text{At } (\gamma_A, \rho_B) \quad : \quad & c \in C_A^{(\gamma_A,\rho_B)}(\{c, s\}) \implies g(\sigma''^*(\gamma_A, \rho_B)) \in C_A^{(\gamma_A,\rho_B)}(O^\mu_A(\sigma''^*_B(\rho_B))), \\
& c \in C_B^{(\gamma_A,\rho_B)}(\{c, n\}) \implies g(\sigma''^*(\gamma_A, \rho_B)) \in C_B^{(\gamma_A,\rho_B)}(O^\mu_B(\sigma''^*_A(\gamma_A))). \\
\text{At } (\gamma_A, \gamma_B) \quad : \quad & c \in C_A^{(\gamma_A,\gamma_B)}(\{c, s\}) \implies g(\sigma''^*(\gamma_A, \gamma_B)) \in C_A^{(\gamma_A,\gamma_B)}(O^\mu_A(\sigma''^*_B(\gamma_B))), \\
& c \in C_B^{(\gamma_A,\gamma_B)}(\{c, n\}) \implies g(\sigma''^*(\gamma_A, \gamma_B)) \in C_B^{(\gamma_A,\gamma_B)}(O^\mu_B(\sigma''^*_A(\gamma_A))).
\end{aligned}$$

Subcase 3.2. If $\sigma^*_B(\rho_B) = M$ and $\sigma^*_B(\gamma_B) = R$: So, $g(\sigma^*(\rho_A, \rho_B)) = s = g(\sigma^*(\rho_A, \gamma_B))$. We have $O^\mu_A(\sigma^*_B(\rho_B)) = \{c, s\}$ and $O^\mu_A(\sigma^*_B(\gamma_B)) = \{c, n, s\}$. At (γ_A, ρ_B), Alice chooses c from $\{c, s\}$, which implies $\sigma^*_A(\gamma_A)$ must be U. On the other hand, at (γ_A, γ_B), Alice chooses c and s from $\{c, n, s\}$, which implies $\sigma^*_A(\gamma_A)$ must be M or D, a contradiction. Hence, we cannot have $\sigma^*_B(\rho_B) = M$ and $\sigma^*_B(\gamma_B) = R$.

Subcase 3.3. If $\sigma^*_B(\rho_B) = R$ and $\sigma^*_B(\gamma_B) = M$: So, $g(\sigma^*(\rho_A, \rho_B)) = s = g(\sigma^*(\rho_A, \gamma_B))$. We have $O^\mu_A(\sigma^*_B(\rho_B)) = \{c, n, s\}$ and $O^\mu_A(\sigma^*_B(\gamma_B)) = \{c, s\}$. At (γ_A, ρ_B), Alice chooses n from $\{c, n, s\}$, and at (γ_A, γ_B), Alice chooses c from $\{c, n\}$. They both imply we must have $\sigma^*_A(\gamma_A) = U$. Thus, $O^\mu_B(\sigma^*_A(\gamma_A)) = \{c, n\}$. But, at (γ_A, ρ_B), Bob chooses c from $\{c, n\}$ even though it would be $g(\sigma^*(\gamma_A, \rho_B)) = n$. So, we cannot have $\sigma^*_B(\rho_B) = R$ and $\sigma^*_B(\gamma_B) = M$.

Subcase 3.4. If $\sigma^*_B(\rho_B) = R$ and $\sigma^*_B(\gamma_B) = R$: So, $g(\sigma^*(\rho_A, \rho_B)) = s = g(\sigma^*(\rho_A, \gamma_B))$. We have $O^\mu_A(\sigma^*_B(\rho_B)) = O^\mu_A(\sigma^*_B(\gamma_B)) = \{c, n, s\}$. At (γ_A, ρ_B), Alice chooses n from $\{c, n, s\}$, which implies $\sigma^*_A(\gamma_A)$ must be U. On the other hand, at (γ_A, γ_B), Alice chooses c, s from $\{c, n, s\}$, which implies $\sigma^*_A(\gamma_A)$ must be M or D, a contradiction. Thus, we cannot have $\sigma^*_B(\rho_B) = R$ and $\sigma^*_B(\gamma_B) = R$ as well.

Therefore, there are exactly three EPE of the mechanism $\mu = (M, g)$, σ'^*, σ''^*, and σ'''^*, as identified above where $g(\sigma''^*(\theta)) = g(\sigma'''^*(\theta))$ for all $\theta \in \Theta$: $g(\sigma''^*(\rho_A, \rho_B)) = g((D, M)) = g((M, M)) = s = g(\sigma'''^*(\rho_A, \rho_B))$; $g(\sigma''^*(\rho_A, \gamma_B)) = g((D, M)) = g((M, M)) = s = g(\sigma'''^*(\rho_A, \gamma_B))$; $g(\sigma''^*(\gamma_A, \rho_B)) = g((U, M)) = c = g(\sigma'''^*(\gamma_A, \rho_B))$; and $g(\sigma''^*(\gamma_A, \gamma_B)) = g((U, M)) = c = g(\sigma'''^*(\gamma_A, \gamma_B))$. $\square$

B.2 Proof of Theorem 1

Let $\mu = (M, g)$ be a mechanism that ex-post implements the non-empty SCS $F \in \mathcal{F}$. Consider any SCF $f \in F$. By (i) of Definition 2, there exists an EPE σ^f of μ such that $f = g \circ \sigma^f$.

By definition of EPE, we have for each $\theta \in \Theta$, $g(\sigma^f(\theta))$ is an element of $C_i^\theta(O_i^\mu(\sigma_{-i}^f(\theta_{-i})))$ for all $i \in N$. Therefore, for each $\theta \in \Theta$, $f(\theta) \in C_i^\theta(O_i^\mu(\sigma_{-i}^f(\theta_{-i})))$ for all $i \in N$. Setting $S_i(f, \theta_{-i}) := O_i^\mu(\sigma_{-i}^f(\theta_{-i}))$, we get for each $\theta \in \Theta$, $f(\theta) \in C_i^\theta(S_i(f, \theta_{-i}))$ for all $i \in N$. Since $f \in F$ is arbitrary, this means for each $f \in F$, $i \in N$, $\theta_{-i} \in \Theta_{-i}$, there exists $S_i(f, \theta_{-i}) \subset X$ such that $f(\theta) \in C_i^\theta(S_i(f, \theta_{-i}))$ as long as θ is compatible with θ_{-i}, i.e., $\theta = (\theta'_i, \theta_{-i})$ for some $\theta'_i \in \Theta_i$. Therefore, $f(\theta'_i, \theta_{-i}) \in C_i^{(\theta'_i, \theta_{-i})}(S_i(f, \theta_{-i}))$ for each $\theta'_i \in \Theta_i$ holds for every set in the collection $\{S_i(f, \theta_{-i}) | f \in F, i \in N, \theta_{-i} \in \Theta_{-i}\}$.

On the other hand, if a deception profile α is such that $f \circ \alpha \notin F$, $\sigma^f \circ \alpha$ cannot be an EPE of $\mu = (M, g)$. Otherwise, by (ii) of Definition 2, there exists $\tilde{f} \in F$ with $\tilde{f} = g \circ \sigma^f \circ \alpha$. But, since $f = g \circ \sigma^f$, we have $\tilde{f} = f \circ \alpha \in F$, a contradiction. Therefore, there exists $\theta^* \in \Theta$, $i^* \in N$ such that $g(\sigma^f(\alpha(\theta^*))) \notin C_{i^*}^{\theta^*}(O_{i^*}^\mu(\sigma_{-i^*}^f(\alpha_{-i^*}(\theta_{-i^*}))))$. Since $g \circ \sigma^f = f$ and $O_{i^*}^\mu(\sigma_{-i^*}^f(\alpha_{-i^*}(\theta_{-i^*}))) = S_{i^*}(f, \alpha_{-i^*}(\theta_{-i^*}))$, we get $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta_{-i^*})))$. $\square$

B.3 Proofs of Propositions 1, 2, and 3 and Lemma 1

Proof of Proposition 1. Let $\mathbb{S}$ be a non-empty collection of sets consistent with an SCS F under incomplete information and let $S^* := S_{i^*}(f, \alpha_{-i^*}(\theta_{-i^*})) \in \mathbb{S}$. Then, condition (i) of ex-post choice monotonicity follows from condition (ii) of Definition 3 while condition (ii) of ex-post choice monotonicity follows from (i) of Definition 3. $\square$

Proof of Proposition 2. Let $\mathbb{S}$ be a non-empty collection of sets consistent with an SCS F under incomplete information and take any $f \in F$, $\theta \in \Theta$, $i \in N$ and let $S := S_i(f, \theta_{-i}) \in \mathbb{S}$. By (i) of Definition 3, $f(\theta) \in C_i^\theta(S_i(f, \theta_{-i}))$ implies $f(\theta) \in C_i^\theta(S)$ establishing condition (i) of quasi-ex-post choice incentive compatibility. Furthermore, since $f(\theta'_i, \theta_{-i}) \in C_i^{(\theta'_i, \theta_{-i})}(S_i(f, \theta_{-i}))$ for each $\theta'_i \in \Theta_i$ due to (i) of Definition 3, we have $f(\theta'_i, \theta_{-i}) \in S$ for each $\theta'_i \in \Theta_i$ establishing condition (ii) of quasi-ex-post choice incentive compatibility. $\square$

Proof of Lemma 1. The proof directly follows from the fact that whenever f is partially truthfully (ex-post) implemented by the direct mechanism $g^d : \Theta \to X$, the opportunity set of any individual $i \in N$ under truthtelling is $\{f(\theta'_i, \theta_{-i}) | \theta'_i \in \Theta_i\}$, i.e., $O_i^{g^d}(\theta_{-i}) = \{f(\theta'_i, \theta_{-i}) | \theta'_i \in \Theta_i\}$. $\square$

Proof of Proposition 3. Suppose the individual choices satisfy the IIA and let f be partially (ex-post) implemented by the mechanism μ. Then, Theorem 1 together with Proposition 2 implies that f is quasi-ex-post choice incentive compatible. That is, for every $\theta \in \Theta$, $i \in N$ there exists $S \in \mathcal{X}$ such that $f(\theta) \in C_i^{\theta}(S)$ and $\{f(\theta_i', \theta_{-i}) | \theta_i' \in \Theta_i\} \subseteq S$. Hence, by the IIA, we must have $f(\theta) \in C_i^{\theta}(\{f(\theta_i', \theta_{-i}) | \theta_i' \in \Theta_i\})$. Therefore, by Lemma 1, the revelation principle holds. □

B.4 Proof of Theorem 2

Consider the mechanism $\mu = (M, g)$ constructed in section A.1.

First, we show that for any $f \in F$, there exists an EPE, σ^f, of $\mu = (M, g)$ such that $f = g \circ \sigma^f$. This implies that condition (i) of ex-post implementability (see Definition 2) holds: Take any $f \in F$, let $\sigma_i^f(\theta_i) = (f, \theta_i, x, 1)$ for each $i \in N$ and for some arbitrary $x \in \bar{X}$. By Rule 1, we have $g(\sigma^f(\theta)) = f(\theta)$ for each $\theta \in \Theta$, i.e., $f = g \circ \sigma^f$. Observe that for any unilateral deviation by individual i from σ^f, either Rule 1 or Rule 2 applies, i.e., Rule 3 is not attainable by any unilateral deviation from σ^f. If individual i deviates to $m_i = (f, \theta_i, x', n')$ when his/her type is θ_i, then Rule 1 continues to apply at θ and the outcome continues to be $f(\theta)$, which is in $S_i(f, \theta_{-i})$ since, by condition (i) of consistency, $f(\theta) \in C_i^{\theta}(S_i(f, \theta_{-i}))$. If individual i deviates to $m_i = (f, \theta_i', x', n')$ with $\theta_i' \neq \theta_i$ when his/her type is θ_i, then Rule 1 continues to apply at θ and the outcome at θ becomes $f(\theta_i', \theta_{-i})$, which is in $S_i(f, \theta_{-i})$ as well since $f(\theta_i', \theta_{-i}) \in C_i^{(\theta_i', \theta_{-i})}(S_i(f, \theta_{-i}))$, again by condition (i) of consistency. If individual i deviates to $m_i = (f', \theta_i', x', n')$ with $f' \neq f$ when his/her type is θ_i, then Rule 2 applies at θ and the outcome at θ becomes x' if x' is in $S_i(f, \theta_{-i})$, and otherwise $\bar{x}(i, f, \theta_{-i})$, which is already in $S_i(f, \theta_{-i})$ as well. This means, as $S_i(f, \theta_{-i}) \subset \bar{X}$ for each $\theta \in \Theta$, $i \in N$, under σ^f, at any $\theta \in \Theta$, by unilateral deviations, individual i can generate every alternative in $S_i(f, \theta_{-i})$ and nothing else. That is, by construction, $O_i^{\mu}(\sigma_{-i}^f(\theta_{-i})) = S_i(f, \theta_{-i})$ for each $\theta \in \Theta$, $i \in N$. Since, by (i) of consistency, $f(\theta) \in C_i^{\theta}(S_i(f, \theta_{-i}))$ for each $i \in N$, we have for each $\theta \in \Theta$, $g(\sigma^f(\theta)) \in C_i^{\theta}(O_i^{\mu}(\sigma_{-i}^f(\theta_{-i})))$ for all $i \in N$, i.e., σ^f is an EPE of μ such that $f = g \circ \sigma^f$.

Consider now any EPE σ^* of μ denoted as $\sigma_i^*(\theta_i) = (f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$ for each $i \in N$. That is, $f_i(\theta_i)$ denotes the SCF proposed by individual i when his/her type is θ_i; $\alpha_i(\theta_i)$ denotes the reported type of individual i when his/her type is θ_i; $x_i(\theta_i)$ denotes the alternative proposed by individual i when his/her type is θ_i; and $k_i(\theta_i)$ denotes the number proposed by individual i when his/her type is θ_i.

Next, we show that, under any EPE σ^* of μ, Rule 1 must apply at each $\theta \in \Theta$: Suppose, for contradiction, that either Rule 2 or Rule 3 applies at some $\tilde{\theta} \in \Theta$ under

σ^*. If Rule 2 applies at $\tilde{\theta}$, by construction, we have $O_j^\mu(\sigma^*_{-j}(\tilde{\theta}_{-j})) = S_j(f, \alpha_j(\tilde{\theta}_{-j}))$ for the odd-man-out $j \in N$ and $O_i^\mu(\sigma^*_{-i}(\tilde{\theta}_{-i})) = \bar{X}$ for all $i \neq j$, i.e., for all the other $n-1$ individuals. On the other hand, if Rule 3 applies at $\tilde{\theta}$, we have, by construction, $O_i^\mu(\sigma^*_{-i}(\tilde{\theta}_{-i})) = \bar{X}$ for all $i \in N$. Therefore, under both Rule 2 and Rule 3, at least $n-1$ individuals have the opportunity set $\bar{X}$. Since σ^* is an EPE of μ, it follows that $g(\sigma^*(\tilde{\theta})) \in C_i^\theta(\bar{X})$ for at least $n-1$ individuals. But this contradicts the choice incompatible pair property of $\bar{X}$ at $\tilde{\theta}$. Therefore, under any EPE σ^* of μ, Rule 1 must apply at each $\theta \in \Theta$.

Moreover, under any EPE σ^* of μ, there is a unique $f \in F$ such that $f_i(\theta_i) = f$ for all $i \in N$ and for all $\theta_i \in \Theta_i$. To see why, fix an EPE σ^* of μ, pick an arbitrary $\theta \in \Theta$, and as Rule 1 must apply at $\theta \in \Theta$ under σ^*, let $f_i(\theta_i) = f$ for all $i \in N$ under σ^*. Suppose, for contradiction, that there exists $i_0 \in N$, $\theta_{i_0} \in \Theta_{i_0}$ such that $f_{i_0}(\theta_{i_0}) \neq f$. Without loss of generality, suppose $i_0 = 1$ and $\hat{\theta}_1 \in \Theta_1$ such that $f_1(\hat{\theta}_1) \neq f$. But, then, under the EPE σ^*, Rule 1 cannot apply at state $(\hat{\theta}_1, \theta_{-1}) \in \Theta$, as $f_1(\hat{\theta}_1) \neq f$ and $f_j(\theta_j) = f$ for all $j \neq 1$ under σ^*, a contradiction.

Therefore, for any EPE σ^* of μ, there exists a unique $f \in F$ such that $f_i(\theta_i) = f$ for all $i \in N$ and for all $\theta_i \in \Theta_i$. Hence, by Rule 1, $g(\sigma^*(\theta)) = f(\alpha(\theta))$ for each $\theta \in \Theta$. That is, $g \circ \sigma^* = f \circ \alpha$.

Finally, we show that it must be that $f \circ \alpha \in F$: Since Rule 1 applies at each $\theta \in \Theta$, and each $i \in N$ reports the type $\alpha_i(\theta_i) \in \Theta_i$ as the second entry of their messages at $\theta \in \Theta$ under σ^*, by construction, we have, at each $\theta \in \Theta$, $O_i^\mu(\sigma^*_{-i}(\theta_{-i})) = S_i(f, \alpha_{-i}(\theta_{-i}))$ for all $i \in N$. If $f \circ \alpha \notin F$, then by (ii) of consistency (see Definition 3), there exists $\theta^* \in \Theta$, $i^* \in N$ such that $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta_{-i^*})))$. But this implies $g(\sigma^*(\theta^*)) \notin C_{i^*}^{\theta^*}(O_{i^*}^\mu(\sigma^*_{-i^*}(\theta^*_{-i^*})))$, a contradiction to σ^* being an EPE of μ. That is, we must have $f \circ \alpha \in F$, as desired. Therefore, $g \circ \sigma^* = f \circ \alpha \in F$, which implies that condition (ii) of ex-post implementability holds as well. □

B.5 Proof of Theorem 3

Consider the mechanism $\mu = (M, g)$ constructed in section A.1.

As shown in the proof of Theorem 2, for any $f \in F$, $\sigma_i^f(\theta_i) = (f, \theta_i, x, 1)$ for each $i \in N$ (for arbitrary $x \in \bar{X}$) is an EPE of μ such that $f = g \circ \sigma^f$. That is, for any $f \in F$, there exists an EPE, σ^f, of μ such that $f = g \circ \sigma^f$, which implies that condition (i) of ex-post implementability (refer to Definition 2) holds.

Now, consider an EPE σ^* of $\mu = (M, g)$ represented as before by $\sigma^*(\theta_i) = (f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$. For any $f \in F$ and $i \in N$, let $\bar{\Theta}_i^f := \{\theta_i \in \Theta_i | f_i(\theta_i) = f\}$.

That is, $\bar{\Theta}_i^f \subset \bar{\Theta}_i$ is the set of types of individual i where the first entry of his/her message—his/her proposed SCF—is f under σ^*. Let $\bar{\Theta}_f := \times_{i\in N}\bar{\Theta}_i^f$. That is, $\bar{\Theta}_f$ is the set of states where all of the individuals propose the SCF $f \in F$ under σ^*. Consider the collection of product sets $\{\bar{\Theta}_f\}_{f\in F}$. Observe that $\bar{\Theta} := \bigcup_{f\in F}\bar{\Theta}_f$ describes the set of states where Rule 1 applies under σ^*.

Thus, at any $\theta \in \Theta\setminus\bar{\Theta}$, either Rule 2 or Rule 3 applies, which means $O_i^\mu(\sigma^*_{-i}(\theta_{-i})) = \bar{X}$ for at least $n-1$ individuals for any $\theta \in \Theta\setminus\bar{\Theta}$. Furthermore, σ^* being an EPE of μ implies $g(\sigma^*(\theta)) \in C_i^\theta(\bar{X})$ for at least $n-1$ individuals. Hence, we have, by (iii) of consistency-no-veto (see Definition 8), there exists $f^* \in F$ such that $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \Theta\setminus\bar{\Theta}$.

Next, we show that it must also be that $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \bar{\Theta}$. Suppose not, for contradiction, then there exists $\tilde{\theta} \in \bar{\Theta}_f$ for some $f \in F$ such that $g(\sigma^*(\tilde{\theta})) \neq f^*(\tilde{\theta})$. Since $\tilde{\theta} \in \bar{\Theta}_f$, we have $f_i(\tilde{\theta}_i) = f$ for all $i \in N$. Thus, Rule 1 applies at $\tilde{\theta}$ under σ^*, and hence $g(\sigma^*(\tilde{\theta})) = f(\alpha(\tilde{\theta}))$ where α is the deception profile induced by σ^*. This means, as $g(\sigma^*(\tilde{\theta})) \neq f^*(\tilde{\theta})$, we have $f(\alpha(\tilde{\theta})) \neq f^*(\tilde{\theta})$. Then, by (iv) of consistency-no-veto, there exists $i^* \in N$ and $\theta^* \in \bar{\Theta}_f$ such that $f(\alpha(\tilde{\theta})) \notin C_{i^*}^{\theta^*}(S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*})))$. But, since Rule 1 applies at $\tilde{\theta}$ under σ^*, by construction, $O_{i^*}^\mu(\sigma^*_{-i^*}(\tilde{\theta}_{-i^*})) = S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*}))$, which implies $g(\sigma^*(\tilde{\theta})) \notin O_{i^*}^\mu(\sigma^*_{-i^*}(\tilde{\theta}_{-i^*}))$, a contradiction to σ^* being an EPE of μ.

Therefore, $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \Theta$. That is, condition (ii) of ex-post implementability holds as well. □

B.6 Proof of Theorem 4

Let F be ex-post implementable by $\mu = (M, g)$. For any $f \in F$, set $S_i(f, \theta_j) := O_i^\mu(\sigma_j^f(\theta_j))$ for $i \neq j$, where σ^f is the EPE of μ such that $f = g \circ \sigma^f$, which exists by (i) of ex-post implementability (see Definition 2). It is also easy to see that (i) and (ii) follows from (i) of consistency (see Definition 3) and Theorem 1; (iv) follows from (ii) of consistency and Theorem 1. The only new condition that requires a proof is (iii).

Take any $f, f' \in F$, if $f = f'$, then, by (i) and (ii), $f(\theta_1, \theta_2) = f'(\theta_1, \theta_2) \in S_1(f, \theta_2) \cap S_2(f', \theta_1)$ for each $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$. Suppose $f \neq f'$. Recall that $S_1(f, \theta_2) := O_1(\sigma_2^f(\theta_2))$ and $S_2(f', \theta_1) := O_2(\sigma_1^{f'}(\theta_1))$ where σ^f and $\sigma^{f'}$ are some EPE of μ such that $f = g \circ \sigma^f$ and $f' = g \circ \sigma^{f'}$. Consider any $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$ and let $m_1' = \sigma_1^{f'}(\theta_1)$, $m_2' = \sigma^f(\theta_2)$. Then, it follows from $S_1(f, \theta_2)$ and $S_2(f', \theta_1)$ being opportunity sets as defined above that $g(m_1', m_2') = g(m_1', \sigma_2^f(\theta_2)) \in S_1(f, \theta_2)$ and $g(m_1', m_2') = g(\sigma_1^{f'}(\theta_1), m_2') \in S_2(f', \theta_1)$. Therefore, we must have $g(m_1', m_2') = g(\sigma_1^{f'}(\theta_1), \sigma_2^f(\theta_2)) \in S_1(f, \theta_2) \cap S_2(f', \theta_1)$. □

B.7 Proof of Theorem 5

Consider the mechanism $\mu = (M, g)$ constructed in section A.2.

First, we show that for any $f \in F$, there exists an EPE, σ^f, of the mechanism $\mu = (M, g)$ such that $f = g \circ \sigma^f$, which implies condition (i) of ex-post implementation (see Definition 2): Take any $f \in F$, let $\sigma_i^f(\theta_i) = (0, f, \theta_i, x_i, 0)$ for both $i \in \{1, 2\}$ with arbitrary $x_i \in \bar{X}$ for each $i \in \{1, 2\}$. Then, Rule 1 applies at each θ under σ^f. Hence, we have $g(\sigma^f(\theta)) = f(\theta)$ for each $\theta \in \Theta$, i.e., $f = g \circ \sigma^f$. Below, we show that for each $i \in \{1, 2\}$, $O_i(\sigma_j^f(\theta_j)) = S_i(f, \theta_j)$ at each $\theta \in \Theta$ with $i \neq j$.

Without loss of generality, we can focus on individual 1. Observe that $f(\theta) \in S_1(f, \theta_2)$ since, by (i) of two-individual consistency, $f(\theta) \in C_1^\theta(S_1(f, \theta_2))$. That is, if a unilateral deviation by individual 1 does not change the outcome at θ, the outcome is already in $(S_1(f, \theta_2))$. On the other hand, if a unilateral deviation by individual 1 changes the outcome at θ, then either Rule 1 or Rule 2.1 or Rule 3 applies at θ, i.e., Rule 2.2 and Rule 4 cannot be attained by a unilateral deviation of individual 1 at any θ since the first entry of individual 2's message is 0 at any θ under σ^f. Therefore, at any $\theta \in \Theta$, by Rule 2.1, individual 1 can attain any outcome $x \in S_1(f, \theta_2)$ by simply changing his message to $(1, f, \theta_1, x, 0)$. Therefore, $S_1(f, \theta_2) \subset O_1(\sigma_2^f(\theta_2))$ for each $\theta \in \Theta$. To see that, at any $\theta \in \Theta$, individual 1 cannot obtain any other alternative by a unilateral deviation, i.e., $O_1(\sigma_2^f(\theta_2)) \subset S_1(f, \theta_2)$ as well: observe that if Rule 1 continues to apply the outcome at θ is $f(\theta_1', \theta_2)$ for some $\theta_1' \in \Theta_1$ and $f(\theta_1', \theta_2) \in S_1(f, \theta_2)$ as, by (i) of two-individual consistency, $f(\theta_1', \theta_2) \in C_1^{(\theta_1', \theta_2)}(S_1(f, \theta_2))$; when the otherwise part of Rule 2.1 implies, $\bar{x}(1, f, \theta_2) \in S_1(f, \theta_2)$; and when Rule 3 applies, $\bar{x}(f', f, \theta_1', \theta_2) \in S_1(f, \theta_2)$ for each $f' \in F$, $\theta_1' \in \Theta_1$ as well because, by construction, $\bar{x}(f', f, \theta_1', \theta_2) \in S_1(f, \theta_2) \cap S_2(f', \theta_1')$, a non-empty set due to (iii) of two-individual consistency. That is, the mechanism is designed such that, under σ^f, at any θ, by a unilateral deviation, individual 1 can obtain every alternative in $S_1(f, \theta_2)$ and nothing else. Due to symmetry, the same line of proof applies to individual 2 as well. That is, for each $\theta \in \Theta$, $O_i(\sigma_j^f(\theta_j)) = S_i(f, \theta_j)$ for both $i, j \in \{1, 2\}$ with $i \neq j$.

Since, by (i) and (ii) of two-individual consistency, for both $i \in \{1, 2\}$ and for each $\theta \in \Theta$ we have $f(\theta) \in C_i^\theta(S_i(f, \theta_j))$, we have, for each $\theta \in \Theta$, $g(\sigma^f(\theta)) \in C_i^\theta(O_i^\mu(\sigma_j^f(\theta_j)))$ for both $i \in \{1, 2\}$. That is, σ^f is an EPE of μ such that $f = g \circ \sigma^f$, as desired.

Next, we show that for any EPE σ^* of μ, Rule 1 must apply at each $\theta \in \Theta$. Below, we show that other rules are ruled out one by one:

Let σ^* be an EPE of μ denoted as $\sigma_i^*(\theta_i) = (n_i(\theta_i), f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$, $i \in \{1, 2\}$. First, consider Rule 2.1 and Rule 2.2; if Rule 2.1 or Rule 2.2 applies at θ, then the opportunity set of individual i, whose message's first entry is 1, equals $S_i(f, \alpha_j(\theta_j))$, where $\alpha_j(\theta_j)$ denotes the reported type of j at θ and $f_j(\theta_j) = f$. On the other hand, the opportunity set of individual j, whose message's first entry is 0, is $\bar{X}$. Hence, if Rule 2.1 or Rule 2.2 applies at θ under σ^*, $g(\sigma^*(\theta)) \in C_i^\theta(S)$ for some $S = S_i(f, \alpha_j(\theta_j)) \in \mathbb{S}_i$ and $g(\sigma^*(\theta)) \in C_j^\theta(\bar{X})$. This violates (ii) of choice incompatibility at θ.

Next, let us deal with Rule 3: If Rule 3 applies at θ, the opportunity sets of individuals i and j are of the form $S_i(f, \alpha_j(\theta_j))$ and $S_j(f', \alpha_i(\theta_i))$, where $f' = f_i(\theta_i)$ and $f = f_j(\theta_j)$ are the reported SCFs such that $f \neq f'$; $\alpha_i(\theta_i)$ and $\alpha_j(\theta_j)$ are the reported types at θ in the messages of i and j, respectively. This is due to the following: When Rule 3 applies, $g(\sigma^*(\theta)) = \bar{x}(f', f, \alpha_1(\theta_1), \alpha_2(\theta_2))$ which is in $S_i(f, \alpha_j(\theta_j)) \cap S_j(f', \alpha_i(\theta_i))$ due to (iii) of two-individual consistency. When individual i deviates to $(1, f_i(\theta_i), \alpha_i(\theta_i), \tilde{x}, k_i(\theta_i))$ with $\tilde{x} \in S_i(f, \alpha_j(\theta_j))$, either Rule 2.1 or Rule 2.2 applies and the outcome is $\tilde{x}$. Thus, $S_i(f, \alpha_j(\theta_j)) \subset O_i^\mu(\sigma_j^*(\theta_j))$. On the other hand, a deviation of the form $(1, f_i(\theta_i), \alpha_i(\theta_i), \hat{x}, k_i(\theta_i))$ with $\hat{x} \notin S_i(f, \alpha_j(\theta_j))$ implies that either Rule 2.1 or Rule 2.2 applies and the outcome is $\bar{x}(i, f, \alpha_j(\theta_j)) \in S_i(f, \alpha_j(\theta_j))$, by construction. The only possible deviation that leads to another outcome consists of $(0, f, \tilde{\theta}_i, \cdot, \cdot)$. But then, Rule 1 applies and the outcome equals $f(\tilde{\theta}_i, \alpha_j(\theta_j))$ which is again in $S_i(f, \alpha_j(\theta_j))$ due to either (i) or (ii) of two-individual consistency. Hence, $O_i^\mu(\sigma_j^*(\theta_j)) = S_i(f, \alpha_j(\theta_j))$. Thus, if Rule 3 applies at some θ under σ^*, we must have $g(\sigma^*(\theta)) \in C_i^\theta(S_i(f, \alpha_j(\theta_j)))$ and $g(\sigma^*(\theta)) \in C_j^\theta(S_j(f', \alpha_i(\theta_i)))$ with $f \neq f'$. But this violates (iii) of choice incompatibility at θ.

Finally, whenever Rule 4 applies, the opportunity sets of individual 1 and individual 2 under our mechanism are equal to $\bar{X}$. Therefore, if Rule 4 applies at some θ under σ^*, then we must have $g(\sigma^*(\theta)) \in C_i^\theta(\bar{X})$ for both $i \in \{1, 2\}$. But this violates (i) of choice incompatibility at θ.

Therefore, under any EPE σ^* of μ, Rule 1 must apply at every $\theta \in \Theta$.

Now, let σ^* be an arbitrary EPE of mechanism μ represented by $\sigma^*(\theta_i) = (n_i(\theta_i), f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$. Since Rule 1 applies at every $\theta \in \Theta$ under σ^*, there must exist a unique $\bar{f} \in F$ such that $f_1(\theta_1) = f_2(\theta_2) = \bar{f}$ for every $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$. To see why, suppose that for an arbitrary $\theta = (\theta_1, \theta_2) \in \Theta$, as Rule 1 must apply at θ under σ^*, $f_1(\theta_1) = f_2(\theta_2) = \bar{f}$ but there also exists $i_0 \in \{1, 2\}$ and $\theta_{i_0} \in \Theta_{i_0}$ such that $f_{i_0}(\theta_{i_0}) \neq \bar{f}$. Without loss of generality, suppose it is individual 1 type $\hat{\theta}_1 \in \Theta_1$ for whom we have $f_1(\hat{\theta}_1) \neq \bar{f}$. But, then, Rule 1 cannot apply at $(\hat{\theta}_1, \theta_2) \in \Theta$, as $f_1(\hat{\theta}_1) \neq \bar{f}$ and $f_2(\theta_2) = \bar{f}$, a contradiction to Rule 1 applying at all $\theta \in \Theta$ under the EPE σ^*.

Since there is a unique $\bar{f} \in F$ such that $f_i(\theta_i) = \bar{f}$ for each $\theta_i \in \Theta_i$ and $i \in \{1, 2\}$, by Rule 1, $g(\sigma^*(\theta)) = \bar{f}(\alpha(\theta))$ for each $\theta \in \Theta$. That is, $g \circ \sigma^* = \bar{f} \circ \alpha$.

Furthermore, it must be that $\bar{f} \circ \alpha \in F$ where α is the deception profile specified by the EPE σ^*. To see why, observe that at any $\theta \in \Theta$, each individual $i \in \{1, 2\}$ reports his/her type as $\alpha_i(\theta_i) \in \Theta_i$ as part of their messages under σ^*. Since Rule 1 applies at θ, by construction, we have $O_i^\mu(\sigma_j^*(\alpha_j(\theta_j))) = S_i(\bar{f}, \alpha_j(\theta_j))$ for each $i \in \{1, 2\}$. If $\bar{f} \circ \alpha \notin F$, then by (iv) of two-individual consistency, there exists $\theta^* \in \Theta$, such that either $\bar{f}(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(\bar{f}, \alpha_2(\theta_2^*)))$ or $\bar{f}(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(\bar{f}, \alpha_1(\theta_1^*)))$. Since Rule 1 applies at θ^* as well, we have $g(\sigma^*(\theta^*)) = \bar{f}(\alpha(\theta^*))$. Therefore, $\bar{f}(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(\bar{f}, \alpha_2(\theta_2^*)))$ implies $g(\sigma^*(\theta^*)) \notin C_1^{\theta^*}(O_1^\mu(\sigma_2^*(\theta_2^*)))$ while $\bar{f}(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(\bar{f}, \alpha_1(\theta_1^*)))$ implies that $g(\sigma^*(\theta^*)) \notin C_2^{\theta^*}(O_2^\mu(\sigma_1^*(\theta_1^*)))$. Both induce a contradiction to σ^* being an EPE of μ. Hence, we must have $\bar{f} \circ \alpha \in F$. Therefore, for any EPE σ^* of μ, there exists $f \equiv \bar{f} \circ \alpha \in F$ with $g \circ \sigma^* = f$, i.e., (ii) of ex-post implementation holds as well. $\square$

B.8 Proof of Theorem 6

Consider the mechanism $\mu = (M, g)$ constructed in section A.2.

As shown in the proof of Theorem 5, for any $f \in F$, $\sigma_i^f(\theta_i) = (0, f, \theta_i, x_i, 0)$ for both $i \in \{1, 2\}$ (with arbitrary $x_i \in \bar{X}$) is an EPE of μ such that $f = g \circ \sigma^f$. That is, for any $f \in F$, there exists an EPE, σ^f, of μ such that $f = g \circ \sigma^f$, which implies that condition (i) of Definition 2 (ex-post implementability) holds.

Consider an EPE σ^* of μ represented by $\sigma^*(\theta_i) = (n_i(\theta_i), f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$. For any $f \in F$ and $i \in N$, let $\bar{\Theta}_i^f := \{\theta_i \in \Theta_i | n_i(\theta) = 0, f_i(\theta_i) = f\}$ for $i \in \{1, 2\}$. That is, $\bar{\Theta}_i^f \subset \bar{\Theta}_i$ is the set of types of individual $i \in \{1, 2\}$ where the first entry of his/her message—his/her proposed SCF—is f under σ^*. Let $\bar{\Theta}_f := \times_{i \in N} \bar{\Theta}_i^f$. That is, $\bar{\Theta}_f$ is the set of states where both individuals propose the SCF $f \in F$ under σ^*. Consider the collection of product sets $\{\bar{\Theta}_f\}_{f \in F}$. Observe that $\bar{\Theta} := \bigcup_{f \in F} \bar{\Theta}_f$ is the set of states where Rule 1 applies under σ^*.

Hence, at any $\theta \in \Theta \setminus \bar{\Theta}$, either one of Rule 2.1, Rule 2.2, Rule 3 or Rule 4 applies under σ^*. For each of these rules, consider the corresponding opportunity sets under μ:

If Rule 2.1 or Rule 2.2 applies at θ, then the opportunity set of the individual i, whose message's first entry is 1, equals $S_i(f, \alpha_j(\theta_j))$, where $\alpha_j(\theta_j)$ denotes the reported type of j at θ, while the opportunity set of the individual j, whose message's first entry is 0, is $\bar{X}$, $i, j = 1, 2$ and $i \neq j$. Thus, if Rule 2.1 or Rule 2.2 applies at θ under σ^*, we must have $g(\sigma^*(\theta)) \in C_i^\theta(T)$ for some $T = S_i(f, \alpha_j(\theta_j)) \in \mathbb{S}_i$ and $g(\sigma^*(\theta)) \in C_j^\theta(\bar{X})$. By (iv) of consistency-unanimity, there exists $f^* \in F$ with $T \in \mathbb{S}_i$ such that $g(\sigma^*(\theta)) = f^*(\theta)$ whenever Rule 2.1 or Rule 2.2 applies at θ.

If Rule 3 applies at θ, then, as was shown in the proof of Theorem 5, the opportunity sets of individuals i and j are $S_i(f, \alpha_j(\theta_j))$ and $S_j(f', \alpha_i(\theta_i))$, where $f' \in F$ and $f \in F$ with $f \neq f'$ are reported SCFs and $\alpha_i(\theta_i)$ and $\alpha_j(\theta_j)$ are the reported types at θ in the messages of i and j, respectively, $i, j = 1, 2$ and $i \neq j$. Thus, if Rule 3 applies at some θ under σ^*, we must have $g(\sigma^*(\theta)) \in C_i^\theta(S_i(f, \alpha_j(\theta_j)))$ and $g(\sigma^*(\theta)) \in C_j^\theta(S_j(f', \alpha_i(\theta_i)))$. Thus, by (iv) of consistency-unanimity, there exists $f^* \in F$ such that $g(\sigma^*(\theta)) = f^*(\theta)$ when Rule 3 applies at θ.

Finally, whenever Rule 4 applies at θ, the opportunity sets of individuals 1 and 2 under our mechanism are both equal to $\bar{X}$. Therefore, if Rule 4 applies at $\theta \in \Theta \setminus \bar{\Theta}$ under σ^*, then $g(\sigma^*(\theta)) \in C_i^\theta(\bar{X})$ for both $i \in \{1, 2\}$. By (iv) of consistency-unanimity, there exists $f^* \in F$ such that $g(\sigma^*(\theta)) = f^*(\theta)$ whenever Rule 4 applies at θ as well.

To sum up, there exists $f^* \subset F$ such that $g(\sigma^*(\theta)) = f^*(\theta)$ for every $\theta \in \Theta \setminus \bar{\Theta}$.

Next, we show that it must also be that $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \bar{\Theta}$. Recall that $\sigma^*(\theta_i) = (n_i(\theta_i), f_i(\theta_i), \alpha_i(\theta_i), x_i(\theta_i), k_i(\theta_i))$, $i \in \{1, 2\}$: Suppose, for contradiction, that there exists $\tilde{\theta} \in \bar{\Theta}_f$ for some $f \in F$ such that $g(\sigma^*(\tilde{\theta})) \neq f^*(\tilde{\theta})$. Since Rule 1 applies at $\tilde{\theta}$ under σ^* and, hence, $g(\sigma^*(\tilde{\theta})) = f(\alpha(\tilde{\theta}))$, we have $f(\alpha(\tilde{\theta})) \neq f^*(\tilde{\theta})$ where α is the deception profile induced by σ^*. Therefore, by (v) of consistency-unanimity, there exists $\theta^* \in \bar{\Theta}_f$ such that either $f(\alpha(\theta^*)) \notin C_1^{\theta^*}(S_1(f, \alpha_2(\theta_2^*)))$ or $f(\alpha(\theta^*)) \notin C_2^{\theta^*}(S_2(f, \alpha_1(\theta_1^*)))$. Since $\theta^* \in \bar{\Theta}_f$, we have $f_i(\theta_i^*) = f$ for both $i \in \{1, 2\}$. That is, Rule 1 applies at θ^* and hence $g(\sigma^*(\theta^*)) = f(\alpha(\theta^*))$. But then, as shown in the proof of Theorem 5, $O_1^\mu(\sigma_2^*(\theta_2^*)) = S_1(f, \alpha_2(\theta_2^*))$ and $O_2^\mu(\sigma_1^*(\theta_1^*)) = S_2(f, \alpha_1(\theta_1^*))$. Therefore, $f(\alpha(\theta^*)) \notin C_1^{\theta^*}(O_1^\mu(\sigma_2^*(\theta_2^*)))$ implies $g(\sigma^*(\theta^*)) \notin C_1^{\theta^*}(O_1^\mu(\sigma_2^*(\theta_2^*)))$ while $f(\alpha(\theta^*)) \notin C_2^{\theta^*}(O_2^\mu(\sigma_1^*(\theta_1^*)))$ implies $g(\sigma^*(\theta^*)) \notin C_2^{\theta^*}(O_2^\mu(\sigma_1^*(\theta_1^*)))$. In both cases, σ^* cannot be an EPE of μ, a contradiction.

Therefore, $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \Theta$ with $f^* \in F$. That is, condition (ii) ex-post implementability (see Definition 2) holds as well. $\square$

B.9 Proof of Theorem 7

The proof is provided as a discussion right after Theorem 7.

C An Illustration of Sufficiency with Two Individuals

We show how one of our sufficiency results, Corollary 2, can be employed on an example that is inspired from Masatlioglu and Ok (2014).

The individual choices of Alice and Bob in this section are as specified in Table 9. The states of the world regarding the individual choices are $\Theta = \{(\Diamond,\Diamond),(\Diamond,c),(c,\Diamond),(c,c)\}$. That is, $\Theta_A = \Theta_B = \{\Diamond, c\}$, where type $\Diamond$ stands for *not having a status-quo* and type c stands for *status-quo being coal*. We consider a social planner who wants to ex-post implement the SCF f described in Table 10—a selection from the BR-optimal outcomes.

In what follows, we show that the collections $\mathbb{S}_A := \{S_A(f,\Diamond), S_A(f,c)\}$ and $\mathbb{S}_B := \{S_B(f,\Diamond), S_B(f,c)\}$, specified below, satisfy conditions (i), (ii), and (iii) of Corollary 2.

$$\begin{aligned} S_A(f,\Diamond) &= \{n,s\}, \quad & S_B(f,\Diamond) &= \{n,s\}, \\ S_A(f,c) &= \{c,n,s\}, \quad & S_B(f,c) &= \{c,n,s\}. \end{aligned}$$

Condition (i):

For $\theta_B = \Diamond$, we must have $f(\Diamond,\Diamond) \in C_A^{(\Diamond,\Diamond)}(S_A(f,\Diamond))$ and $f(c,\Diamond) \in C_A^{(c,\Diamond)}(S_A(f,\Diamond))$. Since $f(\Diamond,\Diamond) = s$, $f(c,\Diamond) = s$ and $s \in C_A^{(\Diamond,\Diamond)}(\{n,s\})$, $s \in C_A^{(c,\Diamond)}(\{n,s\})$, this is satisfied for $S_A(f,\Diamond) = \{n,s\}$.

For $\theta_B = c$, we must have $f(\Diamond,c) \in C_A^{(\Diamond,c)}(S_A(f,c))$ and $f(c,c) \in C_A^{(c,c)}(S_A(f,c))$. Since $f(\Diamond,c) = s$, $f(c,c) = n$ and $s \in C_A^{(\Diamond,c)}(\{c,n,s\})$, $n \in C_A^{(c,c)}(\{c,n,s\})$, this is satisfied for $S_A(f,c) = \{c,n,s\}$.

For $\theta_A = \Diamond$, we must have $f(\Diamond,\Diamond) \in C_B^{(\Diamond,\Diamond)}(S_B(f,\Diamond))$ and $f(\Diamond,c) \in C_B^{(\Diamond,c)}(S_B(f,\Diamond))$. Since $f(\Diamond,\Diamond) = s$, $f(\Diamond,c) = s$ and $s \in C_B^{(\Diamond,\Diamond)}(\{n,s\})$, $s \in C_B^{(\Diamond,c)}(\{n,s\})$, this is satisfied for $S_B(f,\Diamond) = \{n,s\}$.

For $\theta_A = c$, we must have $f(c,\Diamond) \in C_B^{(c,\Diamond)}(S_B(f,c))$ and $f(c,c) \in C_B^{(c,c)}(S_B(f,c))$. Since $f(c,\Diamond) = s$, $f(c,c) = n$ and $s \in C_B^{(c,\Diamond)}(\{c,n,s\})$, $n \in C_B^{(c,c)}(\{c,n,s\})$, this is satisfied for $S_B(f,c) = \{c,n,s\}$.

That is, $f(\theta'_A,\theta_B) \in C_A^{(\theta'_A,\theta_B)}(S_A(f,\theta_B))$ for each $\theta'_A \in \{\Diamond,c\}$ while $f(\theta_A,\theta'_B) \in C_B^{(\theta_A,\theta'_B)}(S_B(f,\theta_A))$ for each $\theta'_B \in \{\Diamond,c\}$, as desired.

Finally, since both n and s are in every set in the collections $\mathbb{S}_A$ and $\mathbb{S}_B$ we have $S_A(f,\theta_B) \cap S_B(f,\theta_A) \neq \emptyset$ for each $\theta_A,\theta_B \in \{\Diamond,c\}$ as well.

Therefore, condition (i) of Corollary 2 is satisfied by the collections $\mathbb{S}_A$ and $\mathbb{S}_B$.

Condition (ii):

For any product set $\bar{\Theta} \subseteq \Theta$, we have to consider the individual choices of Alice and Bob from $\bar{X}$ and $\bar{X}$; $\bar{X}$ and $S_B(f,\Diamond)$; $\bar{X}$ and $S_B(f,c)$; $S_A(f,\Diamond)$ and $\bar{X}$; $S_A(f,c)$ and $\bar{X}$; $S_A(f,\Diamond)$ and $S_B(f,\Diamond)$; $S_A(f,\Diamond)$ and $S_B(f,c)$; $S_A(f,c)$ and $S_B(f,\Diamond)$; $S_A(f,c)$ and $S_B(f,c)$ at every state of the world in $\Theta \setminus \bar{\Theta}$, i.e., outside of $\bar{\Theta}$.

Since $\bigcup_{S \in \mathbb{S}_A \cup\, \mathbb{S}_B} S \subseteq \bar{X}$, we must have $\bar{X} = \{c,n,s\}$. Furthermore, $S_A(f,c) = S_B(f,c) = \{c,n,s\} = \bar{X}$.

Therefore, for any product set $\bar{\Theta} \subseteq \Theta$, it is enough to check Alice's choices from

$\{c,n,s\}$ and Bob's choices from $\{c,n,s\}$; Alice's choices from $\{c,n,s\}$ and Bob's choices from $\{n,s\}$; Alice's choices from $\{n,s\}$ and Bob's choices from $\{c,n,s\}$ at every state of the world in $\Theta \setminus \bar{\Theta}$.

Below, we show that f satisfies choice unanimity whenever the individual choices from the aforementioned sets overlap at any state of the world. This means condition (ii) is satisfied for any subset of Θ, in particular, for any product set $\bar{\Theta} \subseteq \Theta$, as desired.

$\underline{\{c,n,s\} \text{ for Alice, } \{c,n,s\} \text{ for Bob:}}$ Alice's and Bob's choices overlap at $(\Diamond,\Diamond)$ and (c,c) with $s \in C_A^{(\Diamond,\Diamond)}(\{c,n,s\}) \cap C_B^{(\Diamond,\Diamond)}(\{c,n,s\})$ and $n \in C_A^{(c,c)}(\{c,n,s\}) \cap C_B^{(c,c)}(\{c,n,s\})$. Since $f(\Diamond,\Diamond) = s$ and $f(c,c) = n$, choice unanimity for these particular sets is satisfied at every state of the world.

$\underline{\{c,n,s\} \text{ for Alice, } \{n,s\} \text{ for Bob:}}$ Alice's and Bob's choices overlap at $(\Diamond,\Diamond)$ and $(\Diamond,c)$ and (c,c) with $s \in C_A^{(\Diamond,\Diamond)}(\{c,n,s\}) \cap C_B^{(\Diamond,\Diamond)}(\{n,s\})$ and $s \in C_A^{(\Diamond,c)}(\{c,n,s\}) \cap C_B^{(\Diamond,c)}(\{n,s\})$ and $n \in C_A^{(c,c)}(\{c,n,s\}) \cap C_B^{(c,c)}(\{n,s\})$. Since $f(\Diamond,\Diamond) = s$ and $f(\Diamond,c) = s$ and $f(c,c) = n$, choice unanimity for these particular sets is also satisfied at every state of the world.

$\underline{\{n,s\} \text{ for Alice, } \{c,n,s\} \text{ for Bob:}}$ Alice's and Bob's choices overlap at $(\Diamond,\Diamond)$ and $(c,\Diamond)$ and (c,c) with $s \in C_A^{(\Diamond,\Diamond)}(\{n,s\}) \cap C_B^{(\Diamond,\Diamond)}(\{c,n,s\})$ and $s \in C_A^{(c,\Diamond)}(\{n,s\}) \cap C_B^{(c,\Diamond)}(\{c,n,s\})$and $n \in C_A^{(c,c)}(\{n,s\}) \cap C_B^{(c,c)}(\{c,n,s\})$. Since $f(\Diamond,\Diamond) = s$ and $f(c,\Diamond) = s$ and $f(c,c) = n$, choice unanimity for these particular sets is satisfied at every state of the world as well.

Therefore, condition *(ii)* of Corollary 2 is also satisfied by the collections $\mathbb{S}_A$ and $\mathbb{S}_B$.

Condition *(iii)*:

For any $\bar{\Theta} \subset \Theta$, we show that if $f(\alpha(\theta)) \neq f(\theta)$ for some $\theta \in \bar{\Theta}$, then $\theta^* = \theta \in \bar{\Theta}$ works as the informant state by a case by case analysis:

If $f(\alpha(\theta)) \neq f(\theta)$, then (at least) one of the following must be true: (1) $\theta = (\Diamond,\Diamond)$ and hence $f(\alpha(\Diamond,\Diamond)) \neq f(\Diamond,\Diamond)$; (2) $\theta = (\Diamond,c)$ and hence $f(\alpha(\Diamond,c)) \neq f(\Diamond,c)$; (3) $\theta = (c,\Diamond)$ and hence $f(\alpha(c,\Diamond)) \neq f(c,\Diamond)$; or (4) $\theta = (c,c)$ and hence $f(\alpha(c,c)) \neq f(c,c)$.

$\underline{\text{Case 1: If } \theta = (\Diamond,\Diamond) \text{, i.e., } f(\alpha(\Diamond,\Diamond)) \neq f(\Diamond,\Diamond)\text{:}}$ Then, $f(\alpha(\Diamond,\Diamond)) = n$. Hence, we must have $\alpha_A(\Diamond) = \alpha_B(\Diamond) = c$. Then, $\theta^* = \theta = (\Diamond,\Diamond)$ and $i^* = A$ work since $S_A(f,\alpha(\theta_B)) = S_A(f,c) = \{c,n,s\}$ and $n \notin C_A^{(\Diamond,\Diamond)}(\{c,n,s\})$.

$\underline{\text{Case 2: If } \theta = (\Diamond,c) \text{, i.e., } f(\alpha(\Diamond,c)) \neq f(\Diamond,c)\text{:}}$ Then, $f(\alpha(\Diamond,c)) = n$. Hence, we must have $\alpha_A(\Diamond) = c$ and $\alpha_B(c) = c$. Then, $\theta^* = \theta = (\Diamond,c)$ and $i^* = A$ work since $S_A(f,\alpha(\theta_B)) = S_A(f,c) = \{c,n,s\}$ and $n \notin C_A^{(\Diamond,c)}(\{c,n,s\})$.

$\underline{\text{Case 3: If } \theta = (c,\Diamond) \text{, i.e., } f(\alpha(c,\Diamond)) \neq f(c,\Diamond)\text{:}}$ Then, $f(\alpha(c,\Diamond)) = n$. Hence, we must have $\alpha_A(c) = c$ and $\alpha_B(\Diamond) = c$. Then, $\theta^* = \theta = (c,\Diamond)$ and $i^* = B$ work since

$S_B(f, \alpha(\theta_A)) = S_B(f, c) = \{c, n, s\}$ and $n \notin C_B^{(c,\Diamond)}(\{c, n, s\})$.

Case 4: If $\theta = (c, c)$, i.e., $f(\alpha(c, c)) \neq f(c, c)$: Then, $f(\alpha(c, c)) = s$. Hence, we must have either $\alpha_A(c) = \Diamond$ or $\alpha_B(c) = \Diamond$, or both. We consider each of these three cases separately:

Subcase 4.1: If $\alpha_A(c) = \Diamond$ and $\alpha_B(c) = c$: Then, $\theta^* = \theta = (c, c)$ and $i^* = A$ work since $S_A(f, \alpha(\theta_B)) = S_A(f, c) = \{c, n, s\}$ and $s \notin C_A^{(c,c)}(\{c, n, s\})$.

Subcase 4.2: If $\alpha_A(c) = c$ and $\alpha_B(c) = \Diamond$: Then, $\theta^* = \theta = (c, c)$ and $i^* = B$ work since $S_B(f, \alpha(\theta_A)) = S_B(f, c) = \{c, n, s\}$ and $s \notin C_B^{(c,c)}(\{c, n, s\})$.

Subcase 4.3: If $\alpha_A(c) = \Diamond$ and $\alpha_B(c) = \Diamond$: Then, $\theta^* = \theta = (c, c)$ and $i^* = A$ work since $S_A(f, \alpha(\theta_B)) = S_A(f, \Diamond) = \{n, s\}$ and $s \notin C_A^{(c,c)}(\{n, s\})$.

Therefore, $\mathbb{S}_A$ and $\mathbb{S}_B$ satisfy condition (iii) of Corollary 2 as well.

References

Afif, Z. (2017). *"Nudge units" – where they came from and what they can do.* Retrieved 2019-03-27, from http://blogs.worldbank.org/developmenttalk/nudge-units-where-they-came-and-what-they-can-do

Ariely, D. (2009). *Predictably irrational: The hidden forces that shape our decisions.* Harper Perennial.

Barlo, M., & Dalkiran, N. A. (2009). Epsilon-Nash implementation. *Economics Letters*, *102*(1), 36–38.

Bergemann, D., & Morris, S. (2005). Robust mechanism design. *Econometrica*, *73*(6), 1771–1813.

Bergemann, D., & Morris, S. (2008). Ex post implementation. *Games and Economic Behavior*, *63*(2), 527–566.

Bergemann, D., & Morris, S. (2009). Robust implementation in direct mechanisms. *The Review of Economic Studies*, *76*(4), 1175–1204.

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

Borgers, T., & Li, J. (2018). Strategically simple mechanisms. *Unpublished*.

Cabrales, A., & Serrano, R. (2011). Implementation in adaptive better-response dynamics: Towards a general theory of bounded rationality in mechanisms. *Games and Economic Behavior*, *73*(2), 360–374.

Danilov, V. (1992). Implementation via Nash equilibria. *Econometrica*, 43–56.

Dean, M., Kıbrıs, Ö., & Masatlioglu, Y. (2017). Limited attention and status quo bias. *Journal of Economic Theory*, *169*, 93–127.

de Clippel, G. (2014). Behavioral implementation. *American Economic Review*, *104*(10), 2975–3002.

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

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

Glazer, J., & Rubinstein, A. (2012). A model of persuasion with boundedly rational agents. *Journal of Political Economy*, *120*(6), 1057–1082.

Herne, K. (1997). Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, *13*(3), 575–589.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98.

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*.

Jackson, M. O. (1991). Bayesian implementation. *Econometrica*, 461–477.

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

Jehiel, P., Meyer-ter Vehn, M., & Moldovanu, B. (2008). Ex-post implementation and preference aggregation via potentials. *Economic Theory*, *37*(3), 469–490.

Jehiel, P., Meyer-ter Vehn, M., Moldovanu, B., & Zame, W. R. (2006). The limits of ex post implementation. *Econometrica*, *74*(3), 585–610.

Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.

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

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

Korpela, V. (2012). Implementation without rationality assumptions. *Theory and Decision*, *72*(2), 189–203.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, *107*(11), 3257–87.

Masatlioglu, Y., & Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, *121*(1), 1–29.

Masatlioglu, Y., & Ok, E. A. (2014). A Canonical Model of Choice with Initial Endowments. *The Review of Economic Studies*, *81*(2), 851-883.

Maskin, E. (1985). The theory of implementation in Nash equilibrium: A survey. *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, 173–204.

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

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

Moore, J. (1992). Implementation, contracts, and renegotiation in environments with complete information. *Advances in economic theory*, *1*, 182–281.

Moore, J., & Repullo, R. (1990). Nash implementation: a full characterization. *Econo-*

metrica, 1083–1099.

Ohashi, Y. (2012). Two-person ex post implementation. *Games and Economic Behavior*, *75*(1), 435–440.

Ok, E. A., Ortoleva, P., & Riella, G. (2015). Revealed (p)reference theory. *American Economic Review*, *105*(1), 299–321.

Palfrey, T. R. (2002). Implementation theory. *Handbook of game theory with economic applications*, *3*, 2271–2326.

Palfrey, T. R., & Srivastava, S. (1987). On bayesian implementable allocations. *The Review of Economic Studies*, *54*(2), 193–208.

Postlewaite, A., & Schmeidler, D. (1986). Implementation in differential information economies. *Journal of Economic Theory*, *39*(1), 14–33.

Ray, K. T. (2018). Nash implementation under choice functions. *Unpublished*.

Repullo, R. (1987). A simple proof of Maskin's theorem on Nash implementation. *Social Choice and Welfare*, *4*(1), 39–41.

Rubinstein, A., & Salant, Y. (2011). Eliciting welfare preferences from behavioural data sets. *The Review of Economic Studies*, *79*(1), 375–387.

Saijo, T. (1988). Strategy space reduction in Maskin's theorem: sufficient conditions for Nash implementation. *Econometrica*, 693–700.

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

Saran, R. (2011). Menu-dependent preferences and revelation principle. *Journal of Economic Theory*, *146*(4), 1712–1720.

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

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

Spiegler, R. (2011). *Bounded rationality and industrial organization.* Oxford University Press.

Sugden, R. (2004). The opportunity criterion: consumer sovereignty without the assumption of coherent preferences. *American Economic Review*, *94*(4), 1014–1033.

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, *76*(1), 31–48.