Behavioral Implementation under Incomplete Information*

Mehmet Barlo† Nuh Aygün Dalkıran‡

October 15, 2020

Abstract

We investigate mechanism design under incomplete information allowing for individuals to display different behavioral biases in different states of the world. Our primitives are individual choices, which do not have to satisfy the weak axiom of revealed preferences. In this setting, we provide necessary as well as sufficient conditions for behavioral (ex-post) implementation under incomplete information.

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

JEL Classification: D60, D79, D82, D90

*We would like to thank Fuad Aleskerov, Ahmet Alkan, Dirk Bergemann, Eddie Dekel, Mehmet Ekmekci, Jeff Ely, Ehud Kalai, Semih Koray, Stephen Morris, Wojciech Olszewski, Alessandro Pavan, Luciano Pomatto, Yuval Salant, Marciano Siniscalchi, Asher Wolinsky, Kemal Yıldız, and the seminar participants at Northwestern University, The North American Summer Meeting of the Econometric Society 2019, The European Meeting of the Econometric Society 2019, Bosphorus Workshop on Economic Design 2019, Social Choice and Welfare 2018 for helpful comments. Muhammed Dönmez and Ece Yegane provided excellent research assistance. The authors gratefully acknowledge the financial support by TÜBİTAK under grant no. 119K957. 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

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 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 weak axiom of revealed preferences (WARP), the condition corresponding to the standard axioms of rationality. We allow individuals' choices to be interdependent while independence is not ruled out. Our results provide useful insights into behavioral mechanism design as information asymmetries are inescapable in many economic settings.

In particular, we analyze mechanism design under incomplete information when individuals may have different behavioral biases in different states of the world, such as falling for an attraction effect, displaying a status-quo bias, or revealing cyclic preferences, among others. In doing so, we focus on full implementation employing ex-post equilibrium (EPE) and provide necessary as well as sufficient conditions. Therefore, our paper can be viewed as the *incomplete information* counterpart of de Clippel (2014), which is one of the pioneering papers on behavioral implementation under *complete information*.

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

[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 book *Nudge* has been influential in 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 the EU, OECD, UN agencies, and the World Bank integrating 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 (see Footnote 10). Besides *Nudge* (Thaler & Sunstein, 2008), two other 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).

appeal of partial implementation involves *the revelation principle*, which implies, in the rational domain and under incomplete information, the following: if there is a mechanism that partially implements a predetermined goal, then there is a *direct 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, elegantly pointed out by Postlewaite and Schmeidler (1986) among others.[4]

We show that the revelation principle (for partial implementation) fails when individuals' choices do not satisfy WARP.[5] Hence, in our environment, one cannot restrict attention to direct mechanisms without a loss of generality. Thus, focusing on full implementation rather than partial implementation becomes crucial in our setup.

The EPE is a strategy profile such that each individual's plan of action is measurable only with respect to her type and it results in Nash equilibrium play at every state of the world. It is well-suited to our environment with choices, not necessarily derived from preference maximization, for the following reasons: It makes no use of any probabilistic information, it is belief-free, it does not involve any belief updating or expectation considerations, and it does not require any common prior assumption. It induces robust behavior on account of the ex-post no-regret property: "no agent would like to change his message even if he were to know the true type profile of the remaining agents" (Bergemann & Morris, 2008). Further, it provides a plausible extension of dominant equilibrium to the case of interdependence: With independent choices, we show that the EPE is equivalent to (behavioral) dominant equilibrium under some full-range conditions (that hold in direct mechanisms), while dominant equilibrium with interdependent choices imposes excessively stringent requirements reducing its appeal.

The Bayesian Nash equilibrium seems impractical in our setup because it employs an aggregation of individuals' welfare in different states using associated probabilities. At the very least, this necessitates the need for complete and transitive preferences over the set of certain outcomes to obtain a utility representation. However, this is neither coherent nor consistent in a setting in which individuals' choices over certain outcomes do not satisfy WARP. In our environment, individuals' choices over certain outcomes may not even be representable by well-defined preference relations (see Footnote 13).

In order to highlight the new grounds our results cover relative to the complete in-

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

[4] "[The] problem [of multiple equilibria] 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" (Postlewaite & Schmeidler, 1986).

[5] To the best of our knowledge, the failure of the revelation principle in behavioral environments is first documented by Saran (2011).

formation analysis of de Clippel (2014), we emphasize that behavioral (full) ex-post implementation is not the same as behavioral (Nash) implementation on every complete information type space: Each individual's strategy is measurable with respect to her type and cannot vary with others' type profiles. This, therefore, results in a requirement akin to the addition of incentive compatibility constraints. Indeed, the associated necessary conditions are not nested even in the rational domain (Bergemann & Morris, 2008).[6]

Our necessity result, Theorem 1, shows that if a mechanism ex-post implements a social choice set (SCS), then the opportunity sets sustained in EPE of this mechanism, alternatives that an individual can obtain by changing her messages while her opponents' remain the same, form a collection of sets with two desirable properties. We refer to such family of choice sets as the collection of sets *consistent with the given SCS under incomplete information*. Each member of this collection is associated with an individual and a social choice function (SCF) in the SCS and a type profile of the other individuals with the following property novel to the case of incomplete information—causing a significant difference with its complete information counterpart, consistency of de Clippel (2014): Each such choice set is independent of the type of the individual whom this set is associated with. Moreover, the following hold: (i) For all individuals, all her types, all SCFs in the SCS, and all type profiles for the other individuals, this individual's choices at the resulting type profile (state) from the corresponding choice set contain the alternative that corresponds to the outcome of the SCF for that type profile; (ii) whenever there is a deception that leads to an outcome that is incompatible with the SCS, there is an informant state (i.e., a type profile) and an informant individual such that she does not choose at the informant state the alternative generated by this deception from her set associated with others' types identified via their deception from the informant state. We show that the first of these, (i), implies a pseudo ex-post incentive compatibility (Proposition 2), while the second, (ii), implies an ex-post choice monotonicity condition (Proposition 1). Another implication (Proposition 3) is that the revelation principle holds if individuals' choices satisfy the independence of irrelevant alternatives (IIA). Furthermore, we establish that under rationality our ex-post choice incentive compatibility is equivalent to the ex-post incentive compatibility of Bergemann and Morris (2008), and our ex-post choice monotonicity implies their ex-post monotonicity while their ex-post monotonicity coupled with their ex-post incentive compatibility implies our ex-post choice monotonicity.

We provide two methods that strengthen consistency to deliver sufficiency: The first,

[6]See Bergemann and Morris (2008, Propositions 3 and 4) and notice that, under WARP, our necessary conditions are equivalent to theirs. Besides, Bergemann and Morris (2005) shows that even though (partial) ex-post implementation is equivalent to interim (Bayesian) implementation for all possible type spaces in some environments, this equivalence does not hold in the case of full implementation.

Theorem 2, involves a mild condition that requires some level of disagreement in the society (*choice incompatibility*). Theorem 3 presents the second method that uses a combination of consistency and a choice counterpart of no-veto (*consistency-no-veto*).

To showcase the applicability of our results in economically relevant domains, we analyze the behavioral ex-post implementation of efficiency and allocation problems with endowment effects. *First*, we introduce an ex-post analog of efficiency of de Clippel (2014) for incomplete information settings and establish that the failure of incentive compatibility prevents its ex-post implementability (echoing the well-known incompatibility of efficiency and strategy-proofness). That is why, we consider the constrained efficient SCS obtained by additionally imposing our incentive compatibility constraints, and show that it is ex-post implementable on all domains containing weak levels of disagreement. *Second*, we formalize implementation of allocation problems with endowment effects à la Masatlioglu and Ok (2014) and independent choices and describe how to design individuals' opportunity sets for the planner to have their status-quo biases induce choices aligned with the desired goal. This demonstrates the use of initial endowments as critical tools in the design of the mechanism, an observation that cannot be obtained with rationality.

In the implementation literature, positive sufficiency results often rely on "augmented" mechanisms asking individuals to report more than their types. Such mechanisms seem less intuitive and less practical than direct mechanisms to many researchers. That is why we analyze the scope of situations where ex-post implementation is achievable via direct mechanisms: Theorem 4 identifies necessary and sufficient conditions for the social goal to be ex-post implementable via the associated direct mechanism. Moreover, restricting attention to independent choices, we reexamine consistency and obtain choice counterpart of strategy-proofness as a necessary condition.[7] We also elaborate on the relation between (behavioral) dominant equilibrium and EPE identifying a full-range condition that ensures the equivalence of these two equilibrium notions.

When dealing with individuals having cognitive limitations, simplicity becomes a serious concern. That is why, besides direct mechanisms, we consider the size of the joint message space, the number of message profiles, of a mechanism as a measure of its simplicity. This measure is similar in spirit with the total size of message spaces used to analyze communication complexity in Segal (2007, 2010). In Theorem 5, we identify lower bounds with respect to this measure for mechanisms that ex-post implement a given SCS. This

[7] Interdependence plays a critical role in full ex-post implementation in the rational domain and brings about a "stark contrast" to the case with independent choices: The generalized VCG allocation is ex-post implementable with interdependence but not with independence (Bergemann & Morris, 2008, Section 7.3). Our analysis encompasses the rational domain and hence this observation as well. Further, none of our results demands interdependence (e.g., the failure of the revelation principle) while independence provides convenience and paves the way to (behavioral) dominant strategy implementation.

provides a better understanding of the scope of the well-known criticism based on complexity of mechanisms used in the literature. Moreover, we corroborate that behavioral aspects induce less simple mechanisms by comparing the rational case with the situation where only one individual suffers severely from a status-quo bias in an allocation problem with endowment effects and independent choices.

To further demonstrate the practicality and applicability of our results in tangible (finite) environments, we also provide Python codes associated with our necessity and sufficiency theorems: Using the SCS and individuals' state-contingent choices as inputs, these codes deliver consistent collections as well as those satisfying our sufficiency conditions—choice incompatibility and consistency-no-veto, reminiscent of the economic environment assumption and monotonicity-no-veto, respectively (Jackson, 1991). Therefore, these induce better understanding and application capabilities by mitigating the effects of complications observed in implementation under incomplete information that arise due to conditions such as monotonicity-no-veto and ex-post-monotonicity-no-veto.

In the implementation literature, it is well-known that the two-individual case involves additional complications. Nevertheless, we extend our results to the case of two individuals, while we defer its presentation to Appendix B due to expositional purposes.

Our paper is mostly related to de Clippel (2014), which provides necessary as well as sufficient conditions for behavioral implementation under complete information. Besides de Clippel (2014), another closely related paper is Bergemann and Morris (2008), analyzing ex-post implementation in the rational domain under incomplete information.[8] 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 WARP. Another related paper is Jackson (1991), which analyzes Bayesian implementation for the case of three or more individuals in the rational domain. It generalizes the analysis of Maskin (1999) (on Nash implementation under complete information) to 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).

Another significant and related paper is Saran (2011), which considers behavioral partial implementation under incomplete information formalizing behavioral aspects with

[8] Ohashi (2012) provides sufficiency results for ex-post implementation with two rational individuals in an environment that is economic and has a bad outcome. Our sufficiency results with two individuals differ from 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 WARP. Some of the other influential and related work on ex-post implementation and robust mechanism design in the rational domain include Bergemann and Morris (2005, 2009, 2011), Jehiel, Meyer-ter Vehn, Moldovanu, and Zame (2006), Jehiel, Meyer-ter Vehn, and Moldovanu (2008).

menu-dependent preferences over interim Anscombe-Aumann acts. It establishes that weak contraction consistency, a condition implied by the IIA, is sufficient for the revelation principle. In Appendix D, we discuss how our setup compares with that of Saran.

A recent paper, de Clippel (2020), issues a warning for the use of EPE and (behavioral) dominant equilibrium in environments with probabilistically sophisticated individuals having singleton-valued choices over alternatives: The failure of the IIA may be at odds with the plausibility of the EPE and (behavioral) dominant equilibrium. That is why, in environments with individuals acting in lieu of groups of rational and probabilistically sophisticated agents, the planner has to take this warning seriously when designing a mechanism. Our setup, on the other hand, does not involve any probabilities and hence is a good fit for environments where probabilistic sophistication comes at a cost (possibly due to the failure of the IIA). Besides, adopting the analog of the sure-thing principle induced by the EPE and the dominant equilibrium as an axiom implies that the failure of the IIA is not compatible with probabilistic sophistication, reaffirming the significance of the IIA in regard to bounded rationality. We deliberate over these in Appendix E.

Hurwicz (1986), Eliaz (2002), Korpela (2012), and Ray (2018) have also investigated the problem of behavioral implementation under complete information. Hurwicz (1986) considers choices that can be represented by a well-defined preference relation that does not have to be acyclic. Eliaz (2002), a seminal paper containing pioneering research on behavioral implementation, provides an analysis of full implementation when some of the individuals might be "faulty" and hence fail to act optimally. Then, the mechanism has to deal with the complications that emerge due to each individual "optimally respond[ing] to the non-faulty players regardless of the identity and actions of the faulty players." On the other hand, Korpela (2012) shows that when individual choices fail rationality axioms, the IIA, also known as Sen's α, is key to obtaining the necessary and sufficient condition synonymous to that of Moore and Repullo (1990).[9]

The organization of the paper is as follows: Section 2 presents a motivating example. In Section 3, we provide the notation and the definitions. Section 4 contains necessity results; Section 5, sufficiency results with at least three or more individuals. In Section 6, we analyze the behavioral ex-post implementation of efficiency, while Section 7 contains

[9]There have been other papers investigating implementation under complete information that allow for "non-rational" behavior of individuals. 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. Some of the other related work include Benoit and Ok (2006), Cabrales and Serrano (2011), Kucuksenel (2012), Saran (2016), and Bochet and Tumennasan (2018). For more on full implementation, we refer the reader to surveys such as Moore (1992), Jackson (2001), Maskin and Sjöström (2002), Palfrey (2002), and Serrano (2004).

an application to the allocation problems with endowment effects. Section 8 presents our results regarding when behavioral ex-post implementation is possible via direct mechanisms and implications of independent choices to behavioral ex-post implementation. In Section 9, we discuss simple mechanisms and Section 10 concludes. Meanwhile, the proofs are presented in Appendix A. Our analysis with *two individuals* is relegated to Appendix B. Moreover, Appendix C contains a comparison with Bergemann and Morris (2008); Appendix D, a comparison of our setup with that of Saran (2011); Appendix E, a comparison with de Clippel (2020); Appendix F, a comparison of our simplicity notion with the communication complexity of Segal (2007, 2010); Appendix G, the description of the canonical model of choice with initial endowments of Masatlioglu and Ok (2014); finally, Appendix H, outlines of our Python codes.

2 Motivating Example

The following example aims to display the intricacies concerning the design of a mechanism which implements a behavioral welfare notion, the strict generalized Pareto optimality due to Bernheim and Rangel (2009), in EPE with two individuals whose choices do not satisfy WARP.[10] These choices involve three types of behavioral biases: (1) attraction effect, (2) status-quo bias, and (3) Condorcet cycles. Indeed, our example demonstrates that behavioral implementation under incomplete information can be achieved with different behavioral biases in different states of the world.

Two individuals, *Ann* and *Bob*, are to decide what type of energy to employ or jointly invest in, be it *coal* energy, *nuclear* energy, or *solar* energy. Thus, the grand set of alternatives is $X = \{coal, nuclear, solar\}$.[11] Let the set of all relevant states of the world regarding the individuals' choices be given by Θ. Suppose that Ann and Bob have two possible types each, denoted by $\Theta_i = \{\rho_i, \gamma_i\}$ for $i \in \{A, B\}$. 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 Ann 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

[10] Sen (1971) shows that a choice correspondence satisfies 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). Letting $\mathcal{X}$ be the set of all non-empty subsets of alternatives, we say that the individual choice correspondence $C : \mathcal{X} \to \mathcal{X}$ satisfies (i) Sen's α if whenever $x \in S \subset T$ for some $S, T \in \mathcal{X}$, $x \in C(T)$ implies $x \in C(S)$; (ii) 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 $y \in C(T)$.

[11] Ann and Bob can also be interpreted as regions A and B within the same legislation, such as two states in the U.S. or two countries in the E.U. 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."

specific choices where c stands for *coal*, n for *nuclear* power, and s for *solar* energy.

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 Ann and Bob.

At state (ρ_A, ρ_B), Ann'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 Ann and Bob at (ρ_A, γ_B) can be explained by the *attraction effect*, one of the commonly observed behavioral biases.[12,13] 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 Ann's choices feature a *status-quo bias* where the status-quo is c.[14] 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 Ann and Bob violate the

[12]Decoy alternatives, alternatives that are known to be dominated by other alternatives, can cause preference reversals when they are introduced into the choice set. Herne (1997) demonstrates how the presence of a decoy alternative causes the attraction effect in a policy-making context: 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, the supporters of nuclear energy used coal as a decoy alternative. Motivated by this, in our example, at (ρ_A, γ_B), Ann and Bob choose n from $\{c, n, s\}$ and s from $\{n, s\}$. They also choose n from $\{c, n\}$. This means either of Ann and Bob chooses n whenever it is presented with c, the decoy option, even though s is chosen from $\{n, s\}$. Thus, their choices cannot be rationalized by a complete and transitive preference relation, as they violate the IIA. For more on attraction effect, see Huber, Payne, and Puto (1982), de Clippel and Eliaz (2012), and Ok, Ortoleva, and Riella (2015).

[13]There is no *well-defined preference relation* representing Ann and Bob's choices at (ρ_A, γ_B): 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 10, 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.

[14]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. At (γ_A, ρ_B), Ann suffers from status-quo bias and her choices do not obey WARP (in particular, Sen's β); they cannot be rationalized by a complete and transitive preference relation. For more on status-quo bias, see Samuelson and Zeckhauser (1988), Kahneman, Knetsch, and Thaler (1991), Masatlioglu and Ok (2005), and Dean, Kıbrıs, and Masatlioglu (2017).

IIA and Sen's β. Furthermore, Ann's choices lead to a Condorcet cycle.[15]

Our model allows individuals' choices to be *interdependent*: between (ρ_A, ρ_B) and (ρ_A, γ_B), Ann's private information (type) does not change; yet, the choice behavior of Ann is not identical at these two states. That is, even though Ann does not know Bob's private information (type), she knows the set of all possible types of Bob. Therefore, Ann 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 Ann's choices as in the case of a common value auction.

Our social choice notion is based on the welfare criterion developed by Bernheim and Rangel (2009) delivering the notion of *BR-optimality*.[16] The BR-optimal alternatives

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.

contingent on the states are as summarized in Table 2.

As in Palfrey and Srivastava (1987), an SCS refers to a selection of state-contingent allocations. In our example, the planner aims to implement the SCS, $F = \{f, f'\}$, described in Table 3, consisting of state-contingent allocations each of which is BR-optimal

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 for Ann and Bob.

at every state.[17] We note that F is mutually exhaustive of BR-optimal outcomes as $\{f(\theta)\} \cup \{f'(\theta)\}$ equals the set of BR-optimal outcomes at every $\theta \in \Theta$.[18]

[15] Ann chooses n from $\{c, n\}$, c from $\{c, s\}$, and s from $\{n, s\}$. Such a pattern may arise when Ann makes her choices by consulting *a group of rational individuals*, such as pairwise voting with her parents, or a parliamentary vote. Hurwicz (1986) investigates the problem of implementation in such cases.

[16] An alternative x is *strictly unambiguously chosen over* another alternative z, if z is never chosen whenever x is available; 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 extend the notion of Pareto efficiency beyond the rational domain as follows. 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. Another paper that provides a welfare analysis that is in line with non-rational choices is Rubinstein and Salant (2011).

[17] BR-optimal alternatives are defined under certainty. It is not clear how to extend this notion of efficiency to the case of uncertainty as the individual choices of Ann 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.

[18] 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 depends on the given SCS.

A mechanism μ makes Ann and Bob send individual *messages* (in M_A and M_B, respectively) to the social planner and describes the outcome to be implemented as a function of these messages (given by the *outcome function* $g : M \to X$ with $M = M_A \times M_B$). Given a mechanism $\mu = (M, g)$, de Clippel (2014) points out an intuitive and straightforward extension of the notion of Nash Equilibrium (NE) involving individuals' choices that cannot be rationalized by a complete and transitive preference relation: For each individual, the equilibrium outcome should be among the chosen within the set of alternatives he/she can generate by unilateral deviations. Formally, $m^* = (m_A^*, m_B^*)$ is an NE of the mechanism $\mu = (M, g)$ at θ if $g(m^*) \in C_A^\theta(O_A^\mu(m_B^*))$ and $g(m^*) \in C_B^\theta(O_B^\mu(m_A^*))$ where the opportunity sets of Ann and Bob are given by $O_A^\mu(m_B) := \{g(m_A, m_B) \mid m_A \in M_A\}$ and $O_B^\mu(m_A) := \{g(m_A, m_B) \mid m_B \in M_B\}$, respectively. If m^* is an NE of μ at θ, we refer to $g(m^*)$ as an NE outcome of μ at θ.

In our setting, the state of the world is distributed knowledge between Ann and Bob, and they observe only 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. That is, a strategy for Ann and Bob in a mechanism $\mu = (M, g)$ is a function $\sigma_i : \Theta_i \to M_i$ for $i \in \{A, B\}$. There is no clear way of defining a Bayesian Nash Equilibrium of a mechanism in our example as Ann and Bob cannot be modeled as (expected) utility maximizers. Due to the same reason, we restrict our attention to pure strategies. 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 Ann 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.

2.1 The revelation principle fails

The revelation principle (for partial implementation) fails in our example: Consider the SCF f given in Table 3. Later in the section, we show that the mechanism μ in Table 5 possesses an EPE sustaining f. Hence, the mechanism μ *partially* ex-post implements f. However, the corresponding direct mechanism, $g^d : \Theta \to X$, given in Table 4, fails to partially ex-post implement f truthfully as truthful revelation is not an EPE of g^d: At

		Bob	
		ρ_B	γ_B
Ann	ρ_A	n	n
	γ_A	n	s

Table 4: The direct mechanism g^d.

state (ρ_A, γ_B), reporting truthfully delivers n (circled in Table 4), but Ann's opportunity set at state (ρ_A, γ_B) is $\{n, s\}$ and $n \notin C_A^{(\rho_A,\gamma_B)}(\{n, s\}) = \{s\}$.[19]

This conclusion does not depend on the interdependence of choices: In Section 7, we present an example of an allocation problem with endowment effects where the revelation principle fails with independent choices.

To the best of our knowledge, the failure of the revelation principle on the behavioral implementation front is first documented by Saran (2011).[20] That study establishes that *weak contraction consistency*, a condition implied by the IIA, is sufficient for the revelation principle. In Section 4, we reaffirm that the revelation principle holds under the IIA in our setup as well. This justifies our search for indirect mechanisms, even for partial implementation, when individuals' choices do not satisfy the IIA. In this context, our results identifying (indirect) mechanisms for full implementation are also useful as full implementation implies partial implementation.

2.2 An indirect mechanism that ex-post implements F

We now show that the following indirect mechanism $\mu = (M, g)$ fully ex-post implements the SCS $F = \{f, f'\}$, described in Table 3: $M_A = \{U, M, D\}$ and $M_B = \{L, M, R\}$; $g : M \to X$ is described in Table 5.

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

Table 5: The mechanism μ for Ann and Bob.

To exemplify how to identify NE of this mechanism, below we identify Bob's best responses at (γ_A, γ_B): If Ann sends the message U, Bob can unilaterally generate the set $\{c, n\}$ under the mechanism μ, i.e., $O_B^\mu(U) = \{c, n\}$. 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. Similarly, when Ann sends the message M, Bob can unilaterally generate the set $\{c, s\}$ under the mechanism μ, i.e., $O_B^\mu(M) = \{c, s\}$, 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 Ann's action M. Finally, if Ann

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

[20] Saran (2011) considers a setup with menu-dependent preferences over interim Anscombe-Aumann acts. In Appendix D, we discuss how our setting can be captured by menu-dependent preferences over interim (deterministic) Anscombe-Aumann acts. Meanwhile, Bierbrauer and Netzer (2016) notes that the revelation principle fails with intention-based social preferences.

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 $\{n, s\}$ at (γ_A, γ_B); hence, both M and R are the best responses for Bob.

Repeating this exercise, one can show that 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).

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.

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 Ann and Bob, our mechanism would be (fully) implementing the BR-optimal outcomes in NE.[21]

The true state of the world is not common knowledge between Ann and Bob under incomplete information. Then, employing EPE demands individuals' strategies (which are measurable with respect to individuals' private information) induce an NE of the mechanism at every state of the world. In the following, we show that there are three EPE of our mechanism, two of which are 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 $\theta \in \Theta$.*

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

Table 7 summarizes the EPE outcomes of μ where message profiles corresponding to

[21]We would like to emphasize that μ Nash implements the BR-optimal outcomes under complete information and ex-post implements F under incomplete information. In general, a mechanism that ex-post implements an SCS F does not have to Nash implement the social choice correspondence associated with F. For example, the mechanism presented in Table B.3 of Appendix B ex-post implements the SCS F of our motivating example under incomplete information but does not Nash implement the BR-optimal outcomes (the social choice correspondence associated with F) under complete information.

σ'^* are depicted with circles while those associated with σ''^* are indicated with squares and those corresponding to σ'''^* with diamonds 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 $g(\sigma''^*(\theta)) = g(\sigma'''^*(\theta)) = f'(\theta)$ for each $\theta \in \Theta$. That is, (i) each SCF in the SCS is induced by a particular EPE of the mechanism μ; and (ii) for each EPE of the mechanism μ, there is a particular SCF in the SCS that induces the same outcomes state by state. Thus, μ fully ex-post implements the SCS F. In Appendix B.3, we show that μ is the "simplest mechanism" ex-post implementing F.

3 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 $\langle N, X, \Theta, (C_i^\theta)_{i\in N,\theta\in\Theta}\rangle$ 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.

An SCF is $f : \Theta \to X$ that specifies a socially optimal alternative—as evaluated by the planner—for each state, i.e., f is a state-contingent allocation. For any $\theta_{-i} \in \Theta_{-i}$, we let $f(\Theta_i, \theta_{-i}) := \{f(\theta_i', \theta_{-i}) | \theta_i' \in \Theta_i\}$. As there may be many socially optimal state-contingent allocations that a designer may wish to consider simultaneously, we focus on SCSs rather than SCFs. An SCS, denoted by F, is a non-empty set of SCFs, i.e., $F \subset \{f \mid f : \Theta \to X\}$ and $F \neq \emptyset$.[22] $\mathcal{F}$ denotes the set of all 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.

The opportunity set of an individual under a mechanism is the set of alternatives that she can generate by unilateral deviations given the messages of the other individuals: The *opportunity set* of individual i under μ for each $m_{-i} \in M_{-i}$ is given by $O_i^\mu(m_{-i}) = \{g(m_i, m_{-i}) \in X \mid m_i \in M_i\}$. Consequently, an NE of a mechanism at a particular state of the world is defined as follows: A message profile m^* is a *Nash equilibrium* of μ at θ if $g(m^*) \in C_i^\theta(O_i^\mu(m_{-i}^*))$ for all $i \in N$.

The mechanism μ in our environment induces an incomplete information game-form. A strategy of individual i under the mechanism μ, a contingent plan of actions, specifies a message for each possible type of i, and is denoted by $\sigma_i : \Theta_i \to M_i$. Due to the aforementioned reasons, we restrict attention to pure EPE.

Definition 1. *A strategy profile $\sigma^* : \Theta \to M$ is an* **ex-post equilibrium** *of μ 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 an 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 μ such that:*

(i) For every $f \in F$, there exists an EPE σ^* of μ that satisfies $f = g \circ \sigma^*$, and

(ii) For every EPE σ^* of μ, there exists $f \in F$ such that $g \circ \sigma^* = f$.

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 for $F = \{f\}$.

Any mechanism that ex-post implements an SCS should take into consideration the private information of the individuals. However, individuals may misreport their private

[22] 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).

information. We denote a *deception* by 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, $\alpha(\theta) := (\alpha_1(\theta_1), \alpha_2(\theta_2), \dots, \alpha_n(\theta_n))$ is a profile of reported types, which might be deceptive.

4 Necessity

We show that the notion of *consistency under incomplete information* is necessary for ex-post implementation. When the meaning is clear, we refer to it simply as *consistency*.

Definition 3. *We say that a collection of sets* $\mathbb{S} := \{S_i(f, \theta_{-i}) \mid i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\} \subset \mathcal{X}$ *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: (*i*) 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 she is of type θ_i' at state (θ_i', θ_{-i}) contains $f(\theta_i', \theta_{-i})$ for all $\theta_i' \in \Theta_i$; (*ii*) 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 exist an informant state θ^* and an informant individual i^* such that i^* does not choose at state θ^* the alternative $f(\alpha(\theta^*))$ (generated by this deception) from $S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*}))$.[23]

If a mechanism $\mu = (M, g)$ ex-post implements a given SCS $F \in \mathcal{F}$, then for any SCF $f \in F$, there exists an EPE σ^f of μ such that $f = g \circ \sigma^f$. Thus, for each $\theta \in \Theta$, $g(\sigma^f(\theta)) = f(\theta)$ is in $C_i^\theta(O_i^\mu(\sigma^f_{-i}(\theta_{-i})))$ for all $i \in N$. Defining $\mathbb{S}$ by $S_i(f, \theta_{-i}) := O_i^\mu(\sigma^f_{-i}(\theta_{-i}))$ with $i \in N$, $f \in F$, and $\theta_{-i} \in \Theta_{-i}$ (i.e., the collection sustained by the opportunity sets associated with the EPE of μ), we observe that (*i*) of consistency of $\mathbb{S}$ with F holds.

On the other hand, if a deception profile α is such that $f \circ \alpha \notin F$, then $\sigma^f \circ \alpha$ cannot be an EPE of μ. Otherwise, by (*ii*) of ex-post implementability (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. So, there is a state θ^* and an individual i^* who does not choose at θ^* (going along with

[23] Consistency of de Clippel (2014), a necessary condition for behavioral implementation under complete information, requires that, given a social choice correspondence $\Phi : \Theta \to \mathcal{X}$, the collection consistent with Φ is $\{S_i(x, \theta) \in \mathcal{X} \mid i \in N, \theta \in \Theta, x \in \Phi(\theta)\}$, such that (*i*) for all $i \in N$, all $\theta \in \Theta$, and all $x \in \Phi(\theta)$, $x \in C_i^\theta(S_i(x, \theta))$; (*ii*) $x \in \Phi(\theta) \setminus \Phi(\theta')$ with $\theta, \theta' \in \Theta$ implies there is $i^* \in N$ such that $x \notin C_{i^*}^{\theta'}(S_{i^*}(x, \theta))$. The critical difference between de Clippel's consistency and ours is that, with incomplete information, each choice set must be independent of the type of the individual whom this set is associated with.

the deception and obtaining) the alternative $f(\alpha(\theta^*))$ from $O^{\mu}_{i^*}(\sigma^f_{-i^*}(\alpha_{-i^*}(\theta_{-i^*})))$ which equals $S_{i^*}(f, \alpha_{-i^*}(\theta_{-i^*}))$. This delivers (ii) of consistency of $\mathbb{S}$ with F.

This discussion proves that the existence of a collection consistent with an SCS under incomplete information is a necessary condition for this SCS to be ex-post implementable:

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

Theorem 1 affirms the following intuition: if the designer cannot identify sets from which individuals make choices compatible with the social goal, then she cannot succeed in the corresponding implementation attempt.

In Supplementary Materials, we provide Python codes computing consistent collections given individuals' choices and the SCS as inputs (see Appendix H.1.1 for an overview). We also display how to compute consistent collections using these codes on a variety of examples, one of which is our motivating example.

Next, to establish that our study extends the analysis of Bergemann and Morris (2008) to the irrational domain, we show that our necessary condition implies *analogs* of theirs: ex-post choice monotonicity and quasi-ex-post choice incentive compatibility. Under WARP, these conditions are equivalent to ex-post monotonicity and ex-post incentive compatibility of Bergemann and Morris (2008).

Proposition 1. *If there exists a collection of sets consistent with an SCS $F \in \mathcal{F}$ under incomplete information, then F is* **ex-post choice monotonic***; i.e., 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 set of alternatives $S^* \in \mathcal{X}$ such that*

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

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

Our 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.

Proposition 2. *If there exists a collection of sets consistent with an SCS $F \in \mathcal{F}$ under incomplete information, then F is* **quasi-ex-post choice incentive compatible***; i.e., for every SCF $f \in F$ and state $\theta \in \Theta$ and individual $i \in N$, there exists a set of alternatives $S \in \mathcal{X}$ such that $f(\theta) \in C^{\theta}_i(S)$ and $f(\Theta_i, \theta_{-i}) \subseteq S$.*

Quasi-ex-post choice incentive compatibility of an SCS F demands that for all $f \in F$, all state $\theta \in \Theta$, and all individuals $i \in N$, there is a set S from which i chooses $f(\theta)$ at θ while S contains $f(\Theta_i, \theta_{-i})$, all the alternatives achievable by i under f given θ_{-i}.

In Appendix C, we show that under WARP (i) quasi-ex-post choice incentive compatibility is equivalent ex-post incentive compatibility; (ii) ex-post choice monotonicity implies ex-post monotonicity while ex-post monotonicity coupled with ex-post incentive compatibility implies ex-post choice monotonicity. Consequently, the necessary conditions of Bergemann and Morris (2008) are equivalent to our ex-post choice monotonicity coupled with quasi-ex-post choice incentive compatibility under WARP.

If $F = \{f\}$, quasi-ex-post choice incentive compatibility also describes a necessary condition for partial ex-post implementation of the SCF f. As we show in Section 2.1, the revelation principle (for partial ex-post implementation) does not hold in our setup. In general, when the containment relation in quasi-ex-post choice incentive compatibility holds strictly, truthtelling may not be an EPE of the associated direct mechanism. On the other hand, the following is a *necessary and sufficient condition* for the revelation principle: An SCF f is partially truthfully (ex-post) implementable in a direct mechanism if and only if for every $\theta \in \Theta$, $i \in N$, $f(\theta) \in C_i^\theta(f(\Theta_i, \theta_{-i}))$.[24] This condition neither implies nor is implied by the quasi-ex-post choice incentive compatibility. Yet, under the IIA, we obtain the following:

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

In summary, if a mechanism μ partially ex-post implements an SCF f and individuals' choices satisfy the IIA, then there is a direct mechanism g^d which partially implements f in truthful EPE. That is, the IIA is sufficient for the revelation principle.

5 Sufficiency

Ex-post implementation of an SCS F is not feasible when there is no collection of sets consistent with F under incomplete information. 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.[25]

Definition 4. *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 $x \notin C_j^\theta(S)$.*

[24]If f is partially truthfully (ex-post) implemented by the direct mechanism $g^d : \Theta \to X$, the opportunity set of i under truthtelling given θ_{-i} is $O_i^{g^d}(\theta_{-i}) = f(\Theta_i, \theta_{-i})$.

[25]There is room for other sufficient conditions since we do not restrict choices using universal axioms. But, it seems neither easy nor practical to close the gap between the necessary and sufficient conditions.

This condition implies some level of disagreement among individuals regarding their choices from a given set of alternatives at a given state. In words, a set satisfies the choice incompatible pair property at a state, if for each alternative in this set, there is 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 this state.

The choice incompatible pair property coupled with consistency is 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 $\mathbb{S} := \{S_i(f, \theta_{-i}) \mid i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ consistent with F under incomplete information, and

(ii) a set of alternatives $\bar{X} \subseteq X$ with $\bigcup_{S \in \mathbb{S}} S \subseteq \bar{X}$ that 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: If (i) there exists a collection of sets $\mathbb{S}$ consistent with an 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, then F is ex-post implementable. To check whether or not the hypothesis of Theorem 2 holds, we present Python codes that take individuals' choices and the SCS as inputs to compute consistent collections and sets of alternatives $\bar{X}$ satisfying the choice incompatible pair property. Appendix H.1.2 contains an overview of these codes, and we display how to use them on two three-individual examples, Examples SM-4 and SM-5. These demonstrate that the choice incompatible pair property is a mild sufficiency condition: For all pairs of consistent collections and resulting $\bar{X}$'s that satisfy the hypothesis of Theorem 2, $\bar{X}$ is a strict subset of X, whereas individuals' choices on X are *aligned* (i.e., at every state, an alternative is chosen by at least $n-1$ individuals from X). Therefore, the choice incompatible pair property imposes a weaker sufficiency requirement than the one imposed by the *choice counterpart of the economic environment assumption*:[26] The latter assumption is not satisfied in these examples but the hypothesis of Theorem 2 holds, since choice incompatible pair property concerns only $\bar{X}$ obtained from a consistent collection of sets. Indeed, Example SM-5 with rational individuals establishes that, the choice incompatible pair property under rationality does not imply the economic environment assumption.

[26]Individuals' choices satisfy the *choice counterpart of the economic environment assumption* if for all $\theta \in \Theta$ and for all $x \in X$, there are $i, j \in N$ such that $x \notin C_i^{\theta}(X)$ and $x \notin C_j^{\theta}(X)$. Under WARP, this requirement coincides with the economic environment assumption of Bergemann and Morris (2008).

Hence, under rationality, our Theorem 2 (coupled with our necessity results, Theorem 1, Propositions 1, 2, C.1, and C.2) extends Bergemann and Morris (2008)'s sufficiency result that uses the economic environment assumption (their Theorem 2).

Theorem 2 identifies conditions that make sure that all EPE of the mechanism used in its proof (described in Appendix A.3) 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 Rules 2 and 3 as well. To do so, we turn to the counterpart of the no-veto power property in our environment.

Definition 5. *We say that an SCF f satisfies the* **choice no-veto-power property** *on a 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 this state, then this alternative must be f*-optimal* at this state.

Our second sufficiency result employs a combination of consistency and the choice no-veto-power property. Below, we present this condition followed by the result.[27]

Definition 6. *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.*

In words, an SCS F satisfies the *consistency-no-veto* property if there are a collection of sets $\mathbb{S}$ and a set of alternatives $\bar{X}$ containing every alternative in $\mathbb{S}$ such that:

[27] 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 from $S_i(f, \theta_{-i})$ at (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 $f(\alpha(\theta^*))$ (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 the same arguments presented in the proof of Proposition 2, quasi-ex-post choice incentive compatibility follows from (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 hypotheses of Theorem 3 simplify 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 $\mathbb{S} := \{S_i(\theta_{-i}) : i \in N, \theta_{-i} \in \Theta_{-i}\}$ such that for all individuals $i \in N$, $f(\theta_i', \theta_{-i}) \in C_i^{(\theta_i', \theta_{-i})}(S_i(\theta_{-i}))$ 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 $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(S_{i^*}(\alpha_{-i^*}(\theta^*_{-i^*})))$.

Our Python codes related to consistency-no-veto (see Appendix H.1.3 for an overview) help us gain better understanding of implementation under incomplete information when the canonical mechanism has equilibria under rules other than Rule 1. By shedding light on consistency-no-veto (and hence ex-post-monotonicity-no-veto), these codes alleviate associated complications and supply further application possibilities. Taking individuals' choices and the SCS as inputs, our codes compute collections of sets, $\mathbb{S}$'s, the sets, $\bar{X}$'s, and associated f^*'s for each product sets of states, $\bar{\Theta}$'s, satisfying consistency-no-veto. For a three individual example, see Example SM-4 in Supplementary Materials.

6 Efficiency

de Clippel (2014) introduces the following notion of efficiency: $\Phi^{eff}(\theta) := \{x \in X \mid \exists (Y_i)_{i\in N} \text{ with } x \in C_i^\theta(Y_i) \text{ for all } i \in N \text{ and } X = \cup_{i\in N} Y_i\}$. In words, an alternative x is efficient at θ if each individual has an implicit opportunity set such that she chooses x from this set and each alternative is in at least one of the implicit opportunity sets of an individual. $\Phi^{eff} : \Theta \to \mathcal{X}$ is a social choice correspondence (SCC) that we refer to as the de Clippel efficient SCC. de Clippel (2014) discusses the relation of this concept with the opportunity criterion of Sugden (2004).

In what follows, we extend this notion to the case of incomplete information. Let $F^{eff} := \{f : \Theta \to X \mid f(\theta) \in \Phi^{eff}(\theta) \text{ for all } \theta \in \Theta\}$. In words, F^{eff} is the set of all SCFs that are selections from the de Clippel efficient SCC.

Proposition 4. *F^{eff} is* not *ex-post implementable.*

To prove this result, we provide an example where F^{eff} does not have a consistent collection and hence cannot be ex-post implemented:

Let $X = \{x, y\}$, $N = \{A, B\}$, and $\Theta_i = \{\theta_i, \omega_i\}$ with $i = A, B$. The individuals' choices are as given in Table 8. Given these choices, it is straightforward to see that the de Clippel

S	$C_A^{(\theta_A,\theta_B)}$	$C_B^{(\theta_A,\theta_B)}$	$C_A^{(\theta_A,\omega_B)}$	$C_B^{(\theta_A,\omega_B)}$	$C_A^{(\omega_A,\theta_B)}$	$C_B^{(\omega_A,\theta_B)}$	$C_A^{(\omega_A,\omega_B)}$	$C_B^{(\omega_A,\omega_B)}$
$\{x,y\}$	$\{x\}$	$\{x\}$	$\{x\}$	$\{x\}$	$\{y\}$	$\{x\}$	$\{y\}$	$\{x\}$

Table 8: Individual choices of A and B among the allocations x and y.

efficient SCC, Φ^{eff}, is: $\Phi^{eff}(\theta_A, \theta_B) = \Phi^{eff}(\theta_A, \omega_B) = \{x\}; \Phi^{eff}(\omega_A, \theta_B) = \Phi^{eff}(\omega_A, \omega_B) = \{x, y\}$. So, $F^{eff} = \{f, f', f'', f'''\}$ is as in Table 9. If $\mathbb{S}_A$, $\mathbb{S}_B$ were consistent with F^{eff} under incomplete information, then by (i) of consistency, $x \in C_B^{(\omega_A,\theta_B)}(S_B(f', \omega_A))$ and

	(θ_A, θ_B)	(θ_A, ω_B)	(ω_A, θ_B)	(ω_A, ω_B)
f	x	x	x	x
f'	x	x	x	y
f''	x	x	y	x
f'''	x	x	y	y

Table 9: The SCS F^{eff}.

$y \in C_B^{(\omega_A,\omega_B)}(S_B(f',\omega_A))$. Therefore, $S_B(f',\omega_A) = \{x,y\}$. However, $C_B^{(\omega_A,\omega_B)}(\{x,y\}) = \{x\}$, and hence, $y \notin C_B^{(\omega_A,\omega_B)}(S_B(f',\omega_A))$, a contradiction.

The reason behind F^{eff} not being ex-post implementable is due to the lack of quasi-ex-post choice incentive compatibility: Given $f' \in F^{eff}$ and state (ω_A,ω_B), $f'(\omega_A,\theta_B) = x$ and $f'(\omega_A,\omega_B) = y$ implies that the only candidate set $S' \in \mathcal{X}$ that satisfies the conditions in quasi-ex-post choice incentive compatibility is $\{x,y\}$. However, Bob chooses x, and not $y = f'(\omega_A,\omega_B)$, from this set at (ω_A,ω_B).

In the rational domain, ex-post incentive compatibility with independent types implies (weak) dominant strategy incentive compatibility that parallels strategy-proofness. Given the results regarding the incompatibility between efficiency and strategy-proofness in the mechanism/market design literature, our result pointing out the incompatibility between F^{eff} and quasi-ex-post choice incentive compatibility is not very surprising. This leads us to the following notion of ***constrained efficiency***:

$$E^{c.eff} \equiv \left\{ e : \Theta \to X \;\middle|\; \begin{array}{ll} (i) & \forall i \in N, \forall \theta_{-i} \in \Theta_{-i},\ \exists\, Y_i^{\theta_{-i}} \text{ with } \forall \theta \in \Theta,\ \cup_{i\in N} Y_i^{\theta_{-i}} = X, \text{ and} \\ (ii) & e(\tilde{\theta}_i,\theta_{-i}) \in C_i^{(\tilde{\theta}_i,\theta_{-i})}(Y_i^{\theta_{-i}}), \forall \tilde{\theta}_i \in \Theta_i. \end{array} \right\} \tag{1}$$

This notion internalizes ex-post incentive compatibility concerns into efficiency: A state-contingent allocation e is constrained efficient if for any individual and for any type profile of the others', there exists an implicit opportunity set such that her choices from this set for each of her types are aligned with e with the additional property that at every state each alternative is in at least one of the implicit opportunity sets of an individual. Therefore, constrained efficient SCS, $E^{c.eff}$, consists of all the constrained efficient state-contingent allocations and can be considered as an ex-post counterpart of de Clippel efficiency entangled with quasi-ex-post choice incentive compatibility.[28] Below, we show that $E^{c.eff}$ satisfies our necessity condition:

Proposition 5. *$E^{c.eff}$ has a consistent collection of sets under incomplete information.*

Theorem 2 and Proposition 5 deliver the following result presented without a proof: $E^{c.eff}$ is ex-post implementable when there is a weak form of disagreement in the society:

Proposition 6. *Let $n \geq 3$. $E^{c.eff}$ is ex-post implementable on all domains with X satisfying the choice incompatible pair property at every state $\theta \in \Theta$.*

[28]In the example used to prove Proposition 4, $E^{c.eff} = \{f, f'''\}$.

7 Allocation Problems

In this section, we investigate the implications of our results in an allocation problem under the assumption that the individual choices are independent. Then, we consider the canonical model of choice with initial endowments of Masatlioglu and Ok (2014) to obtain further insights on *allocation problems with endowment effects*.

There are n objects to be allocated among n individuals: Let $H = \{h_1, \ldots, h_n\}$ be a set of *objects* (e.g., houses or offices) while $\mathcal{H}$ denotes the set of all non-empty subsets of H. The set of allocations we focus on are those in which each individual gets only one object, i.e., $X := \{x \in H^n \mid x_i \neq x_j, \text{ for all } i, j \in N \text{ with } i \neq j\}$.[29]

There are two different types of individual choice behavior one needs to consider in this setup: *choices on objects* and *choices on allocations of objects*. Individuals' choices are independent and each individual cares only about her own object allocation—as is standard in canonical assignment/matching models without externalities: For any $i \in N$, $Z \in \mathcal{H}$, and $\theta_i \in \Theta_i$, $c_i(Z, \theta_i) \subset Z$ with $c_i(Z, \theta_i) \neq \emptyset$ constitutes the chosen object(s) from Z by i when her type is θ_i. For any set of allocations $S \in \mathcal{X}$, $H_i(S)$ is the set of objects that is assigned to i in allocations in S; i.e., $H_i(S) := \{x_i \in H \mid x \in S\}$. The choices of $i \in N$ on allocations at $\theta_i \in \Theta_i$ is as follows: for any $S \in \mathcal{X}$, $C_i^{\theta_i}(S) := \{x \in S \mid x_i \in c_i(H_i(S), \theta_i)\}$. So, the individual choices on allocations are independent.

In this setting, *constrained efficiency* has a natural appeal. It takes the following specific form: $\tilde{E}_H^{\text{c.eff}}$ consists of $\tilde{e} : N \times \Theta \to H$ with $\tilde{e}_i(\theta) \neq \tilde{e}_j(\theta)$ for all $i \neq j$ and all $\theta \in \Theta$ such that (i) for all $i \in N$ and all $\theta_{-i} \in \Theta_{-i}$, there is $H_i(\theta_{-i}) \in \mathcal{H}$ with $\tilde{e}_i(\tilde{\theta}_i, \theta_{-i}) \in c_i(H_i(\theta_{-i}), \tilde{\theta}_i)$ for all $\tilde{\theta}_i \in \Theta_i$; and (ii) for all $\theta \in \Theta$, $\cup_{i \in N} H_i(\theta_{-i}) = H$.[30]

We say that the *choice incompatibility on the set of objects* H holds whenever for all $\theta \in \Theta$ and all $h \in H$, there are $i, j \in N$ with $i \neq j$ such that $h \notin c_i(H, \theta_i)$ and $h \notin c_j(H, \theta_j)$. That is, for all states $\theta \in \Theta$ and all objects $h \in H$, there are two individuals not choosing h from H at θ. Indeed, this condition implies that the choice incompatible pair property holds on the set of all allocations X. Hence, by Proposition 6, we obtain the following:

Corollary 2. *Let $n \geq 3$. $\tilde{E}_H^{c.eff}$ is ex-post implementable on all domains satisfying the choice incompatibility on the set of all objects H.*

In what follows, we use the canonical model of choice with initial endowments of Masatlioglu and Ok (2014) to highlight insights about ex-post implementation in an allocation problem with endowment effects. We consider a setting in which individuals'

[29]This setup parallels the housing market analyzed in Shapley and Scarf (1974) and de Clippel (2014).

[30]$E^{\text{c.eff}}$ given in (1) is equivalent to $\tilde{E}_H^{\text{c.eff}}$: To see $E^{\text{c.eff}}$ implies $\tilde{E}_H^{\text{c.eff}}$, for all $i \in N$ and all $\theta_{-i} \in \Theta_{-i}$, set $H_i(\theta_{-i}) = H_i(Y_i^{\theta_{-i}})$; to see $\tilde{E}_H^{\text{c.eff}}$ implies $E^{\text{c.eff}}$, let $Y_i^{\theta_{-i}} = \{x \in X : x_i = h \text{ for some } h \in H_i(\theta_{-i})\}$.

choices on objects are given by (singleton-valued) choice functions as in Masatlioglu, Nakajima, and Ozbay (2012).

The initial endowment profile, $h^* = (h_i^*)_{i\in N} \in X$, is common knowledge among the individuals and the planner. We note that $H = \cup_{i\in N}\{h_i^*\}$. Whether or not the initial endowment of individual i affects her behavior is privately known only by herself. When she is a *rational* type, her behavior does not feature any endowment effects; we denote this case by $\theta_i = \lozenge_i$. But when she is a *behavioral* type and hence $\theta_i = h_i^*$, then her choices may be affected by her initial endowment. Masatlioglu and Ok (2014) shows that under some reasonable assumptions the choices of individual i can be represented by a *utility function* $U_i : H \to \mathbb{R}$ and a *consideration set* $Q_i(h_i^*) := \{\tilde{h} \in H : \tilde{h} = c_i(\{\tilde{h}, h_i^*\}, h_i^*)\}$ such that i's choices when she is a rational type are $c_i(S, \lozenge_i) = \arg\max_{h\in S} U_i(h)$, while if she is a behavioral type her choices are $c_i(S, h_i^*) = \arg\max_{h\in Q_i(h_i^*)\cap S} U_i(h)$. The consideration set of i consists of all the houses that are chosen against the initial endowment h_i^* in a binary comparison. Meanwhile, $LCS_i(x_i) := \{h \in H : U_i(x_i) \geq U_i(h)\}$.

Masatlioglu and Ok (2014, Theorem 1) implies that a behavioral type's choices when her initial endowment is offered is as if she is maximizing her utility subject to her consideration set, a (psychological) constraint induced by her initial endowment. In all other cases, she is a standard utility maximizer. Appendix G provides further details.

This construction implies that individual i's choices on allocations (obtained from her choices on objects) are as follows: For all $\theta \in \Theta$ and all $S \in \mathcal{X}$,

$$C_i^{\theta_i}(S) := \begin{cases} \{x \in S \mid x_i = c_i(H_i(S), h_i^*)\} & \text{if } \theta_i = h_i^* \text{ and } h_i^* \in H_i(S), \\ \{x \in S \mid x_i = c_i(H_i(S), \lozenge_i)\} & \text{if } \theta_i = h_i^* \text{ and } h_i^* \notin H_i(S), \\ \{x \in S \mid x_i = c_i(H_i(S), \lozenge_i)\} & \text{if } \theta_i = \lozenge_i. \end{cases}$$

We note that if a mechanism does not provide individual i the option to keep her initial endowment, h_i^*, then she makes her choices as if she is rational.

The following refines our necessary conditions implied in this setup:

Proposition 7. *Let U_i be individual i's utility function and $Q_i(h_i^*)$ her consideration set representing her choice behavior over the set of objects H as in Masatlioglu and Ok (2014, Theorem 1). If an SCS $F \in \mathcal{F}$ is ex-post implementable, then there exists a collection of sets $\mathbb{S} := \{S_i(f, \theta_{-i}) \mid i \in N, f \in F, \theta_{-i} \in \Theta_{-i}\}$ consistent with F under incomplete information such that*

(i) if $f(\theta) = x$, then $H_i(S_i(f, \theta_{-i})) \subset LCS_i(x_i)$,

(ii) if $f(\theta) = x$ and $x_i \notin Q_i(h_i^*)$, then $h_i^* \notin H_i(S_i(f, \theta_{-i}))$,

(iii) if $f(\theta) = x$ and $x_i = h_i^*$, then $H_i(S_i(f, \theta_{-i})) \cap Q_i(h_i^*) = \{h_i^*\}$,

(iv) if $f(\theta) = x$ and $x_i \neq h_i^*$, then either $h_i^* \notin H_i(S_i(f, \theta_{-i}))$ or $x_i \in Q_i(h_i^*)$.

In words, Proposition 7 says that an SCS F being ex-post implementable (and hence possessing a consistent collection of sets) implies that the following hold: (i) if an SCF f in SCS F equals allocation x at θ, then the objects offered to i at θ_{-i} associated with f, $H_i(S_i(f, \theta_{-i}))$, is a subset of $LCS_i(x_i)$. (ii) if f equals x at θ and x_i is not in i's consideration set, $Q_i(h_i^*)$, then the set of objects offered to i at θ_{-i} associated with f cannot include i's initial endowment, h_i^*. (iii) if f equals x at θ such that i is allocated her initial endowment h_i^*, then the set of objects offered to i at θ_{-i} associated with f cannot include any object in the consideration set $Q_i(h_i^*)$ other than h_i^*. Finally, (iv) says that if f equals x at θ while $x_i \neq h_i^*$, then either x_i is in i's consideration set $Q_i(h_i^*)$ or her initial endowment h_i^* is not among the objects offered to i at θ_{-i} associated with f.

Below, we present an allocation problem with endowment effects that showcases that offering initial endowments as options in individuals' opportunity sets helps the designer with ex-post implementation.

Consider three individuals whose choices can be described by the canonical model of choice with initial endowments: $N = \{1,2,3\}$, $H = \{I, II, III\}$, $\Theta_i = \{\Diamond_i, h_i^*\}$ where $h_1^* = II$, $h_2^* = I$, and $h_3^* = III$. The corresponding consideration sets are given by $Q_1(h_1^*) = \{I, II\}$, $Q_2(h_2^*) = \{I, II, III\}$, and $Q_3(h_3^*) = \{III\}$ while the corresponding utilities are as given in Table 10. When individual 1 is a behavioral type, she suffers

Object	U_1	U_2	U_3
I	2	2	3
II	1	3	2
III	3	1	1

Table 10: Individuals utilities U_i, $i = 1,2,3$, of objects.

severely from status-quo bias as $c_1(\{I, II, III\}, h_1^*) = \{I\}$ but $c_1(\{I, III\}, h_1^*) = \{III\}$. Meanwhile, the choices of individual 3 of type h_3^* satisfies WARP even though she suffers from an "extreme status-quo bias" as in Masatlioglu and Ok (2014, Example 2).

The planner wishes to ex-post implement the SCF $f : \Theta \to X$ such that $f(\theta) = (I, II, III)$ if $\theta = h^*$ and $f(\theta) = (III, II, I)$ for any other $\theta \neq h^*$, where $h^* = (h_1^*, h_2^*, h_3^*)$. That is, if all three individuals are of behavioral types, then the planner desires each individual to get the object that maximizes her utility function subject to her consideration set. Otherwise, the planner demands each individual i to get the object that maximizes her utility function without any constraints. We note that $f \in \tilde{E}_H^{\text{c.eff}}$.[31]

It follows from Proposition 7 that $\{I, III\} \subset H_1(S_1(h_{-1}^*))$ and $\{I, III\} \subset H_3(S_3(h_{-3}^*))$ because $f_1(h^*) = I$, $f_1(\Diamond_1, h_{-1}^*) = III$, $f_3(h^*) = III$, $f_3(\Diamond_3, h_{-3}^*) = I$. As a result, consis-

[31] To see this, set $H_1(h_{-1}^*) = \{I, II, III\}$, $H_2(h_{-2}^*) = \{II\}$, and $H_3(h_{-3}^*) = \{III\}$; $H_1(\theta_{-1}) = \{III\}$ for all $\theta_{-1} \neq h_{-1}^*$, $H_2(\theta_{-2}) = \{II\}$ for all $\theta_{-2} \neq h_{-2}^*$, and $H_3(\theta_{-3}) = \{I\}$ for all $\theta_{-3} \neq h_{-3}^*$.

tency demands that $H_1(S_1(h^*_{-1}))$ equals $\{I, II, III\}$ because if it were to equal $\{I, III\}$, $c_1(\{I, III\}, h^*_1) = c_1(\{I, III\}, \Diamond_1) = \{III\}$ implies an impasse due to $f_1(h^*) = I$. Therefore, given that both of the others are of behavioral types, individual 1's initial endowment, object II, must be offered to her as an option to make her choose as the planner wishes, and ex-post implementation of f is impossible if II is not in her opportunity set. On the other hand, $H_3(S_3(h^*_{-3}))$ equals either $\{I, III\}$ or $\{I, II, III\}$. For any other $\theta_{-1} \neq h^*_{-1}$, $f_1(\Diamond_1, \theta_{-1}) = f_1(h^*_1, \theta_{-1}) = III$. Thus, $H_1(S_1(\theta_{-1}))$ is either $\{III\}$ or $\{I, III\}$. So, if one of the others is rational, then 1's endowment cannot be offered to her. For any other $\theta_{-3} \neq h^*_{-3}$, $H_3(S_3(\theta_{-3}))$ could be either $\{I, II\}$ or $\{I\}$ as $f_3(\Diamond_3, \theta_{-3}) = f_3(h^*_3, \theta_{-3}) = I$.

One can show that the collection of sets $\mathbb{S}$ (of allocations) that induces the following are consistent with the SCF f under incomplete information: $H_1(S_1(h^*_{-1})) = \{I, II, III\}$ and $H_1(S_1(\theta_{-1})) = \{III\}$ for all $\theta_{-1} \neq h^*_{-1}$; $H_2(S_2(\theta_{-2})) = \{II\}$ for all θ_{-2}; $H_3(S_3(h^*_{-3})) = \{I, III\}$ and $H_3(S_3(\theta_{-3})) = \{I\}$ for all $\theta_{-3} \neq h^*_{-3}$.

This consistent collection of sets enables us to construct the following mechanism $\mu = (M, g)$ to ex-post implement SCF f: $M_1 = \{U, C, D\}$, $M_2 = \{M\}$, $M_3 = \{L, R\}$ and $g : M \to X$ is as in Table 11. Notice that $O^\mu_1(M, L) = \{I, II, III\}$ and $O^\mu_1(M, R) = \{III\}$. One can show that any EPE strategy $\sigma^*_1(\theta_1)$ equals D if $\theta_1 = \Diamond_1$ and U if $\theta_1 = h^*_1$. This conclusion is based on employing individual 1's initial endowment, object II, in her opportunity set $O^\mu_1(M, L)$. Also, the unique EPE strategy of individual 3, $\sigma^*_3(\theta_3)$, equals R if $\theta_3 = \Diamond_3$ and L if $\theta_3 = h^*_3$ as $O^\mu_3(U, M) = \{I, III\}$ and $O^\mu_3(D, M) = \{I\}$.

Ind. 2 chooses M

		Ind. 3: L	Ind. 3: R
Ind. 1	U	(I, II, III)	(III, II, I)
	C	(II, I, III)	(III, II, I)
	D	(III, II, I)	(III, II, I)

Table 11: A mechanism that ex-post implements f.

This example also displays that the revelation principle fails with independent choices: The direct mechanism associated with f is as in Table 12. At state (h^*_1, h^*_2, h^*_3), $O^\mu_1(h^*_2, h^*_3) = \{I, III\}$ and the EPE strategy of 1 must be $\sigma^*_1(\theta_1) = \Diamond_1$ for all $\theta_1 \in \{\Diamond_1, h^*_1\}$ as $III = c_1(\{I, III\}, \Diamond_1) = c_1(\{I, III\}, h^*_1)$. Thus, we obtain an impasse due to $f_1(h^*_1, h^*_2, h^*_3) = I$.

To provide comparative statics using this example, we refer to the above specifications as the *behavioral-bias case*. We construct an associated *no-behavioral-bias case* in the rational domain with the property that the only difference pertaining to individuals' choices involves individual 1's behavior.[32] Consider individuals' rational choices as repre-

[32]In the no-behavioral-bias case, individuals' choices are affected by their initial endowments while WARP holds in all states of the world.

		Ind. 3 chooses h_3^*				Ind. 3 chooses $\Diamond_3$	
		Ind. 2				Ind. 2	
		h_2^*	$\Diamond_2$			h_2^*	$\Diamond_2$
Ind. 1	h_1^*	(I, II, III)	(III, II, I)	Ind. 1	h_1^*	(III, II, I)	(III, II, I)
	$\Diamond_1$	(III, II, I)	(III, II, I)		$\Diamond_1$	(III, II, I)	(III, II, I)

Table 12: The direct mechanism associated with f.

sented by the payoffs given in Table 13. When individual 1's type is h_1^*, her choice from

Object	$U_1^{\Diamond_1}$	$U_1^{h_1^*}$	$U_2^{\Diamond_2}$	$U_2^{h_2^*}$	$U_3^{\Diamond_3}$	$U_3^{h_3^*}$
I	2	3	2	2	3	2
II	1	2	3	3	2	1
III	3	1	1	1	1	3

Table 13: Individuals utilities U_i, $i = 1, 2, 3$, of objects with no behavioral biases.

$\{I, III\}$ is $\{III\}$ in the behavioral-bias case and it is $\{I\}$ in the no-behavioral-bias case, while the choices of every individual in the no-behavioral-bias and behavioral-bias cases coincide in all other contingencies, as can be seen in Table 14. Thus, the transition from

Z	$c_1(Z, \Diamond_1)$	$c_1(Z, h_1^*)$	$c_2(Z, \Diamond_2)$	$c_2(Z, h_2^*)$	$c_3(Z, \Diamond_3)$	$c_3(Z, h_3^*)$
$\{I, II, III\}$	III	I	II	II	I	III
$\{I, II\}$	I	I	II	II	I	I
$\{I, III\}$	III	$\boxed{III}$	I	I	I	III
$\{II, III\}$	III	II	II	II	II	III

The behavioral-bias case

Z	$c_1(Z, \Diamond_1)$	$c_1(Z, h_1^*)$	$c_2(Z, \Diamond_2)$	$c_2(Z, h_2^*)$	$c_3(Z, \Diamond_3)$	$c_3(Z, h_3^*)$
$\{I, II, III\}$	III	I	II	II	I	III
$\{I, II\}$	I	I	II	II	I	I
$\{I, III\}$	III	$\boxed{I}$	I	I	I	III
$\{II, III\}$	III	II	II	II	II	III

The no-behavioral-bias case

Table 14: Individuals' choices in the behavioral-bias and no-behavioral-bias cases.

the no-behavioral-bias to the behavioral-bias case entails only individual 1's choices from $\{I, III\}$ when her type is h_1^*. Then, one can show that, unlike in the behavioral-bias case, the direct mechanism associated with f specified in Table 12 fully ex-post implements f in the no-behavioral-bias case.

Therefore, in the behavioral-bias case the planner is obliged to use a mechanism that includes individual 1's initial endowment, II, as a possible outcome in order to secure the consistency of her choices with the desired goal. However, as the SCF f does not assign individual 1 her initial endowment at any state, in the no-behavioral-bias case, ex-post implementation does not compel the planner to offer individual 1 her initial endowment.

8 Direct Mechanisms and Independent Choices

We start this section by evaluating the significance of direct mechanisms pertinent to ex-post implementation in general environments including interdependent choices. Thereby, we portray settings where ex-post implementation is attainable using intuitive mechanisms. We focus on SCFs instead of SCSs since direct mechanisms cannot coordinate selections of SCFs from an SCS. Given an SCF f, there is an intertwined link between the consistency of the collection $\mathbb{F} := \{f(\Theta_i, \theta_{-i}) \mid i \in N, \theta_{-i} \in \Theta_{-i}\}$ and ex-post implementability of f via its direct mechanism. This connection results in (i) of Theorem 4, our *first* characterization of situations in which ex-post implementation is possible only when ex-post implementation via the direct mechanism is possible. Moreover, we provide a *second* characterization, which is akin to Bergemann and Morris (2008, Proposition 1), using the following condition we borrow from that study: An SCF f is *full-range* if for all $x \in X$, all $i \in N$, and all $\theta_{-i} \in \Theta_{-i}$, there is $\theta_i \in \Theta_i$ with $f(\theta) = x$.

Theorem 4. *Let $f : \Theta \to X$ be an SCF.*

- (i) *f is (fully) ex-post implementable by its associated direct mechanism possessing a truthful EPE if and only if the collection $\mathbb{F} := \{f(\Theta_i, \theta_{-i}) : i \in N, \theta_{-i} \in \Theta_{-i}\}$ is consistent with f under incomplete information.* [33], [34]
- (ii) *If f is full-range, then f is ex-post implementable if and only if it is (fully) ex-post implementable via its direct mechanism.*

Next, we analyze the implications of independent choices with regard to ex-post implementation. Let i's choices at $\theta \in \Theta$ be described by $C_i^{\theta_i} : \mathcal{X} \to \mathcal{X}$ such that $C_i^{\theta_i}(S) \subset S$ for any $S \in \mathcal{X}$. Then, consistency under incomplete information simplifies to: A collection of sets $\mathbb{S} := \{S_i(\theta_{-i}) \mid i \in N, \theta_{-i} \in \Theta_{-i}\} \subset \mathcal{X}$ is ***independent–consistent with the SCF f under incomplete information*** if

[33] The direct mechanism associated with f may also have an untruthful EPE. But, the outcome of this EPE must coincide with f whenever f is (fully) ex-post implementable by its direct mechanism.

[34] The example of Section 7 shows that indirect ex-post implementation is achievable while direct ex-post implementation is not possible even when individuals' choices are independent: f is ex-post implementable via the indirect mechanism given in Table 11, whereas $\mathbb{F}$ is not consistent with f because $H_1(f(\Theta_1, h^*_{-1})) = \{I, III\}$ but, in any consistent collection $\mathbb{S}$, $H_1(S_1(f, h^*_{-1}))$ must equal $\{I, II, III\}$.

(i) for all $i \in N$ and all $\theta_i' \in \Theta_i$, $f(\theta_i', \theta_{-i}) \in C_i^{\theta_i'}(S_i(\theta_{-i}))$ for all $\theta_{-i} \in \Theta_{-i}$, and

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

By (i) of independent-consistency, for all $i \in N$ and $\theta_{-i} \in \Theta_{-i}$, $f(\Theta_i, \theta_{-i}) \subset S_i(\theta_{-i})$. Thus, independent-consistency implies weak choice strategy-proofness (quasi-ex-post choice incentive compatibility with independent choices): An SCF $f : \Theta \to X$ is ***weak choice strategy-proof*** if there is a collection of sets $\mathbb{S}^* = \{S_i^*(\theta_{-i}) \mid i \in N, \theta_{-i} \in \Theta_{-i}\}$ with for all $i \in N$ and all $\theta_{-i} \in \Theta_{-i}$, $f(\Theta_i, \theta_{-i}) \subset S_i^*(\theta_{-i})$ and $f(\hat{\theta}_i, \theta_{-i}) \in C_i^{\hat{\theta}_i}(S_i^*(\theta_{-i}))$ for all $\hat{\theta}_i \in \Theta_i$; ***choice strategy-proofness*** holds if for all $i \in N$, $f(\theta_i', \theta_{-i}) \in C_i^{\theta_i'}(f(\Theta_i, \theta_{-i}))$ for all $\theta_i' \in \Theta_i$. Under the IIA, these notions are equivalent; under rationality, they equal *dominant strategies incentive compatibility* (Bergemann & Morris, 2008, Definition 21).[35]

Another equilibrium concept suitable to incomplete information and independent choices is the notion of dominant equilibrium analyzed in Mizukami and Wakayama (2007) in the rational domain: Given a mechanism μ, a strategy profile $s = (s_i)_{i \in N}$ with $s_i : \Theta_i \to M_i$ is a ***dominant equilibrium*** of μ if for all $i \in N$ and all $\theta_i \in \Theta_i$, $g(s_i(\theta_i), \tilde{m}_{-i}) \in C_i^{\theta_i}(O_i^{\mu}(\tilde{m}_{-i}))$ for all $\tilde{m}_{-i} \in M_{-i}$. But s is an EPE of μ if for all $i \in N$ and $\theta_i \in \Theta_i$, $g(s_i(\theta_i), s_{-i}(\theta_{-i})) \in C_i^{\theta_i}(O_i^{\mu}(s_{-i}(\theta_{-i})))$. Thus, the intricate relation between the dominant equilibrium and EPE with independent choices is: (i) any dominant equilibrium of μ is an EPE of μ; (ii) if s is an EPE of μ such that for all $\tilde{m}_{-i} \in M_{-i}$ there is $\tilde{\theta}_{-i} \in \Theta_{-i}$ with $s_{-i}(\tilde{\theta}_{-i}) = \tilde{m}_{-i}$ (s satisfies a full-range condition), then s is a dominant equilibrium of μ. So, given f, the truthtelling strategy profile is a dominant equilibrium of f's direct mechanism μ^f if and only if it is an EPE of μ^f.

With interdependence, dominant equilibrium loses its appeal due to its stringent requirements we describe below: Given a mechanism μ, s^* is a ***dominant equilibrium of*** μ ***with interdependent choices*** if for all $i \in N$ and all $\theta \in \Theta$, $g(s_i^*(\theta_i), \tilde{m}_{-i}) \in C_i^{(\theta_i, \theta_{-i})}(O_i^{\mu}(\tilde{m}_{-i}))$ for all $\tilde{m}_{-i} \in M_{-i}$. Since s_i^* is measurable only with respect to Θ_i, s^* being a dominant equilibrium of μ with interdependent choices is equivalent to for all $i \in N$ and for all $\theta_i \in \Theta_i$, $g(s_i^*(\theta_i), \tilde{m}_{-i}) \in C_i^{(\theta_i, \tilde{\theta}_{-i})}(O_i^{\mu}(\tilde{m}_{-i}))$ for all $\tilde{m}_{-i} \in M_{-i}$ and all $\tilde{\theta}_{-i} \in \Theta_{-i}$. Therefore, s_i^* is a strategy such that for any $\theta_i \in \Theta_i$, $s_i^*(\theta_i)$ leads to a chosen alternative from each opportunity set of i generated by any one of others' messages $\tilde{m}_{-i} \in M_{-i}$ at any one of the states $(\theta_i, \tilde{\theta}_{-i})$, for all type profiles of others $\tilde{\theta}_{-i} \in \Theta_{-i}$.

[35] The SCF f of the example in Section 7 is weak choice strategy-proof, but not choice strategy-proof.

9 Simple Mechanisms

There is a growing interest in simple mechanisms in the mechanism design literature.[36] Dealing with individuals having limited cognitive abilities increases the relevance and importance of the simplicity of mechanisms. But, "[t]he question as to what constitutes a "simple" mechanism is a difficult and controversial one" (Dutta, Sen, & Vohra, 1995).

We consider the total number of message profiles of a mechanism as a measure of its *simplicity*. Our measure is similar in spirit with the total size of message spaces used to analyze communication complexity in Segal (2007, 2010) building upon the literature on realization, message processes, and communication protocols (Williams, 1986; Reichelstein & Reiter, 1988).[37] These studies seek the "minimal information that must be elicited by the designer in order to achieve the goals" within the framework of communication protocols with verification properties, whereas our analysis aims to answer the same question restricting attention directly to mechanisms implementing a given social goal. As a result, even though our notions of simplicity are similar, they do not produce perfectly aligned implications (see Appendix F for the details).

Using our necessity result, we derive lower bounds on the cardinality of message profiles needed for behavioral implementation under incomplete information: In the proof of Theorem 1, the collection of sets $\mathbb{S} = \{S_i(f, \theta_{-i}) \mid f \in F, i \in N, \theta_{-i} \in \Theta_{-i}\}$ consistent with the SCS F is constructed from a mechanism that ex-post implements F. When there are multiple such mechanisms, there could be different 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. Let $\{\mathbb{S}^\lambda\}_{\lambda \in \Lambda}$ be the set of all consistent collections of sets represented by $\mathbb{S}^\lambda = \{\mathbb{S}_i^\lambda\}_{i \in N}$ for each $\lambda \in \Lambda$ with $\mathbb{S}_i^\lambda = \{S_i^\lambda(f, \theta_{-i}) \mid f \in F, \theta_{-i} \in \Theta_{-i}\}$. The goal of the planner is to select one of these collections and design a mechanism that ex-post implements F. Suppose that $\mathbb{S}^\lambda$ is the collection of opportunity sets generated by the mechanism that ex-post implements F. Then, individual i is able to generate any set in $\mathbb{S}_i^\lambda$ and hence i must have at least as many messages as the cardinality of the maximal set in $\mathbb{S}_i^\lambda$. This implies that the minimum number of messages required for individual i is $\min_{\lambda \in \Lambda} \max_{S \in \mathbb{S}_i^\lambda} \#S$. At the same time, for each different set in $\mathbb{S}_i^\lambda$, there must exist a particular message profile of the individuals other than i that should allow i to generate this particular set, which implies that the minimum number of message profiles required for the individuals other than i is $\min_{\lambda \in \Lambda} \#\mathbb{S}_i^\lambda$. So, the total number of message profiles in this mechanism must be at least as much as the cardinality of $\mathbb{S}_i^\lambda$ times the cardinality of the maximal set

[36]See for example, Li (2017), Borgers and Li (2018), and Pycia and Troyan (2019).

[37]We thank an anonymous referee for pointing us toward the link with these studies.

in $\mathbb{S}_i^\lambda$ for each $i \in N$. That is, the measure of simplicity of this mechanism is at least $\max_{i \in N}(\#\mathbb{S}_i^\lambda \max_{S \in \mathbb{S}_i^\lambda} \#S)$. Furthermore, the total number of message profiles required in this mechanism must be more than $\prod_{i \in N} \max_{S \in \mathbb{S}_i^\lambda} \#S$. Combining these, we observe that the total number of message profiles must exceed both $\min_{\lambda \in \Lambda} \max_{i \in N}(\#\mathbb{S}_i^\lambda \max_{S \in \mathbb{S}_i^\lambda} \#S)$ and $\min_{\lambda \in \Lambda}(\prod_{i \in N} \max_{S \in \mathbb{S}_i^\lambda} \#S)$.

The following theorem summarizes the lower bounds established above:

Theorem 5. *In any mechanism that ex-post implements the SCS $F \in \mathcal{F}$,*

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

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

(iii) *the minimum number of total message profiles is*
$\max\left\{\min_{\lambda \in \Lambda} \max_{i \in N}(\#\mathbb{S}_i^\lambda \max_{S \in \mathbb{S}_i^\lambda} \#S), \min_{\lambda \in \Lambda}(\prod_{i \in N} \max_{S \in \mathbb{S}_i^\lambda} \#S)\right\}$.

Whether or not the presence of behavioral aspects initiate simpler mechanisms is an interesting and natural question. But it needs a *structure* for being well-defined. The example of Section 7 provides a suitable framework in which we address this question in the context of allocation problems with endowment effects and independent choices by comparing the *no-behavioral-bias case* to the *behavioral-bias case* with only one individual suffering severely from a status-quo bias (see Table 14). In that setting, we provide an answer to this question: behavioral aspects induce less simple mechanisms.

First, notice that in that example individual 2's behavior is independent of her type and hence we now let $\Theta_2 = \{\Diamond_2\}$. Second, the transition from the no-behavioral-bias to the behavioral-bias case is *minimal* in the sense that it entails a deviation from rationality only for individual 1's choices from $\{I, II, III\}$ and $\{I, III\}$ when her type is h_1^*.

In the no-behavioral-bias case, the planner may use a direct mechanism to ex-post implement the given SCF $f : \Theta \to X$, which follows from Theorem 4 and $\mathbb{F} := \{f(\Theta_i, \theta_{-i}) : i \in N, \theta_{-i} \in \Theta_{-i}\}$ being independent-consistent with f under incomplete information.[38] The direct mechanism associated with f is as in Table 15 and one can verify that this mechanism ex-post implements f. The total number of message profiles, its measure of simplicity, is *four*. In the behavioral-bias case, the mechanism in Table 11 ex-post implements f, and the measure of its simplicity equals *six*.

We observe the following: First, the direct mechanism given in Table 15 does not ex-post implement f in the behavioral-bias case. Second, the indirect mechanism presented

[38] Then, f is strategy-proof (Mizukami & Wakayama, 2007) as it is choice strategy-proof (by necessity), a notion which is equivalent to strategy-proofness under rationality. See page 29 for the details.

		Ind. 2 chooses $\lozenge_2$	
		Ind. 3	
		h_3^*	$\lozenge_3$
Ind. 1	h_1^*	(I, II, III)	(III, II, I)
	$\lozenge_1$	(III, II, I)	(III, II, I)

Table 15: The direct mechanism of the no-behavioral bias case.

in Table 11 ex-post implements f both in the behavioral-bias and no behavioral-bias cases. Third, these mechanisms are the simplest ones of the corresponding cases. The main cause of this discrepancy involves whether or not the mechanism offers individual 1 her initial endowment (which is not assigned to her at any state under f) as an option to ensure the consistency of her choices with f: She needs to switch her choices between her types $\lozenge_1$ and h_1^* (as called for by f) in the behavioral-bias case. Thus, she must have an additional action resulting in her initial endowment. Therefore, the simplest mechanism in the behavioral-bias case is less simple than the simplest mechanism in the no-behavioral-bias case.

The general analysis of the question about whether or not behavioral aspects initiate less simple mechanisms necessitates adopting a particular behavioral bias and a systematic method of associating resulting cases with their rational counterparts. We leave the analysis of this interesting subject for future research.

10 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 EPE because (i) the revelation principle fails for partial implementation, and hence, one cannot restrict attention to direct mechanisms without a loss of generality; and (ii) the concept of EPE 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 violate WARP.

We provide necessary as well as sufficient conditions. These help us analyze (constrained) efficiency and allocation problems with endowment effects. We present conditions characterizing instances when the social goal is ex-post implementable via direct mechanisms. Finally, 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 Proofs

A.1 Proof of Claim 1

We identify all EPE of $\mu = (M, g)$ by a case by case analysis on what Ann 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), Ann chooses n from $\{c, n\}$ and hence $\sigma_A^*(\gamma_A)$ must be either U or D. But, at (γ_A, γ_B), Ann chooses c from $\{c, n\}$ which implies $\sigma_A^*(\gamma_A)$ must be M, a contradiction.

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), Ann chooses n from $\{c, n\}$, which implies $\sigma_A^*(\gamma_A)$ must be either U or D. At (γ_A, γ_B), Ann chooses c and s from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be M or D. So, $\sigma_A^*(\gamma_A) = D$.

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

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

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), Ann chooses n from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. On the other hand, at (γ_A, γ_B), Ann 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 $g(\sigma^*(\gamma_A, \rho_B)) = n$, a contradiction.

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), Ann chooses n from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. But, at (γ_A, γ_B), Ann chooses c and s from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be either M or D, a contradiction.

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) Ann chooses c from the set $\{c, s\}$, which implies it must be that $\sigma_A^*(\rho_A) = U$.

Then, 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 $O_A^\mu(\sigma_B^*(\rho_B)) = O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, s\}$. At (γ_A, ρ_B) and (γ_A, γ_B), Ann 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 $\sigma_B''^*(\rho_B) = M$, $\sigma_B''^*(\gamma_B) = M$.

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

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 $O_A^\mu(\sigma_B^*(\rho_B)) = \{c, s\}$ and $O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, n, s\}$. At (γ_A, ρ_B), Ann chooses c from $\{c, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. On the other hand, at (γ_A, γ_B), Ann chooses c and s from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be M or D, a contradiction.

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 $O_A^\mu(\sigma_B^*(\rho_B)) = \{c, n, s\}$ and $O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, s\}$. At (γ_A, ρ_B), Ann chooses n from $\{c, n, s\}$, and at (γ_A, γ_B), Ann chooses c from $\{c, n\}$. They both imply we must have $\sigma_A^*(\gamma_A) = U$. Thus, $O_B^\mu(\sigma_A^*(\gamma_A)) = \{c, n\}$. But, at (γ_A, ρ_B), Bob chooses c from $\{c, n\}$ even though it would be $g(\sigma^*(\gamma_A, \rho_B)) = n$, a contradiction.
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 $O_A^\mu(\sigma_B^*(\rho_B)) = O_A^\mu(\sigma_B^*(\gamma_B)) = \{c, n, s\}$. At (γ_A, ρ_B), Ann chooses n from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be U. On the other hand, at (γ_A, γ_B), Ann chooses c, s from $\{c, n, s\}$, which implies $\sigma_A^*(\gamma_A)$ must be M or D, a contradiction.

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 $\theta \in \Theta$: $g(\sigma''^*(\rho_A, \rho_B)) = g((D, M)) = g((M, M)) = s = g(\sigma'''^*(\rho_A, \rho_B))$; $g(\sigma''^*(\rho_A, \gamma_B)) = g((D, M)) = g((M, M)) = s = g(\sigma'''^*(\rho_A, \gamma_B))$; $g(\sigma''^*(\gamma_A, \rho_B)) = g((U, M)) = c = g(\sigma'''^*(\gamma_A, \rho_B))$; $g(\sigma''^*(\gamma_A, \gamma_B)) = g((U, M)) = c = g(\sigma'''^*(\gamma_A, \gamma_B))$. □

A.2 Proofs of Propositions 1, 2, and 3

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. □

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 the first condition 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 the second condition. □

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 $f(\Theta_i, \theta_{-i}) \subseteq S$; by the IIA, $f(\theta) \in C_i^\theta(\{f(\theta_i', \theta_{-i}) | \theta_i' \in \Theta_i\})$. Therefore, the revelation principle holds. □

A.3 The mechanism with three or more individuals

Our mechanism 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, e.g., by a modulo game or an integer game.[39]

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 $S_i(f, \theta_{-i})$.

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 16.

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

Table 16: 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 $k_i \in N = \{1, 2, \ldots, n\}$. 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, 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 $\bar{x}(i, f, \theta_{-i}) \in S_j(f, \theta_{-j})$. 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 is no undesirable EPE of the mechanism. We need at least three individuals for our mechanism to be well-defined. Otherwise, Rule 2 becomes ambiguous.

[39]Our mechanism resembles those used for sufficiency 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).

A.4 Proof of Theorem 2

Consider the mechanism $\mu = (M, g)$ constructed in Appendix A.3.

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 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 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 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 i when her type is θ_i; $\alpha_i(\theta_i)$, the reported type of i when her type is θ_i; $x_i(\theta_i)$, the alternative proposed by i when her type is θ_i; and $k_i(\theta_i)$, the number proposed by i when 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. This contradicts the choice incompatible pair property of $\bar{X}$ at $\tilde{\theta}$. So, 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 there is $i_0 \in N$, $\theta_{i_0} \in \Theta_{i_0}$ such that $f_{i_0}(\theta_{i_0}) \neq f$. Without loss of generality, let $i_0 = 1$ 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 $(\hat{\theta}_1, \theta_{-1}) \in \Theta$, 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 $\theta \in \Theta$.

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. □

A.5 Proof of Theorem 3

Consider the mechanism $\mu = (M, g)$ constructed in Appendix A.3.

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 her message—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 individuals propose $f \in F$ under σ^*. Consider the collection of product sets $\{\bar{\Theta}_f\}_{f \in F}$. $\bar{\Theta} := \bigcup_{f \in F} \bar{\Theta}_f$ describes the set of states where Rule 1 applies under σ^*.

Thus, at any $\theta \in \Theta \backslash \bar{\Theta}$, 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 $\theta \in \Theta \backslash \bar{\Theta}$. 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, there is $f^* \in F$ with $g(\sigma^*(\theta)) = f^*(\theta)$ 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, $O^{\mu}_{i^*}(\sigma^*_{-i^*}(\tilde{\theta}_{-i^*})) = S_{i^*}(f, \alpha_{-i^*}(\theta^*_{-i^*}))$, which implies $g(\sigma^*(\tilde{\theta})) \notin O^{\mu}_{i^*}(\sigma^*_{-i^*}(\tilde{\theta}_{-i^*}))$, a contradiction to σ^* being an EPE of μ.

So, $g(\sigma^*(\theta)) = f^*(\theta)$ for each $\theta \in \Theta$, i.e., (ii) of ex-post implementability holds. $\square$

A.6 Proof of Proposition 5

Consider $\mathbb{Y} := \{Y_i(e, \theta_{-i}) : i \in N, e \in E^{c.eff}, \theta_{-i} \in \Theta_{-i}\}$ such that $Y_i(e, \theta_{-i})$ equals $Y_i^{\theta_{-i}}$ associated with $e \in E^{c.eff}$ as in the defining condition (1).

Let $e \in E^{c.eff}$ and suppose there is a deception $\alpha = (\alpha_i)_{i \in N}$ where the individual deception $\alpha_i : \Theta_i \to \Theta_i$ is such that $e \circ \alpha \notin E^{c.eff}$. Then, (by letting $e \circ \alpha = e^{\alpha}$)

$$\forall (Y_i)_{i \in N} \text{ with } \cup_i Y_i = X, \exists i^* \in N, \exists \theta^*_{-i^*} \in \Theta_{-i^*}, \exists \theta^*_{i^*} \in \Theta_{i^*}, \text{ s.t. } e^{\alpha}(\theta^*) \notin C^{\theta^*}_{i^*}(Y_{i^*}). \quad (2)$$

Because $e \in E^{c.eff}$, we have that $\{Y_i(e, \theta_{-i}) : i \in N, \theta_{-i} \in \Theta_{-i}\}$ is such that $\cup_{i \in N} Y_i(e, \theta_{-i}) = X$ for all $\theta_{-i} \in \Theta_{-i}$. Thus, by (2), if $e \circ \alpha \notin E^{c.eff}$, then there exists $\tilde{\theta}$ such that $e(\tilde{\theta}) \notin C^{\theta^*}_{i^*}(Y_{i^*}(e, \tilde{\theta}_{-i^*}))$ where $\alpha_j(\theta^*_j) = \tilde{\theta}_j$ for all $j \in N$. Hence, $e(\alpha(\theta^*)) \notin C^{\theta^*}_{i^*}(Y_{i^*}(e, \alpha_{-i^*}(\theta_{-i^*})))$ delivering consistency of $\mathbb{Y}$ associated with $E^{c.eff}$ under incomplete information. $\square$

A.7 Proof of Proposition 7

Let μ ex-post implement F. Then, the existence of a collection of sets consistent with F under incomplete information follows directly from Theorem 1.

For any $f \in F$, let σ^f be the corresponding EPE with $f = g \circ \sigma^f$. Thus, (i) of consistency implies $f(h^*_i, \theta_{-i}) \in C_i^{h^*_i}(S_i(f, \theta_{-i}))$ and $f(\Diamond_i, \theta_{-i}) \in C_i^{\Diamond_i}(S_i(f, \theta_{-i}))$. Meanwhile, $C_i^{h^*_i}(S_i(f, \theta_{-i})) = \{x \in S_i(f, \theta_{-i}) | x_i = c_i(H_i(S_i(f, \theta_{-i})), h^*_i)\}$ whenever $h^*_i \in H_i(S_i(f, \theta_{-i}))$ and $C_i^{h^*_i}(S_i(f, \theta_{-i})) = \{x \in S_i(f, \theta_{-i}) | x_i = c_i(H_i(S_i(f, \theta_{-i})), \Diamond_i)\}$ whenever $h^*_i \notin H_i(S_i(f, \theta_{-i}))$. Finally, $C_i^{\Diamond_i}(S_i(f, \theta_{-i})) = \{x \in S_i(f, \theta_{-i}) | x_i = c_i(H_i(S_i(f, \theta_{-i})), \Diamond_i)\}$.

By Masatlioglu and Ok (2014, Theorem 1), we have that $c_i(H_i(S_i(f, \theta_{-i})), \Diamond_i) = \arg\max_{h \in H_i(S_i(f, \theta_{-i}))} U_i(h)$. So, if $f(\theta) = x$, then $U_i(x_i) \geq U_i(h)$ for all $h \in H_i(S_i(f, \theta_{-i}))$. Therefore, $H_i(S_i(f, \theta_{-i})) \subset LCS_i(x_i)$ establishing (i).

Suppose $f(\theta) = x$ and $x_i \notin Q_i(h^*_i)$, but also $h^*_i \in H_i(S_i(f, \theta_{-i}))$. Since $h^*_i \in H_i(S_i(f, \theta_{-i}))$, $C_i^{h^*_i}(S_i(f, \theta_{-i})) = \{x \in S_i(f, \theta_{-i}) | x_i = c_i(H_i(S_i(f, \theta_{-i})), h^*_i)\}$. $f(h^*_i, \theta_{-i}) \in C_i^{h^*_i}(S_i(f, \theta_{-i}))$ implies $x_i = c(H_i(S_i(f, \theta_{-i})), h^*_i) = \arg\max_{h \in H_i(S_i(f, \theta_{-i})) \cap Q_i(h^*_i)} U_i(h)$. But this contradicts $x_i \notin H_i(S_i(f, \theta_{-i})) \cap Q_i(h^*_i)$, establishing (ii).

Suppose now that $f(\theta) = x$ and $x_i = h^*_i$. Then, as $x_i = h^*_i \in H_i(S_i(f, \theta_{-i}))$, we have $C_i^{h^*_i}(S_i(f, \theta_{-i})) = \{x \in S_i(f, \theta_{-i}) | x_i = c_i(H_i(S_i(f, \theta_{-i})), h^*_i)\}$. Hence, $h^*_i = c(H_i(S_i(f, \theta_{-i})), h^*_i)$. Ergo, $h^*_i = \arg\max_{h \in H_i(S_i(f, \theta_{-i})) \cap Q_i(h^*_i)} U_i(h)$. Furthermore, $U_i(h) \geq U_i(h^*_i)$ for all $h \in Q_i(h^*_i)$.Therefore, $H_i(S_i(f, \theta_{-i})) \cap Q_i(h^*_i) = \{h^*_i\}$, establishing (iii).

Next, suppose $f(\theta) = x$ and $x_i \neq h^*_i$. Then, either $x_i \notin Q_i(h^*_i)$ and hence by (iii) above, $h^*_i \notin H_i(S_i(f, \theta_{-i}))$ or $x_i \in Q_i(h^*_i)$ establishing (iv). $\square$

A.8 Proof of Theorem 4

For the *necessity* of (i) of the theorem, suppose f is ex-post implementable by its direct mechanism $\mu^f = (\Theta, g^f)$ with $g^f = f$. Due to full ex-post implementation, let the truthful EPE be s^f with $s_i^f : \Theta_i \to \Theta_i$ for all $i \in N$ and $f = g^f \circ s^f$. Let $i \in N$ and $\tilde{\theta}_{-i} \in \Theta_{-i}$. So, $f(\theta_i, \tilde{\theta}_{-i}) \in C_i^{(\theta_i, \tilde{\theta}_{-i})}(f(\Theta_i, \tilde{\theta}_{-i}))$, establishing (i) of consistency of $\mathbb{F}$. This follows from $O_i^{\mu^f}((s_j^f(\tilde{\theta}_j))_{j \neq i}) = f(\Theta_i, \tilde{\theta}_{-i})$ as $s_i^f(\theta_i) = \theta_i$ for all $i \in N$ and $\theta_i \in \Theta_i$. For (ii) of consistency, for any deception α with $f \circ \alpha \neq f$, $s^f \circ \alpha$ cannot be an EPE of μ^f because otherwise $g^f \circ s^f \circ \alpha = f \circ \alpha$ and hence by (ii) of ex-post implementation $f \circ \alpha$ equals f, a contradiction. Thus, there is $i^* \in N$, $\theta^* \in \Theta$ with $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(f(\Theta_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*})))$ since $O_{i^*}^{\mu^f}((s_j^f(\alpha_j(\theta_j^*)))_{j \neq i^*}) = f(\Theta_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*}))$. For the *sufficiency* of (i) of the theorem: By hypothesis, $\mathbb{F}$ is consistent with f. Define $\bar{s}$ by $\bar{s}_i(\theta_i) = \theta_i$ for any $i \in N$ and $\theta_i \in \Theta_i$. Then, $\bar{s}$ is a truthful EPE strategy and $g^f \circ \bar{s} = f$: This is because $g^f(\bar{s}(\theta)) = f(\theta)$ for all $\theta \in \Theta$; for all $i \in N$ and all $\tilde{\theta}_{-i} \in \Theta_{-i}$, $O_i^{\mu^f}(\tilde{\theta}_{-i}) = f(\Theta_i, \tilde{\theta}_{-i})$ and $g^f(\bar{s}_i(\theta_i'), (\bar{s}_j(\tilde{\theta}_j)_{j \neq i})) = f(\theta_i', \tilde{\theta}_{-i}) \in C_i^{(\theta_i', \tilde{\theta}_{-i})}(f(\Theta_i, \tilde{\theta}_{-i}))$ for all $\theta_i' \in \Theta_i$, by (i) of consistency. Further, if s^* is an EPE, then $g^f \circ s^* = f$: Suppose s^* is an EPE and $g^f(s^*(\theta)) \neq f(\theta)$ for some $\theta \in \Theta$. Let α be such that $\alpha(\theta) = s^*(\theta) \neq \theta$. Then, $f(\alpha(\theta)) \neq f(\theta)$. Hence, by (ii) of consistency of $\mathbb{F}$, there is $i^* \in N$ and $\theta^* \in \Theta$ with $g^f(s^*(\theta^*)) = f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(f(\Theta_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*})))$, contradicting to s^* being an EPE as $O_{i^*}^{\mu^f}(\alpha_{-i^*}(\theta^*_{-i^*})) = f(\Theta_{i^*}, \alpha_{-i^*}(\theta^*_{-i^*}))$.

For (ii) of the theorem: The sufficiency holds trivially. If f is full-range and ex-post implementable by a mechanism μ and s is an EPE of μ with $g \circ s = f$, then for all $i \in N$, $O_i^{\mu}(s_{-i}(\theta_{-i})) = X$ for all $\theta_{-i} \in \Theta_{-i}$. Hence, $\mathbb{S} = \{S_i(\theta_{-i}) \mid i \in N, \theta_{-i} \in \Theta_i\}$ such that $S_i(\theta_{-i}) = X$ for all $i \in N$ and $\theta_{-i} \in \Theta_{-i}$ is consistent with f under incomplete information. Therefore, (i) of consistency implies $f(\theta_i', \theta_{-i}) \in C_i^{(\theta_i', \theta_{-i})}(X)$; (ii) of consistency implies for any α with $f \circ \alpha \neq f$, there is $i^* \in N$ and θ^* such that $f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(X)$. Now, consider f's direct mechanism, μ^f, and observe that for all $i \in N$ and $O_i^{\mu^f}(\theta_{-i}) = f(\Theta_i, \theta_{-i}) = X$ for all $\theta_{-i} \in \Theta_{-i}$. Therefore, the truthtelling strategy profile s^T with $s_i^T(\theta_i) = \theta_i$ for all $i \in N$ and $\theta_i \in \Theta_i$ is an EPE of μ^f because (by (i) of consistency) for any $i \in N$ and $\theta_i \in \Theta_i$, $g^f(s_i^T(\theta_{-i}), s_{-i}^T(\tilde{\theta}_{-i})) = f(\theta_i, \tilde{\theta}_{-i}) \in C_i^{(\theta_i, \tilde{\theta}_{-i})}(X)$ for all $\tilde{\theta}_{-i} \in \Theta_{-i}$. For any other EPE $\tilde{s}$ and for any $\tilde{\theta} \in \Theta$, it must be that $g^f(\tilde{s}(\tilde{\theta})) = f(\hat{\theta})$ for some $\hat{\theta} \in \Theta$, due to the full-range condition. If $f(\hat{\theta}) \neq f(\tilde{\theta})$, construct α such that $\alpha(\theta) = \tilde{s}(\theta)$ for all $\theta \in \Theta$. Then, $f(\alpha(\tilde{\theta})) = f(\hat{\theta}) \neq f(\tilde{\theta})$; so, by (ii) of consistency, there is $i^* \in N$ and $\theta^* \in \Theta$ such that $g^f(\tilde{s}(\theta^*)) = f(\alpha(\theta^*)) \notin C_{i^*}^{\theta^*}(X)$. Thus, $\tilde{s}$ cannot be an EPE, a contradiction. $\square$

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.

Benoit, J.-P., & Ok, E. A. (2006). Maskin's theorem with limited veto power. *Games and Economic Behavior*, *55*(2), 331–339.

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.

Bergemann, D., & Morris, S. (2011). Robust implementation in general mechanisms. *Games and Economic Behavior*, *71*(2), 261–281.

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.

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

Bochet, O., & Tumennasan, N. (2018). One truth and a thousand lies: Defaults and benchmarks in mechanism design. *Unpublished*.

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.

de Clippel, G. (2020). Incorporating risk in choice theory: Some observations.

de Clippel, G., & Eliaz, K. (2012). Reason-based choice: A bargaining rationale for the attraction and compromise effects. *Theoretical Economics*, *7*(1), 125–162.

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

Dutta, B., Sen, A., & Vohra, R. (1995). Nash implementation through elementary mechanisms in economic environments. *Economic Design*, *1*(1), 173–203.

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.

Kucuksenel, S. (2012). Behavioral mechanism design. *Journal of Public Economic Theory*, *14*(5), 767–789.

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

Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, *102*(5), 2183–2205.

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. (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.

Mizukami, H., & Wakayama, T. (2007). Dominant strategy implementation in economic environments. *Games and Economic Behavior*, *60*(2), 307–325.

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. *Econometrica*, 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.

Pycia, M., & Troyan, P. (2019). A theory of simplicity in games and mechanism design.

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

Reichelstein, S., & Reiter, S. (1988). Game forms with minimal message spaces. *Econometrica: Journal of the Econometric Society*, 661–692.

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.

Saran, R. (2016). Bounded depths of rationality and implementation with complete information. *Journal of Economic Theory*, *165*, 517–564.

Segal, I. (2007). The communication requirements of social choice rules and supporting budget sets. *Journal of Economic Theory*, *136*(1), 341–378.

Segal, I. (2010). Nash implementation with little communication. *Theoretical Economics*, *5*(1), 51–71.

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.

Shapley, L., & Scarf, H. (1974). On cores and indivisibility. *Journal of mathematical economics*, *1*(1), 23–37.

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.

Williams, S. R. (1986). Realization and nash implementation: two aspects of mechanism design. *Econometrica: Journal of the Econometric Society*, 139–151.