Partially-honest Nash implementation with non-connected honesty standards

Michele Lombardi* and Naoki Yoshihara†

July 31, 2015

As per Ariely (2012), an individual may display an honesty standard which allows her to lie or cheat a little without that being harmful to herself view as an honest person. On this basis, the paper considers a society with a finite number of individuals involving partially-honest individuals and in which every individual has her own honesty standard. An individual honesty standard is modeled as a subgroup of the society, including the individual herself, for whom she feels truth-telling concerns. A partially-honest individual is an individual who strictly prefers to tell the truth prescribed by her honesty standard whenever lying has no effect on her material well-being. The paper studies the impact of placing honesty standard restrictions on the mechanism designer for Nash implementation problems of that society. It offers a necessary condition for Nash implementation, called partial-honesty monotonicity, and shows that in an independent domain of preferences that condition is equivalent to Maskin monotonicity, provided that honesty standards of society are non-connected. They are non-connected if every individual is excluded from the honesty standard of another individual.

JEL classification: C72; D71; D82.

Keywords: Nash implementation; partial-honesty; non-connected honesty standards, independent domain.

*Department of Economics, University of Glasgow, Glasgow, G12 8QQ, United Kingdom, e-mail: michele.lombardi@glasgow.ac.uk.

†Institute of Economic Research, Hitotsubashi University, 2-4 Naka, Kunitachi, Tokyo, 186-8603 Japan, e-mail: yosihara@ier.hit-u.ac.jp.

1 Introduction

The implementation problem is the problem of designing a mechanism or game form with the property that for each profile of participants' preferences, the equilibrium outcomes of the mechanism played with those preferences coincide with the recommendations that a given social choice rule (SCR) would prescribe for that profile. If that mechanism design exercise can be accomplished, the SCR is said to be implementable. The fundamental paper on implementation in Nash equilibrium is thanks to Maskin (1999; circulated since 1977), who proves that any SCR that can be Nash implemented satisfies a remarkably strong invariance condition, now widely referred to as Maskin monotonicity. Moreover, he shows that when the mechanism designer faces at least three individuals, an SCR is Nash implementable if it is Maskin monotonic and satisfies the condition of no veto-power, subsequently, *Maskin's theorem.*

Since Maskin's theorem, economists have been interested in understanding how to circumvent the limitations imposed by Maskin monotonicity by exploring the possibilities offered by approximate (as opposed to exact) implementation (Matsushima, 1988; Abreu and Sen, 1991), as well as by implementation in refinements of Nash equilibrium (Moore and Repullo, 1988; Abreu and Sen, 1990; Palfrey and Srivastava, 1991; Jackson, 1992) and by repeated implementation (Kalai and Ledyard, 1998; Lee and Sabourian, 2011; Mezzetti and Renou, 2012). One additional way around those limitations is offered by implementation with partially-honest individuals.

A *partially-honest individual* is an individual who strictly prefers to tell the truth whenever lying has no effect on her material well-being. In a general environment, a seminal paper on Nash implementation problems involving partially-honest individuals is Dutta and Sen (2012), which shows that for implementation problems involving at least three individuals and in which there is at least one partially-honest individual, the Nash implementability is assured by no veto-power; subsequently *Dutta-Sen's theorem.* Similar positive results are uncovered in other environments by Matsushima (2008a,b), Kartik and Tercieux (2012), Kartik et al. (2014), and Ortner (2015). Thus, there are far fewer limitations for Nash implementation where there are partially-honest individuals.

One way to put those studies into perspective with this paper is to realize that a participant chooses a preference announcement as part of her strategy choice. Moreover, a participant's play is honest if she plays a strategy choice which conveys to the mechanism designer the true preference announcement. Therefore, their common ground is that the mechanism designer establishes a unique honesty (equivalently, truth-telling) standard which spells out to participants the boundary between an honest and a dishonest play of preference announcements and enforces participants to endorse it. Simply put, participants are not free to maintain their own view of honesty through their play. However, there are cases in which human factors such as past individual experiences and cultural traits may blur the boundary between honest and dishonest behavior and which, more importantly, make honesty standards not uniform across individuals. As Dan Ariely (2012) aptly noted:[1]

> [...] our sense of our own morality is connected to the amount of cheating we feel comfortable with. Essentially, we cheat up to the level that allows us to retain our self-image as reasonably honest individuals.
>
> (Ariely, 2012, pp.22-23)

How should we think of more general types of individual honesty standards? Uniformity of honesty standards across individuals and their coercion is a strong assumption in regards to both aspects. How sensitive are the main results obtained thus far to that assumption? Do personal views of honesty enhance the scope of Nash implementation with partially-honest individuals, or do they hinder it? Do we learn new principles for actual mechanism design?

This paper models an *individual honesty standard* as a subset of individuals involved with an implementation problem. Our interpretation is that participant i concerns herself with the truth-telling of individuals in her honesty standard when she plays a strategy choice. Our definition endorses the view that an individual concerns herself with at least her own self. Then, our study looks at what SCR can be

1. Ariely (2012; pp. 28-29) also refers to the Jerome's classic British novel, *Three Men in a Boat (to Say Nothing of the Dog)*, to corroborate his experimental findings that "each of us has a limit to how much we can cheat before it becomes absolutely 'sinful.'"

Nash implemented in a society involving partially-honest individuals, in which participants share the responsibility for maintaining their own honesty standards and in which the mechanism designer takes those standards as an institutional constraint.

First, we assume that individual honesty standards are known to the mechanism designer and that he has to respect them. Under this institutional constraint, we show that any SCR that can be Nash implemented with partially-honest individuals satisfies a variant of Maskin monotonicity, called *partial-honesty monotonicity*.

The idea of this axiom is quite intuitive. If x is one of the outcomes selected by a given SCR at one preference profile but is not selected when there is a monotonic change of preferences around x, then that monotonic change has altered preferences of individuals in the honesty standard of a partially-honest individual.

Second, we consider what we call *non-connected honesty standards*. Simply put, individual honesty standards are connected if there is a participant with whom all other participants are jointly concerned. When that is not the case, we call it non-connected honesty standards. In other words, they are non-connected if every participant is excluded from the honesty standard of another participant.

In an independent domain of preferences, where the set of the profiles of participants' preferences takes the structure of the Cartesian product of individual preferences, we show that partial-honesty monotonicity is *equivalent* to Maskin monotonicity. Thus, under those hypotheses, Maskin's theorem provides an almost complete characterization of SCRs that are Nash implementable, though partially-honest individuals are involved in the society.

The remainder of the paper is divided into four sections. Section 2 presents the theoretical framework and outlines the implementation model, with the necessary condition presented in section 3. Section 4 presents the equivalence result. Section 5 concludes.

2 Preliminaries

2.1 Basic framework

We consider a finite set of individuals indexed by $i \in N = \{1, \cdots, n\}$, which we will refer to as a society. The set of outcomes available to individuals is X. The information held by the individuals is summarized in the concept of a state. Write Θ for the domain of possible states, with θ as a typical state. In the usual fashion, individual i's preferences in state θ are given by a complete and transitive binary relation, subsequently an ordering, $R_i(\theta)$ over the set X. The corresponding strict and indifference relations are denoted by $P_i(\theta)$ and $I_i(\theta)$, respectively. The preference profile in state θ is a list of orderings for individuals in N that are consistent with that state and is denoted by $R_N(\theta)$.

We assume that the mechanism designer does not know the true state. We assume, however, that there is complete information among the individuals in N. This implies that the mechanism designer knows the preference domain consistent with the domain Θ. In this paper, we identify states with preference profiles.

The goal of the mechanism designer is to implement an SCR $F : \Theta \twoheadrightarrow X$ where $F(\theta)$ is non-empty for any $\theta \in \Theta$. We shall refer to $x \in F(\theta)$ as an F-optimal outcome at θ. Given that individuals will have to be given the necessary incentives to reveal the state truthfully, the mechanism designer delegates the choice to individuals according to a mechanism $\Gamma \equiv \left(\prod_{i \in N} M_i, g \right)$, where M_i is the strategy space of individual i and $g : M \to X$, the outcome function, assigns to every strategy profile $m \in M \equiv \prod_{i \in N} M_i$ a unique outcome in X. We shall sometimes write (m_i, m_{-i}) for the strategy profile m, where $m_{-i} = (m_1, \cdots, m_{i-1}, m_{i+1}, \cdots, m_n)$.

An *honesty standard of individual* i, denoted by $S(i)$, is a subgroup of society with the property that $i \in S(i)$. Thus, given a state θ, $R_{S(i)}(\theta)$ is a list of orderings consistent with θ for individuals in the honesty standard $S(i)$ of individual i. An *honesty standard of society* is a list of honesty standards for all members of society. Write $S(N)$ for a typical honesty standard of society.

2.2 Intrinsic preferences for honesty

An individual who has an intrinsic preference for truth-telling can be thought as an individual who is torn by a fundamental conflict between her deeply and ingrained propensity to respond to material incentives and the desire to think of herself as a honest person (Ariely, 2012). In this paper, the theoretical construct of the balancing act between those contradictory desires is based on two ideas.

First, the triplet $(\Gamma, \theta, S(i))$ acts as a "context" for individuals' conflicts. The reason for this is that an individual who has intrinsic preferences for honesty can categorize her strategy choices as truthful or untruthful relative to her honesty standard $S(i)$, the state θ and the mechanism Γ designed by the mechanism designer to govern the communication with her. That categorization can be captured by the following notion of truth-telling correspondence:

Definition 1. For each Γ and each individual $i \in N$ with an honesty standard $S(i)$, individual i's *truth-telling correspondence* is a (non-empty) correspondence $T_i^{\Gamma}(\cdot; S(i)) : \Theta \twoheadrightarrow M_i$ with the property that for any two states θ and θ', it holds that

$$T_i^{\Gamma}(\theta; S(i)) = T_i^{\Gamma}(\theta'; S(i)) \iff R_{S(i)}(\theta) = R_{S(i)}(\theta').$$

Strategy choices in $T_i^{\Gamma}(\theta; S(i))$ will be referred to as truthful strategy choices for θ according to $S(i)$.

According to the above definition, in a state θ, every truthful strategy choice of individual i is to encode information of individuals' orderings consistent with that state for members of society in her honesty standard $S(i)$. Moreover, if in two different states, say θ and θ', the orderings consistent with those two states for individuals in $S(i)$ are the same, then the sets of individual i's truthful strategy choices for those two states need to be identical according to her honesty standard $S(i)$.

In modeling intrinsic preferences for honesty, we adapt the notion of partially-honest individuals of Dutta and Sen (2012) to our research questions. First, a partially-honest individual is an individual who responds primarily to material incentives. Second, she strictly prefers to tell the truth whenever lying has no effect on her material well-being. That behavioral choice of a partially-honest individual can be modeled by extending individual's ordering over X to an ordering over the strategy space M, because that individual's preference between being truthful and being untruthful is contingent upon announcements made by other individuals as well as the outcome(s) obtained from them. By following standard conventions of orderings, write $\succcurlyeq_i^{\Gamma,\theta,S(i)}$ for individual i's ordering over M in state θ whenever she is confronted with the mechanism Γ and has set her honesty standard at $S(i)$. Formally, our notion of a partially-honest individual is as follows:

Definition 2. For each Γ, individual $i \in N$ with an honesty standard $S(i)$ is *partially-honest* if, for all $\theta \in \Theta$, her intrinsic preference for honesty $\succcurlyeq_i^{\Gamma,\theta,S(i)}$ on M satisfies the following properties: for all m_{-i} and all $m_i, m_i' \in M_i$, it holds that

(i) If $m_i \in T_i^{\Gamma}(\theta; S(i))$, $m_i' \notin T_i^{\Gamma}(\theta; S(i))$ and $g(m) \, R_i(\theta) \, g(m_i', m_{-i})$, then $m \succ_i^{\Gamma,\theta,S(i)} (m_i', m_{-i})$.

(ii) In all other cases, $m \succcurlyeq_i^{\Gamma,\theta,S(i)} (m_i', m_{-i})$ if and only if $g(m) \, R_i(\theta) \, g(m_i', m_{-i})$.

Intrinsic preference for honesty of individual i is captured by the first part of the above definition, in that, for a given mechanism Γ, honesty standard $S(i)$ and state θ, individual i strictly prefers the message profile (m_i, m_{-i}) to (m_i', m_{-i}) provided that the outcome $g(m_i, m_{-i})$ is at least as good as $g(m_i, m_{-i})$ according to her ordering $R_i(\theta)$ and that m_i is truthful for θ and m_i' is not truthful for θ, according to $S(i)$.

If individual i is *not* partially-honest, this individual cares for her material well-being associated with outcomes of the mechanism and nothing else. Then, individual i's ordering over M is just the transposition into space M of individual i's relative ranking of outcomes. More formally:

Definition 3. For each Γ, individual $i \in N$ with an honesty standard $S(i)$ is *not partially-honest* if, for all $\theta \in \Theta$, her intrinsic preference for honesty $\succcurlyeq_i^{\Gamma,\theta,S(i)}$ on M satisfies the following property: for all $m, m' \in M$, it holds that

$$m \succcurlyeq_i^{\Gamma,\theta,S(i)} m' \iff g(m) \, R_i(\theta) \, g(m').$$

2.3 Implementation problems

In formalizing the mechanism designer's problem, we first introduce our informational assumptions and discuss their implications for our analysis. They are:

Assumption 1. There exists at least one partially-honest individual in society.

Assumption 2. The mechanism designer knows the honesty standard of society.

The above two assumptions combined with the assumption that there is complete information among the individuals imply that the mechanism designer only knows the set Θ, the fact that there is at least one partially-honest individual among the individuals and the honesty standard of society, but he does not know either the true state or the identity of the partially-honest individual(s) (or their identities). Indeed, the mechanism designer cannot exclude any member(s) of society from being partially-honest purely on the basis of Assumption 1. Therefore, the following considerations are in order from the viewpoint of the mechanism designer.

An environment is described by three parameters, $(\theta, S(N), H)$: a state θ, an honesty standard of society $S(N)$ and a conceivable set of partially-honest individuals H. We denote by H a typical conceivable set of partially-honest individuals in N, with h as a typical element, and by $\mathcal{H}$ the class of conceivable sets of partially-honest individuals.

A mechanism Γ and an environment $(\theta, S(N), H)$ induce a strategic game $\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$, where

$$\succsim^{\Gamma,\theta,S(N),H} \equiv \left(\succsim_i^{\Gamma,\theta,S(i)}\right)_{i\in N}$$

is a profile of orderings over the strategy space M as formulated in Definition 2 and in Definition 3. Specifically, $\succsim_i^{\Gamma,\theta,S(i)}$ is individual i's ordering over M as formulated in Definition 2 if individual i is in H, whereas it is the individual i's ordering over M as formulated in Definition 3 if individual i is not in H.

A (pure strategy) Nash equilibrium of the strategic game $\left(\Gamma, \succsim^{\Gamma,\theta,H,S(N)}\right)$ is a strategy profile m such that for all $i \in N$, it holds that

$$\text{for all } m_i' \in M_i : m \succsim_i^{\Gamma,\theta,S(i)} (m_i', m_{-i}).$$

Write $NE\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$ for the set of Nash equilibrium strategies of the strategic game $\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$ and $NA\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$ for its corresponding set of Nash equilibrium outcomes.

The following definition is to formulate the designer's Nash implementation problem involving partially-honest individuals in which the society maintains the standard of honesty summarized in $S(N)$.

Definition 4. Let Assumption 1 and Assumption 2 hold. Let the honesty standard of society be summarized in $S(N)$. A mechanism Γ *partially-honestly Nash implements* the SCR $F : \Theta \twoheadrightarrow X$ provided that for all $\theta \in \Theta$ and $H \in \mathcal{H}$, there exists for any $h \in H$ a truth-telling correspondence $T_h^{\Gamma}(\theta; S(h))$ as formulated in Definition 1 and, moreover, it holds that $F(\theta) = NA\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$. If such a mechanism exists, F is said to be *partially-honestly Nash implementable.*

The objective of the mechanism designer is thus to design a mechanism whose Nash equilibrium outcomes, for each state θ as well as for each conceivable set of partially-honest individuals H, coincide with $F(\theta)$. Note that there is no distinction between the above formulation and the standard Nash implementation problem as long as Assumption 1 is discarded.

3 A necessary condition

In this section, we discuss a condition that is necessary for the partially-honest Nash implementation where the honesty standard of society is prescribed by $S(N)$.

A condition that is central to the implementation of SCRs in Nash equilibrium is Maskin monotonicity. This condition says that if an outcome x is F-optimal at the state θ, and this x does not strictly fall in preference for anyone when the state is changed to θ', then x must remain an F-optimal outcome at θ'. Let us formalize that condition as follows. For any state θ, individual i and outcome x, the weak lower contour set of $R_i(\theta)$ at x is defined by $L_i(\theta, x) = \{x' \in X | x R_i(\theta) x'\}$. Therefore:

Definition 5. The SCR $F : \Theta \twoheadrightarrow X$ is *Maskin monotonic* provided that for all $x \in X$ and all $\theta, \theta' \in \Theta$, if $x \in F(\theta)$ and $L_i(\theta, x) \subseteq L_i(\theta', x)$ for all $i \in N$, then $x \in F(\theta')$.

An equivalent statement of Maskin monotonicity stated above follows the reasoning that if x is F-optimal at θ but not F-optimal at θ', then the outcome x must have fallen strictly in someone's ordering at the state θ' in order to break the Nash equilibrium via some deviation. Therefore, there must exist some (outcome-)preference reversal if an equilibrium strategy profile at θ is to be broken at θ'.

Our variant of Maskin monotonicity for Nash implementation problems involving partially-honest individuals where the standard of honesty of society is represented by $S(N)$ can be formulated as follows:

Definition 6. The SCR $F : \Theta \twoheadrightarrow X$ is *partial-honesty monotonic* given the standard $S(N)$ provided that for all $x \in X$, all $H \in \mathcal{H}$ and all $\theta, \theta' \in \Theta$, if $x \in F(\theta) \setminus F(\theta')$ and $L_i(\theta, x) \subseteq L_i(\theta', x)$ for all $i \in N$, then for one $h \in H : R_{S(h)}(\theta) \neq R_{S(h)}(\theta')$.

This says that if x is F-optimal at θ but not F-optimal at θ' and, moreover, there is a monotonic change of preferences around x from θ to θ' (that is, whenever $xR_i(\theta)x'$, one has that $xR_i(\theta')x'$), then that monotonic change has altered preferences of individuals in the honesty standard of a partially-honest individual $h \in H$ (that is, $R_{S(h)}(\theta) \neq R_{S(h)}(\theta')$). Stated in the contrapositive, this says that if x is F-optimal at θ, there is a monotonic change of preferences around x from θ to θ' and, moreover, for any conceivable partially-honest individual h in H that change has not altered preferences of individuals in her honesty standard $S(h)$, then x must continue to be one of the outcomes selected by F at the state θ'. Note that if x is F-optimal at θ but not F-optimal at θ', one has that $R_N(\theta) \neq R_N(\theta')$, and thus any SCR is partial-honesty monotonic whenever the honesty standard of society is such that $S(i) = N$ for all $i \in N$.

The above condition is necessary for partially-honest Nash implementation. This is because if x is F-optimal at θ but not F-optimal at θ' and, moreover, the outcome x has not fallen strictly in any individual's ordering at the state θ', then only a partially-honest individual in the given conceivable set H can break the Nash equilibrium via a unilateral deviation. Therefore, there must exist some strategy-profile-preference reversal for a partially-honest individual $h \in H$ if an equilibrium strategy profile at $(\theta, S(N), H)$ is to be broken at $(\theta', S(N), H)$. Formally:

Theorem 1. Let Assumption 1 and Assumption 2 be given. Let the honesty standard of society be summarized in $S(N)$. The SCR $F : \Theta \twoheadrightarrow X$ is partial-honesty monotonic given the standard $S(N)$ if it is partially-honestly Nash implementable.

Proof. Let Assumption 1 and Assumption 2 be given. Let the honesty standard of society be summarized in $S(N)$. Suppose that $\Gamma \equiv (M, g)$ partially-honest Nash implements the SCR $F : \Theta \twoheadrightarrow X$. For any $x \in X$, consider any environment $(\theta, S(N), H)$ such that $x \in F(\theta)$. Then, there is $m \in NE\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$ such that $g(m) = x$.

Consider any state $\theta' \in \Theta$ such that

$$\text{for all } i \in N \text{ and all } x' \in X : xR_i(\theta)x' \implies xR_i(\theta')x'. \tag{1}$$

If there exists an individual $i \in N$ such that $g(m_i', m_{-i}) P_i(\theta') g(m)$, then, from (1), $g(m_i', m_{-i}) P_i(\theta) g(m)$, a contradiction of the fact that $m \in NE\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$. Therefore, we conclude that

$$\text{for all } i \in N \text{ and all } m_i' \in M_i : g(m) R_i(\theta') g(m_i', m_{-i}). \tag{2}$$

Suppose that $x \notin F(\theta')$. Then, the strategy profile m is not a Nash equilibrium of $\left(\Gamma, \succsim^{\Gamma,\theta',S(N),H}\right)$, that is, there exists an individual $i \in N$ who can find a strategy choice $m_i' \in M_i$ such that $(m_i', m_{-i}) \succ_i^{\Gamma,\theta',S(i)} m$. Given that (2) holds, it must be the case that $i \in H$. From part (i) of Definition 2 we conclude, therefore, that

$$m_i \notin T_i^{\Gamma}(\theta'; S(i)) \text{ and } m_i' \in T_i^{\Gamma}(\theta'; S(i)) \tag{3}$$

and that

$$g(m_i', m_{-i}) R_i(\theta') g(m). \tag{4}$$

Note that (2) and (4) jointly imply that

$$g(m_i', m_{-i}) I_i(\theta') g(m). \tag{5}$$

We show that $R_{S(i)}(\theta) \neq R_{S(i)}(\theta')$. Assume, to the contrary, that

$$\text{for all } h \in H : R_{S(h)}(\theta) = R_{S(h)}(\theta'). \tag{6}$$

Definition 1 implies that

$$\text{for all } h \in H : T_h^{\Gamma}(\theta; S(h)) = T_h^{\Gamma}(\theta'; S(h)). \tag{7}$$

From (3) and (7), it follows that

$$m_i \notin T_i^{\Gamma}(\theta; S(i)) \text{ and } m_i' \in T_i^{\Gamma}(\theta; S(i)). \tag{8}$$

Furthermore, given that $i \in S(i)$, by definition of an individual honesty standard, (5) and (6) jointly imply that

$$g(m_i', m_{-i}) I_i(\theta) g(m). \tag{9}$$

Given (8) and (9) and the fact that $i \in H$, Definition 2 implies that $(m_i', m_{-i}) \succ_i^{\Gamma,\theta,S(i)} m$, which is a contradiction of the fact that $m \in NE\left(\Gamma, \succsim^{\Gamma,\theta,S(N),H}\right)$. Thus, F is partial-honesty monotonic given the honesty standard $S(N)$. ■

4 Equivalence result

The classic paper on Nash implementation theory is Maskin (1999), which shows that where the mechanism designer faces a society involving at least three individuals, an SCR is Nash implementable if it is monotonic and satisfies the auxiliary condition of no veto-power.[2]

The condition of no veto-power says that if an outcome is at the top of the preferences of all individuals but possibly one, then it should be chosen irrespective of the preferences of the remaining individual: that individual cannot veto it. Formally:

2. Moore and Repullo (1990), Dutta and Sen (1991), Sjöström (1991) and Lombardi and Yoshihara (2013) refined Maskin's theorem by providing necessary and sufficient conditions for an SCR to be implementable in (pure strategies) Nash equilibrium. For an introduction to the theory of implementation see Jackson (2001), Maskin and Sjöström (2002) and Serrano (2004).

Definition 7. The SCR $F : \Theta \twoheadrightarrow X$ satisfies *no veto-power* provided that for all $\theta \in \Theta$ and all $x \in X$, if there exists $i \in N$ such that for all $j \in N \backslash \{i\}$ and all $x' \in X : xR_j(\theta)x'$, then $x \in F(\theta)$.

Theorem 2 (Maskin's theorem, 1999). If $n \geq 3$ and $F : \Theta \twoheadrightarrow X$ is an SCR satisfying Maskin monotonicity and no veto-power, then it is Nash implementable.

In a general environment such as that considered here, a seminal paper on Nash implementation problems involving partially-honest individuals is Dutta and Sen (2012). It shows that for Nash implementation problems involving at least three individuals and in which there is at least one partially-honest individual, the Nash implementability is assured by no veto-power. From the perspective of this paper, Dutta-Sen's theorem can be formally restated as follows:

Theorem 3 (Dutta-Sen's theorem, 2012). Let Assumption 1 and Assumption 2 be given. Let the honesty standard of society be summarized in $S(N)$, where $S(i) = N$ for all $i \in N$. If $n \geq 3$ and $F : \Theta \twoheadrightarrow X$ is an SCR satisfying partial-honesty monotonicity for the standard $S(N)$ and no veto-power, then it is partially-honestly Nash implementable.

As already noted in the previous section, any SCR is partial-honesty monotonic whenever the honesty standard of society is such that every individual considers truthful only messages that encode the whole truth about preferences of individuals in society, that is, $S(i) = N$ for all $i \in N$.

That is a particular kind of honesty standards of individuals but there is no reason to restrict attention to such standards. Indeed, as per Ariely (2012)'s findings, individuals may display honesty standards which allow them to lie or cheat a little without that being harmful to their views of themselves as honest persons. On this basis, we have presented an implementation model which is able to handle such views of honesty and presented a necessary condition for Nash implementation problems involving partially-honest individuals. In this section, we are interested in understanding the kind of honesty standards of individuals which would make it impossible for the mechanism designer to circumvent the limitations imposed by Maskin monotonicity.

To this end, let us introduce the following notion of standards of honesty of a society.

Definition 8. Given a society N involving at least two individuals, an honesty standard of this society is said to be *non-connected* if and only if for all $i \in N$, $i \notin S(j)$ for some $j \in N$.

Given that the honesty standard of individual i includes the individual herself, by definition of $S(i)$, the honesty standard of society is non-connected whenever everyone of its members is excluded from the honesty standard of another member of the society. Simply put, members of society do not concern themselves with the same individual.

It is self-evident that the kind of honesty standards in Dutta-Sen's theorem are not non-connected, because every individual of the society is interested in telling the truth about the whole society. As another example of honesty standards of a society that are not non-connected, consider a three-individual society where individual 1 concerns herself with herself and with individual 2 (that is, $S(1) = \{1, 2\}$), individual 2 concerns herself with everyone (that is, $S(2) = \{1, 2, 3\}$) and, finally, individual 3 concerns herself with herself and with individual 1 (that is, $S(3) = \{1, 3\}$). The honesty standard of this three-individual society is not non-connected because everyone concerns themselves with individual 1.

Moreover, it is not the case that every non-connected honesty standard of society implies that every individual honesty standard be of the form $S(i) \neq N$, as we demonstrate with the next example. Consider a three-individual society where individual 1 is concerned only with herself (that is, $S(1) = \{1\}$), individual 2 with everyone (that is, $S(2) = \{1, 2, 3\}$) and individual 3 with herself and with individual 2 (that is, $S(3) = \{2, 3\}$). The honesty standard of this society is non-connected given that individual 2 and individual 3 are both excluded from the honesty standard of individual 1 and individual 1 is excluded from the honesty standard of individual 3.

As is the case here, the above definition is a requirement for the honesty standard of a society that is sufficient for partial-honesty monotonicity to be equivalent to Maskin monotonicity when two further assumptions are satisfied. The first assumption requires that the family $\mathcal{H}$ include singletons. This requirement is innocuous given that the mechanism designer cannot exclude any individual from being partially-honest purely on the basis of Assumption 1.

The second requirement is that the set of states Θ takes the structure of the Cartesian product of allowable independent characteristics for individuals. More formally, the domain Θ is said to be independent if it takes the form

$$\Theta = \prod_{i \in N} \Theta_i,$$

where Θ_i is the domain of allowable independent characteristics for individual i, with θ_i as a typical element. Thus, in a state $\theta = (\theta_i)_{i \in N}$, individual i's ordering $R_i(\theta_i)$ over X depends solely on individual i's independent characteristic θ_i rather than on the profile θ. By abuse of notation, write $R_i(\theta)$ for $R_i(\theta_i)$ in state θ. Given that a characteristic of individual i is independent from those of other individuals, the equivalence result does not hold for the correlated values case.

The latter two requirements and the requirement that the honesty standard of society needs to be non-connected are jointly sufficient for partial-honesty monotonicity to imply Maskin monotonicity. Each of those requirements is indispensable, and this can be seen as follows:

Consider a two-individual society where Θ is the set of states and X is the set of outcomes available to individuals. Let $S(i)$ be the honesty standard of individual $i = 1, 2$. Consider an outcome x and a state θ such that x is an F-optimal outcome at θ. Consider any other state θ' such that individuals' preferences change in a Maskin monotonic way around x from θ to θ'. Maskin monotonicity says that x must continue to be an F-optimal outcome at θ'. To avoid trivialities, let us focus on the case that $\theta \neq \theta'$, which means that $R_N(\theta) \neq R_N(\theta')$, given that we identify states with preference profiles.

If every individual were concerned with the whole society, we could never invoke (the contrapositive of) partial-honesty monotonicity to conclude that x should remain F-optimal at θ' because $R_N(\theta) \neq R_N(\theta')$. Furthermore, consider the case that individual 1 concerns herself with only herself, that is, $S(1) = \{1\}$, while individual 2 with the whole society, that is, $S(2) = \{1, 2\}$. Reasoning such as the one just used shows that partial-honesty monotonicity cannot be invoked if $R_1(\theta) \neq R_1(\theta')$. The argument for honesty standards of the form $S(1) = \{1, 2\}$ and $S(2) = \{2\}$ is symmetric. Thus, the only case left to be considered is the one in which everyone concerns themselves with only themselves, that is, $S(i) = \{i\}$ for $i = 1, 2$. In this situation, the honesty standard of society is reduced to the non-connected one. Note that standards considered earlier were not.

Suppose that preferences of individual 1 are identical in the two states, that is, $R_1(\theta) = R_1(\theta')$. To conclude that x should be F-optimal at θ' by invoking partial-honesty monotonicity we need to find individual 1 in the family $\mathcal{H}$. The argument for the case $R_2(\theta) = R_2(\theta')$ is symmetric. Thus, if $R_i(\theta) = R_i(\theta')$ for one of the individuals, the requirement that the singleton $\{i\}$ is an element of $\mathcal{H}$ is needed for the completion of the argument.

Suppose that preferences of individuals are not the same in the two states, that is, $R_i(\theta) \neq R_i(\theta')$ for every individual i, though they have changed in a Maskin monotonic way around x from the state θ to θ'. In this case, one cannot directly reach the conclusion of Maskin monotonicity by invoking partial-honesty monotonicity. One way to circumvent the problem is to be able to find a feasible state θ'' with the following properties: i) individuals' preferences change in a Maskin monotonic way around x from θ to θ'' and $R_i(\theta) = R_i(\theta'')$ for an individual i, and ii) individuals' preferences change in that way around x from θ'' to θ' and $R_j(\theta') = R_j(\theta'')$ for individual $j \neq i$. A domain Θ that assures the existence of a such state is the independent domain.

Even if one were able to find such a state θ'' by requiring an independent product structure of Θ, one could not invoke partial-honesty monotonicity and conclude that x must continue to be an F-optimal outcome at θ' whenever the family $\mathcal{H}$ did not have the appropriate structure. This can be seen as follows:

Suppose that Θ is an independent domain. Then, states take the form of profiles of individuals' characteristics, that is, $\theta = (\theta_1, \theta_2)$ and $\theta' = (\theta'_1, \theta'_2)$. Moreover, the characteristic of individual i in one state is independent from the characteristic of the other individual. That is, $R_i(\theta) = R_i(\theta_i)$ and $R_i(\theta') = R_i(\theta'_i)$ for every individual i. The product structure of Θ assures that the states (θ_1, θ'_2) and (θ'_1, θ_2) are both available and each of them has the properties summarized above.

Next, suppose that the family $\mathcal{H}$ has a structure given by $\{\{1\}, \{1, 2\}\}$. One can invoke partial-honesty monotonicity for $H = \{1\}$ to obtain that x is one of the outcomes chosen by the SCR F at (θ_1, θ'_2) when the state changes from θ to (θ_1, θ'_2), but he cannot conclude that x remains also F-optimal at θ' when it changes from (θ_1, θ'_2) to θ'. The reason is that partial-honesty monotonicity cannot be invoked again for the case $H = \{2\}$ because the structure of the family $\mathcal{H}$ does not contemplate such a case. The argument for the case that $\mathcal{H}$ takes the form $\{\{2\}, \{1, 2\}\}$ is symmetric. Thus, each of our requirements is indispensable, and jointly they lead to the following conclusion:

Theorem 4. Let N be a society involving at least two individuals, Θ be an independent domain and $\mathcal{H}$ include singletons. Suppose that the honesty standard of the society is non-connected. Partial-honesty monotonicity is equivalent to Maskin monotonicity.

Proof. Let $n \geq 2$, Θ be an independent domain and $\mathcal{H}$ include singletons. Let $S(N)$ be a non-connected honesty standard of N. One can see that Maskin monotonicity implies partial-honesty monotonicity.

For the converse, consider any SCR $F : \Theta \twoheadrightarrow X$ satisfying partial-honesty monotonicity. Consider any $x \in X$ and any state $\theta \in \Theta$ such that x is an F-optimal outcome at θ. Moreover, consider any state θ' such that individuals' preferences change in a Maskin monotonic way around x from θ to θ', that is,

$$\text{for all } i \in N \text{ and all } x' \in X : xR_i(\theta)\, x' \implies xR_i(\theta')\, x'.$$

We show that x remains F-optimal at θ'.

If characteristics of individuals in the honesty standard of individual $i \in N$ are identical in the two states, that is, $R_{S(i)}(\theta) = R_{S(i)}(\theta')$, partial-honesty monotonicity for the case $H = \{i\}$ assures that x is still F-optimal at θ'. Thus, let us consider the case $R_{S(i)}(\theta) \neq R_{S(i)}(\theta')$ for every individual $i \in N$.

To economize notation, for any subset K of N, write K_C for the complement of K in N. Therefore, for any non-empty subset K of N, we can write any non-trivial combination of the states θ and θ' as $(\theta_K, \theta'_{K_C})$, where it is understood that θ_K is a list of characteristics of individuals in K at the state θ and θ'_{K_C} is a list of characteristics of individuals in K_C at θ'. Note that any state that results by that combination is available in Θ because of its product structure.

Given that the honesty standard of society is non-connected, there must be an individual $j(1) \in N$ who does not concern herself with the whole society, that is, $S(j(1)) \neq N$. Consider the state

$$\left(\theta_{K(1)}, \theta'_{K(1)_C}\right) \text{ where } K(1) \equiv S(j(1)),$$

and call it θ^1. By construction, individuals' preferences change in a Maskin monotonic way around x from θ to θ^1 and, moreover, $\theta_{K(1)} = \theta^1_{K(1)}$. Partial-honesty monotonicity for the case $H = \{j(1)\}$ assures that the x remains an F-optimal outcome at θ^1.

If there is an individual $i \in N \backslash \{j(1)\}$ who is not concerned with any of the individuals in the honesty standard of individual $j(1)$, that is, the intersection $S(i) \cap S(j(1))$ is empty, then partial-honesty monotonicity for the case $H = \{i\}$ assures that x is still F-optimal at θ'. This is because, by construction, individuals' preferences change in a Maskin monotonic way around x from θ^1 to θ' and $\theta^1_{S(i)} = \theta'_{S(i)}$.

Thus, consider any individual $j(2) \in N \backslash \{j(1)\}$, and denote by $K(2)$ the set of individuals who jointly concern individual $j(1)$ and individual $j(2)$ according to their individual honesty standards. Furthermore, consider the state

$$\left(\theta_{K(2)}, \theta'_{K(2)_C}\right) \text{ where } K(2) \equiv K(1) \cap S(j(2)),$$

and call it θ^2. By construction, individuals' preferences change in a Maskin monotonic way around x from θ^1 to θ^2 and, moreover, $\theta^1_{S(j(2))} = \theta^2_{S(j(2))}$. Partial-honesty monotonicity for the case $H = \{j(2)\}$ assures that x remains an F-optimal outcome at θ^2.

If there is an individual $i \in N \backslash \{j(1), j(2)\}$ who is not concerned with any of the individuals with whom individuals $j(1)$ and $j(2)$ are jointly concerned, partial-honesty monotonicity for the case $H = \{i\}$ assures that x is also F-optimal at θ'. This is because, by construction, individuals' preferences change in a Maskin monotonic way around x from θ^2 to θ' and $\theta^2_{S(i)} = \theta'_{S(i)}$.

Thus, consider any individual $j(3) \in N \backslash \{j(1), j(2)\}$, and denote by $K(3)$ the set of individuals that jointly concern individuals $j(1)$, $j(2)$ and $j(3)$ according to their individual honesty standards. Furthermore, consider the state

$$\left(\theta_{K(3)}, \theta'_{K(3)_C}\right) \text{ where } K(3) \equiv K(2) \cap S(j(3)),$$

and call it θ^3. By construction, individuals' preferences change in a Maskin monotonic way around x from θ^2 to θ^3 and, moreover, $\theta^2_{S(j(3))} = \theta^3_{S(j(3))}$. Partial-honesty monotonicity for the case $H = \{j(3)\}$ assures that x remains an F-optimal outcome at θ^3.

As above, if there is an individual $i \in N \backslash \{j(1), j(2), j(3)\}$ who is not concerned with any of the individuals with whom individuals $j(1)$, $j(2)$ and $j(3)$ are jointly concerned, partial-honesty monotonicity for the case $H = \{i\}$ assures that x remains also F-optimal at θ', because, by construction, individuals' preferences change in a Maskin monotonic way around x from θ^3 to θ' and $\theta^3_{S(i)} = \theta'_{S(i)}$. And so on.

Since the society N is a finite set and the above iterative reasoning is based on its cardinality, we are left to show that it must stop at most after $n-1$ iterations.

To this end, suppose that we have reached the start of the $n-1$th iteration. Thus, consider any individual $j(n-1) \in N$, with $j(n-1) \neq j(r)$ for $r = 1, \cdots, n-2$, and denote by $K(n-1)$ the set of individuals that jointly concern individuals $j(1)$, $j(2), \cdots, j(n-2)$ and $j(n-1)$ according to their individual honesty standards. Furthermore, consider the state

$$\left(\theta_{K(n-1)}, \theta'_{K(n-1)_C}\right) \text{ where } K(n-1) \equiv K(n-2) \cap S(j(n-1)),$$

and call it θ^{n-1}. As above, by construction, individuals' preferences change in a Maskin monotonic way around x from $\theta^{n-2} \equiv \left(\theta_{K(n-2)}, \theta'_{K(n-2)_C}\right)$ to θ^{n-1} and, moreover, $\theta^{n-2}_{S(j(n-1))} = \theta^{n-1}_{S(j(n-1))}$. Partial-honesty monotonicity for the case $H = \{j(n-1)\}$ assures that x is an F-optimal outcome at θ^{n-1}.

At this stage there is only one individual in N who is left to be considered. Call her $j(n)$. Suppose that this individual is concerned with one of the individuals for whom individuals $j(1)$, $j(2), \cdots, j(n-2)$ and $j(n-1)$ are jointly concerned. In other words, suppose that the intersection $K(n-1) \cap S(j(n))$ is non-empty. Then, the whole society concerns itself with one of its member, and this contradicts the fact that the honesty standard of society is non-connected. Therefore, it must be the case that individual $j(n)$ is not concerned with any of the individuals with whom individuals $j(1)$, $j(2), \cdots, j(n-2)$ and $j(n-1)$ are jointly concerned according to their individual honesty standards. Partial-honesty monotonicity for the case $H = \{j(n)\}$ assures that x remains also F-optimal at θ' given that, by construction, individuals' preferences change in a Maskin monotonic way around x from θ^{n-1} to θ' and $\theta^{n-1}_{S(j(n))} = \theta'_{S(j(n))}$.

The iterative reasoning would stop at the rth $(< n-1)$ iteration if there were an individual $i \in N \backslash \{j(1), \cdots, j(r)\}$ who did not concern itself with any of the individuals in $K(r)$, that is, if the intersection $S(i) \cap K(r)$ were empty. If that were the case, then the desired conclusion could be obtained by invoking partial-honesty monotonicity for $H = \{i\}$ because, by construction, it would hold that individuals' preferences change in a Maskin monotonic way around x from θ^r to θ' and that $\theta^r_{S(i)} = \theta'_{S(i)}$. ■

In light of Theorem 1 and Maskin's theorem, the main implications of the above conclusion can be formally stated as follows:

Corollary 1. Let N be a society involving at least two individuals, Θ be an independent domain and $\mathcal{H}$ include singletons. Suppose that the honesty standard of the society is non-connected. Let Assumption 1 and Assumption 2 be given. The SCR $F : \Theta \twoheadrightarrow X$ is Maskin monotonic if it is partially-honestly Nash implementable.

Corollary 2. Let N be a society involving at least three individuals, Θ be an independent domain and $\mathcal{H}$ include singletons. Suppose that the honesty standard of the society is non-connected. Let Assumption 1 and Assumption 2 be given. Any SCR $F : \Theta \twoheadrightarrow X$ satisfying no veto-power is partially-honestly Nash implementable if and only if it is Maskin monotonic.

5 Concluding remarks

The assumption that the mechanism designer knows the honesty standard of society is often not met in reality, although it may be plausible in societies with a small number of individuals in which the mechanism designer knows their sensitivity to honesty. Outside of cases like those, we view as more plausible the assumption that the mechanism designer only knows the type of honesty standards shared by individuals. Does the conclusion change in this case? The answer is no. After all, if individuals are honesty-sensitive, the mechanism designer can test for connectedness of their honesty standards. If the test fails, it would be vein for him to attempt to Nash implement any SCR that is not Maskin monotonic. The reason for it is easy to identify: the fact that he solely knows that the honesty standard of society is non-connected can only make implementation harder than if the actual non-connected honesty standards of participants were known.

In an environment in which knowledge is dispersed, how individuals will interact with the mechanism designer is a natural starting point when it comes to Nash implement an SCR. A particular kind of communication is, as we do in this paper, to ask participants to report individuals' preferences. However, there is no reason to restrict attention to such schemes.

A simpler way to go about it is to ask individuals to report only their own preferences. An obvious advantage of such a type of communication is that the mechanism designer does not need to understand the psychological motivations of the individuals, beyond a basic self-interest. However, if the mechanism designer structured the communication in that way, he would then force individuals to behave *as if* their honesty standards were non-connected, though their non-connectedness could be merely an artifact of that communication structure. The reason is that individuals cannot articulate the communication according to their actual honesty standards. And, from the perspective of this paper, that would impair the ability of the mechanism designer to escape the limitations imposed by Maskin monotonicity. Not surprisingly, in an independent domain of strict preferences, Saporiti (2014) shows that any social choice function that can be securely implemented is Maskin monotonic, though all participants are partially-honest.[3]

There are multiple other ways for the mechanism designer to structure the exchange of information with individuals, and there is no limit to how imaginative he can be. However, this paper offers this guidance on how to go about it in environments involving partially-honest individuals: If the honesty standards of participants are connected, the informational requirements need not force individuals to behave *as if* their honesty standards were not.

Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1989) and Jackson (1991) have shown that Maskin's theorem can be generalized to Bayesian environ-

3. Secure implementation is implementation in Nash equilibrium and in dominant strategies. Recall that in an independent domain of strict preferences, strategy-proofness implies Maskin monotonicity (Dasgupta et al., 1979).

ments. A necessary condition for Bayesian Nash implementation is Bayesian monotonicity. In a Bayesian environment involving at least three individuals, Bayesian monotonicity combined with no veto-power is sufficient for Bayesian Nash implementation provided that a necessary condition called closure and the Bayesian incentive compatibility condition are satisfied (Jackson, 1991). Although the implementation model developed in this paper needs to be modified to handle Bayesian environments, we believe a similar equivalence result holds in those environments for suitably defined non-connected honesty standards. This subject is left for future research.

Based upon the evidence that people regard themselves as honest even though they lie or cheat a little, as Ariely (2012) suggests, we identified testable conditions for Nash implementation with partially-honest individuals which, if satisfied, send us back to the limitations imposed by Maskin's theorem. Thus, the exploration of the possibilities offered by that implementation needs to move away from those properties. As yet, where the exact boundaries of those possibilities lay for general environments and economic environments is far from known.[4]

References

D. Abreu, A. Sen, Subgame perfect implementation: a necessary and almost sufficient condition, J. Econ. Theory 50 (1990) 285-299.

D. Abreu, A. Sen, Virtual implementation in Nash equilibrium, Econometrica 59 (1991) 997–1021.

D. Ariely, The (honest) truth about dishonesty, first edition, HarperCollins Publishers, New York, 2012.

P. Dasgupta, P. Hammond, E. Maskin, The implementation of social choice rules: Some general results on incentive compatibility, Rev. Econ. Stud. 46 (1979) 185-216.

B. Dutta, A. Sen, A necessary and sufficient condition for two-person Nash implementation, Rev. Econ. Stud. 58 (1991) 121-128.

B. Dutta, A. Sen, Nash implementation with partially honest individuals, Games Econ Behav 74 (2012) 154-169.

M.O. Jackson, Bayesian implementation, Econometrica 59 (1991) 461-477.

M.O. Jackson, Implementation in Undominated Strategies: A Look at Bounded Mechanisms, Rev. Econ. Stud. 59 (1992) 757-775.

4. An exception is represented by Lombardi and Yoshihara (2014), in which problems in pure exchange economies by price-quantity mechanisms are studied.

M.O. Jackson, A crash course in implementation theory, Soc. Choice Welfare 18 (2001) 655-708.

E. Kalai, J.O. Ledyard, Repeated Implementation, J. Econ. Theory 83 (1998) 308-317.

N. Kartik, O. Tercieux, Implementation with evidence, Theor. Econ. 7 (2012) 323-355.

N. Kartik, O. Tercieux, R. Holden, Simple mechanism and preference for honesty, Games Econ. Behav. 83 (2014) 284-290.

J. Lee, H. Sabourian, Efficient Repeated Implementation, Econometrica 79 (2011) 1967-1994.

M. Lombardi, N. Yoshihara, A full characterization of Nash implementation with strategy space reduction, Econ. Theory 54 (2013) 131-151.

M. Lombardi, N. Yoshihara, Natural implementation with partially-honest agents in economic environments with free-disposal, IER discussion paper series A (2014) n. 616, Hitotsubashi University.

E. Maskin, Nash equilibrium and welfare optimality, Rev. Econ. Stud. 66 (1999) 23-38.

E. Maskin, T. Sjöström, Implementation theory, in: K. Arrow, A.K. Sen, K. Suzumura (Eds), Handbook of Social Choice and Welfare, Elsevier Science, Amsterdam, 2002, pp. 237–288.

H. Matsushima, A new approach to the implementation problem, J. Econ. Theory 45 (1988) 128–144.

H. Matsushima, Behavioral aspects of implementation theory, Econ. Letters 100 (2008a) 161-164.

H. Matsushima, Role of honesty in full implementation, J. Econ. Theory 139 (2008b) 353-359.

J. Moore, R. Repullo, Subgame perfect implementation, Econometrica 56 (1988) 1191-1220.

J. Moore, R. Repullo, Nash implementation: A full characterization, Econometrica 58 (1990) 1083-1100.

J. Ortner, Direct implementation with partially honest individuals, Games Econ. Behav. 90 (2015) 1-16.

T. Palfrey, S. Srivastava Implementation with incomplete information in exchange economies, Econometrica 57 (1989) 115-134.

T. Palfrey, S. Srivastava, Nash-implementation using undominated strategies, Econometrica 59 (1991) 479-501.

A. Postlewaite, D. Schmeidler, Implementation in differential information economies, J. Econ. Theory 39 (1986) 14-33.

L. Renou, C. Mezzetti, Repeated Nash Implementation (June 29, 2012). Available at SSRN: http://ssrn.com/abstract=2096184

A. Saporiti, Securely implementable social choice rules with partially honest agents, J. Econ. Theory 154 (2014) 216-228.

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

T. Sjöström, On the necessary and sufficient conditions for Nash implementation, Soc. Choice Welfare 8 (1991) 333-340.