Implementation with Interdependent Payoffs*

Boaz Zik†

Abstract

We consider the problem of implementation in a model with agents who have interdependent payoffs. We show that in such a model, under mild restrictions on the behavior of the decision rules and the structure of the valuation functions, ex-post implementation is impossible. Given profiles of valuation functions and distributions of signals, the set of Bayes–Nash implementable decision rules in any interdependent payoffs setup is equal to the set of Bayes–Nash implementable decision rules in the independent private values setup. For each decision rule in this set we construct a transfer scheme that implements it in a Bayes–Nash equilibrium in the independent private values setup and in every interdependent payoffs setup. (Keywords: *Mechanism design; Social preferences; Ex-post implementation; Bayesian implementation.*)

1 Introduction

Models of mechanism design usually consider selfish agents, that is, agents whose utilities consist of their own personal payoffs. However, it is well established that in many economic environments subjects often have "other-regarding preferences." In

*I would like to thank Alex Gershkov, Ilan Kremer, Motty Perry, and seminar participants at the Hebrew University for their valuable comments.

†Department of Economics and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem.

these environments agents' utilities depend not only on their own personal payoff but also on the payoffs of other agents in the society. In this paper, we study a problem of mechanism design in such environments. Our study focuses on the implementation of decision rules that depend on information about the agents' personal payoffs and ignore information about their social preferences. This study is motivated, among other things, by the following economic scenarios.

The first scenario concerns an agency problem in a conglomerate. A conglomerate is a collection of independent corporations, engaged in different business ventures, that function as a single economic entity under the control of a central administration. The problem, introduced by Groves (1973), is as follows. The conglomerate's central administration needs to choose an alternative from a set of possible alternatives. The central administration's payoff from an alternative depends on the effect this alternative has on the profits of the conglomerate's corporations (for example, the central administration may want to maximize the sum of the profits of the conglomerate's corporations). The effect of each alternative on a corporation's profit is the private information of the corporation's manager. Therefore, to make an optimal decision, the central administration must elicit from each manager information about the effect of each alternative on her corporation's profit. Our study investigates whether it is possible to elicit this information in environments where managers' utilities depend not only on the profits of their corporations but also on the profits of other corporations in the conglomerate. Such dependency may occur, for example, when a manager is a shareholder in the conglomerate and, therefore, profits from its success; when a manager is rewarded according to the relative success of her corporation with respect to the other corporations in the conglomerate; when a manager is connected in some way (say, through family, friendship, or business ties) to other managers in the conglomerate; or when a manager is invested in some other corporation of the conglomerate.

The second scenario concerns a utilitarian designer who is called to choose a social alternative. Consider a society some of whose members may have antisocial preferences, such as envy, spite, and so on.[1] In such a scenario utilitarian theory suggests

[1]There is evidence in the experimental economics literature that subjects often have such "other-

that these preferences will be "laundered"; i.e., that the antisocial aspects in these preferences will be removed before the preferences are incorporated into the social utility.[2] Harsanyi, one of the greatest advocates of utilitarian theory, suggests that:

> Some preferences . . . must be altogether excluded from our social-utility function. In particular we must exclude all clearly antisocial preferences such as sadism, envy, resentment and malice. . . . Utilitarian ethics makes all of us members of the same moral community. A person displaying ill will toward others does remain a member of this community, but not with his whole personality. That part of his personality that harbors these hostile antisocial feelings must be excluded from membership, and has no claim to a hearing when it comes to defining our concept of social utility (Harsanyi 1977, pp. 647)

Blanchet and Fleurbaey (2006) suggest that even altruistic preferences should be "laundered" since they can lead to rewarding the selfish. Laundering preferences means that when the designer is called to choose the social alternative, he should consider only information about agents' personal gains and disregard information about agents' social preferences. Our research question investigates whether the designer can launder preferences when the information about the parameters of each agent's utility function is privately held by the agent.

We consider environments of agents whose preferences are quasilinear in money, whose valuations are private, and whose utilities depend on the payoffs of other agents.[3] Such environments are different from environments of interdependent values in the following respect. In environments of interdependent values, an agent's utility is affected by the mechanism through the decision rule and his personal transfer. In environments with interdependent payoffs, however, an agent's utility is affected by the mechanism through the decision rule, his personal transfer, and the personal transfers of the agents whose payoffs affect his utility.[4] That is, in environments of interde-

regarding" preferences. See Cooper and Kagel (2009) for a survey.

[2]See, for example, Harsanyi (1977), Goodin (1986), and Blanchet and Fleurbaey (2006).

[3]The dependency of an agent's utility on the payoffs of other agents is a function of a signal that is privately known to the agent.

[4]Note that the effect of other agents' transfers on an agent's utility, unlike the effect of his personal

pendent payoffs mechanisms affect agents' preferences in a more diversified manner, compared to in environments of interdependent values. The property that agents' utilities depend on the transfers of other agents affects the ability of the designer to achieve implementation in two ways. On the one hand, it provides the designer with more tools to incentivize an agent to report truthfully and to achieve implementation. On the other hand, since each agent's transfer affects the incentives of various agents, constructing transfers that incentivize one agent to report truthfully may impair the incentives of other agents to report truthfully.

We find that the possibility of implementation in environments with payoffs dependencies heavily depends on the solution concept that is used for the implementation. We call a setup with agents whose utilities depend on the payoffs of other agents an *interdependent payoffs* setup and we show the following results:

- Under mild and economically reasonable conditions on the properties of the decision rules and the structure of the valuation functions, ex-post implementation is impossible in an interdependent payoffs setup.

- For any given profiles of valuation functions and distribution functions, the set of decision rules that can be implemented in a Bayes–Nash equilibrium in any interdependent payoffs setup is identical to the set of decision rules that can be implemented in a Bayes–Nash equilibrium in the independent private values setup.

- For each decision rule in the above set, there exists a transfer scheme that implements it in a Bayes–Nash equilibrium in every interdependent payoffs setup as well as in the independent private values setup.

In the interdependent payoffs model we must take into account the immense difference between ex-post implementation and Bayesian implementation. While ex-post implementation is virtually impossible, Bayesian implementation allows for the implementation of decision rules that satisfy certain monotonicity conditions and, in

transfer on his utility, depends on the realization of signals.

particular, of the efficient decision rule.[5] In models of independent private values we do not see such a difference. For example, Gershkov et al. (2013) present an equivalency result between dominant strategy and Bayesian implementation when signals are one-dimensional. In models of interdependent values such a difference does not appear either. In these models positive results on efficient ex-post implementation have been shown in the case where signals are one-dimensional; see, for example, Dasgupta and Maskin (2000) and Perry and Reny (2002). On the other hand, negative results on Bayesian efficient implementation have been presented in the case where signals are multidimensional; see Jehiel and Moldovanu (2001). The reason for the difference in the implementation power of these solution concepts in the interdependent payoffs model is as follows. In the model an agent's utility depends on other agents' payoffs, namely, on other agents' valuations and transfers. This means that an agent's transfer affects not only the incentive of the agent who receives the transfer to report truthfully but also the incentives of other agents, whose utilities depend on this transfer, to report truthfully. We show that when an agent's utility depends on other agents' payoffs a necessary condition for implementation is that the agent's report does not affect the payoffs of these agents. This means that when other agents' utilities depend on an agent's payoff the agent's transfer should eliminate the effect of these agents' reports on his valuation. In addition, the agent's transfer must also incentivize the agent himself to report truthfully. When we consider ex-post implementation these requirements for an agent's transfer must be satisfied for every realization of signals. We show that this cannot happen without contradictions and hence ex-post implementation is impossible. However, when we consider Bayesian implementation these requirements should only be met in expectation. We show that in this case it is possible to construct transfer schemes that satisfy these requirements. Hence, Bayesian implementation is possible.

Our impossibility result on ex-post implementation in the interdependent payoffs model joins several other impossibility results on implementation by robust solution concepts in the literature. In environments of private values and unrestricted preferences, Gibbard (1973) and Satterthwaite (1975) show that if the cardinality of the set

[5]We consider the "efficient decision rule" to be the decision rule that maximizes the sum of the agents' payoffs.

of social alternatives is greater than or equal to three, then only dictatorial decision rules are implementable in dominant strategies. Most of the literature on implementation, however, focuses on environments with quasilinear preferences. In such environments an agent's utility is affected by his personal transfer in an additive manner independently of the realization of signals. The designer can use these personal transfers to assist him in aligning agents' preferences with social preferences. The analysis of robust implementation in these environments provides positive results both in the case of private values and in the case of interdependent values and single-dimensional signals. In the case of interdependent values and multidimensional signals, however, Jehiel et al. (2006) show that for generic valuation functions only constant decision rules are ex-post implementable.[6] In this paper, we consider environments of interdependent payoffs, in which an agent's personal transfer affects not only the preferences of the agent who receives the transfer but also the preferences of other agents whose utilities depend on this agent's payoff. On the one hand, this provides the designer with more ways to align agents' preferences with social preferences with respect to the standard quasilinear environment. On the other hand, since an agent's transfer affects the incentives of other agents, this property is also confining. Our result shows that ultimately environments of interdependent payoffs do not allow for robust implementation.

There are a number of other papers that analyze mechanism design problems in models with social preferences. Desiraju and Sapington (2007) consider a screening problem of a monopsonistic firm facing two potential workers who are inequity averse. They show that if the two workers are identical ex-ante, then workers' social preferences are not constraining and the firm can achieve the same expected payoff as in the case where workers are selfish. They also show that the converse is generally true. Siemens (2011) considers the monopsonist's screening problem in a model with a continuum of workers where a proportion of them are averse to inequity. He shows that the presence of inequity-averse workers distorts the firm's production choice. In addition, it may also lead to the exclusion of workers who are both inequity averse and have

[6] It is worth noting that there are important environments that are negligible in Jehiel et al.'s (2006) setting in which implementation of non-constant decision rules is possible, e.g., environments of private goods; see Bikhchandani (2006).

low abilities. Bierbrauer and Netzer (2016) consider the problem of implementation in a model with intention-based preferences. They show that every decision rule that is Bayesian implementable in the case where agents are selfish is also implementable in their model. Moreover, they show that the existence of social preferences can be used to reconcile efficiency, incentive compatibility, and individual rationality. Bartling and Netzer (2016) investigate the trade-off between belief-robust implementation and externality-robust implementation. They examine participants' behavior both in the second-price auction, which is dominant-strategy implementable but is not robust to the existence of social preferences, and in its externality-robust counterpart, which is robust to the existence of social preferences but is only Bayesian implementable. They find that participants overbid in the second-price auction, but that average bids equal value in the externality-robust auction. This result suggests that participants do take into account the externalities of their actions on other participants. In addition, they find that both auctions produce the same level of efficiency. This result suggests that the two notions of robustness are equally important from an efficiency perspective. The above papers focus on Bayesian implementation, while in this paper we concentrate on ex-post implementation. The analysis of ex-post implementation provides insight into whether it is possible to satisfy the robustness criteria that appear in the "Wilson critique" in models with social preferences; see Wilson (1987). In particular, our result on ex-post implementation shows that achieving robustness both in the dimension of beliefs and in the dimension of payoff externalities is impossible.

The rest of the paper is organized as follows. In Section 2 we present the model. In Section 3 we discuss the notion of ex-post implementation and present an impossibility result. In Section 4 we discuss the notion of Bayes–Nash implementation. We characterize the set of Bayes–Nash implementable decision rules and construct a transfer scheme that implements a decision rule that belongs to this set in every interdependent payoffs setup as well as in the independent private values setup. In Section 5 we discuss another interpretation of the model, the case of interdependent utilities. Section 6 concludes. Proofs are relegated to the Appendix.

2 The Model

Let $I = \{1, ..., n\}$ be the set of agents. There is a finite set, A, of social alternatives from which the designer has to choose an alternative. Each agent $i \in I$ receives a signal $\theta_i \in \Theta_i$, where Θ_i is a measurable space. The signal θ_i is drawn from the set Θ_i according to the density function f_i, where $f_i(\theta_i) > 0$ for every $\theta_i \in \Theta_i$. If alternative a is chosen, if the signal realization is θ_i, and if agent i obtains a transfer t_i, then agent i's payoff is given by $\Pi_i = v_i(a, \theta_i) + t_i$. The utility of agent i depends in a linear manner on her personal payoff and on the payoffs of some (possibly empty) subset of the other agents, i.e., $u_i = \Pi_i + \sum_{k \in P_i} \delta_i^k \cdot \Pi_k$, where $P_i \subset I \setminus \{i\}$ and $\delta_i^k \in \left[\underline{\delta_i^k}, \overline{\delta_i^k}\right] \subset \mathbb{R}$ with $\underline{\delta_i^k} < \overline{\delta_i^k}$. The vector of coefficients $\left(\delta_i^k\right)_{k \in P_i} := \delta_i$ is drawn from the set $\mathcal{D}_i := \times_{k \in P_i} \left[\underline{\delta_i^k}, \overline{\delta_i^k}\right]$ according to the density function g_i, where $g_i(\delta_i) > 0$ for every $\delta_i \in \mathcal{D}_i$. The signals θ_i and δ_i are the private information of agent i, they are drawn independently of each other and of other agents' signals, and their distributions are common knowledge.[7] A *setup* is a 3-tuple consisting of a profile of sets of agents $(P_i)_{i \in I}$, a profile of sets of coefficients $(\mathcal{D}_i)_{i \in I}$, and a profile of density functions on these sets of coefficients $(g_i)_{i \in I}$. Each setup characterizes the structure of the payoffs dependencies in the model. We define R to be the set of all possible setups[8]

$$R := \left\{ \left\langle (P_i)_{i \in I}, (\mathcal{D}_i)_{i \in I}, (g_i)_{i \in I} \right\rangle \mid P_i \subset I \setminus \{i\}, \mathcal{D}_i \in M(P_i),\ g_i \in L(\mathcal{D}_i) \right\}$$

We call the setup in R that satisfies $\forall i \in I\ P_i = \emptyset$ the *independent private values setup* (PV). We define a set $P := R \setminus PV$ and call an element in the set P an *interdependent payoffs setup* (IP). We denote $\Theta := \times_{i \in I} \Theta_i$ with generic element θ, and $\Theta_{-i} := \times_{k \in I \setminus \{i\}} \Theta_k$ with generic element θ_{-i}. We denote $\mathcal{D} := \times_{i \in I} \mathcal{D}_i$ with generic element δ, and $D_{-i} := \times_{k \in I \setminus \{i\}} \mathcal{D}_k$ with generic element δ_{-i}. A function $q : \Theta \to A$ is called a *decision rule.* A *social choice function* is a function $s(\theta, \delta) = (q(\theta), t_1(\theta, \delta), ..., t_n(\theta, \delta))$, where $q(\theta) \in A$ and $t_i(\theta, \delta) \in \mathbb{R}$ for every $i \in I$.

[7] We allow the existence of social preferences to be common knowledge. The intensity of these preferences, however, is private knowledge.

[8] where $M(P_i)$ is the set of all the possible sets of the form $\times_{k \in P_i} \left[\underline{\delta_i^k}, \overline{\delta_i^k}\right]$ with $\underline{\delta_i^k} < \overline{\delta_i^k}$ and $L(\mathcal{D}_i)$ is the set of all strictly positive density functions on $\mathcal{D}_i$.

Remark. The decision rules we consider depend only on information about agents' personal payoffs. However, in our analysis we allow agents' transfers to depend also on information about agents' social preferences. Nonetheless, all the results in the paper would still hold even if we restricted transfers to depend only on information about agents' payoffs.[9]

3 Ex-Post Implementation

We start with a definition ex-post implementation in the context of our model. Ex-post equilibrium requires that the strategy of each agent i be optimal with respect to the strategies of the other agents for every possible realization of signals. By the revelation principle, we can restrict our analysis to direct mechanisms. Consider a given profile of Θ, $(v_i)_{i\in I}$, and an IP setup, where IP∈P. We say that a social choice function $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ is *ex-post implementable in IP* if for every $i \in I$, $\theta \in \Theta$ and $\delta \in \mathcal{D}$ we have

$$(\theta_i, \delta_i) \in \underset{\left(\hat{\theta}_i, \hat{\delta}_i\right) \in \Theta_i \times \mathcal{D}_i}{\arg\max}\ v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right) +$$

$$\sum_{k\in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right)\right)\right]$$

A decision rule $q(\theta)$ is *ex-post implementable in IP* if there exists a profile of real valued functions $(t_1(\theta,\delta), ..., t_n(\theta,\delta))$ such that $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ is ex-post implementable in IP.

Consider a given profile of Θ, $(v_i)_{i\in I}$, and the PV setup, i.e., the setup where $\forall i \in I\ P_i = \emptyset$. We say that a social choice function of the form $(q(\theta), t_1(\theta), ..., t_n(\theta))$ is *ex-post (dominant strategy) implementable in PV* if for every $i \in I$ and $\theta \in \Theta$ we have

$$\theta_i \in \underset{\hat{\theta}_i \in \Theta_i}{\arg\max}\ v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right)\right)$$

[9] The impossibility result holds as a particular case, and the possibility result is based on transfer schemes that depend only on information about agents' payoffs.

A decision rule $q(\theta)$ is *ex-post (dominant strategy) implementable in PV* if there exists a profile of real valued functions $(t_1(\theta), ..., t_n(\theta))$, such that $(q(\theta), t_1(\theta), ..., t_n(\theta))$ is ex-post implementable in PV. Ex-post implementability in IP implies ex-post (dominant strategy) implementability in PV.

Lemma 1. *Consider a given profile of Θ, $(v_i)_{i\in I}$, and an IP setup. If a decision rule $q(\theta)$ is ex-post implementable in IP then it is ex-post (dominant strategy) implementable in PV.*

3.1 The impossibility of ex-post implementation

We now present our main result that is an impossibility result of ex-post implementation in an IP setup. The argument behind this result is the following. Ex-post implementation implies that if $j \in P_i$ then for any two signals θ_i and $\theta_i^{'}$ the payoff of agent j must remain equal on a subset of measure one of the interval $\left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$.[10] Therefore, if the decision rule assign different alternatives for θ_i and $\theta_i^{'}$, and if agent j's valuation is different for each alternative, it is left for agent j's transfer function t_j to eliminate this gap in agent j's payoff. However, t_j also plays a role in incentivizing agent j to report truthfully. We describe conditions under which these two roles of t_j lead to a contradiction and hence make ex-post implementation impossible. These conditions are presented below.[11]

> **Property 1:** There exists an agent $\hat{i}$ and $\hat{j} \in P_{\hat{i}}$ such that for every $k \in P_{\hat{j}}$, $0 \in \left[\underline{\delta_{\hat{j}}^k}, \overline{\delta_{\hat{j}}^k}\right]$. In words, there exists an agent $\hat{i}$ whose utility includes the payoff of some agent $\hat{j}$, and it is possible, based on the common knowledge of all the agents, that agent $\hat{j}$'s utility coincides with her personal payoff.

[10] For any fixed $\left(\theta_{-i}, \delta_{-i}^{-j}\right) \in \Theta_{-i} \times \mathcal{D}_{-i}^{-j}$ where $D_{-i}^{-j} := \mathcal{D} \setminus \left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$.

[11] Property 2 can be replaced by the demand that the signal space and the valuation function satisfy any set of restrictions under which dominant strategy implementation in PV implies revenue equivalence. See Krishna and Maenner (2001) and Heydenreich et al. (2009). Properties 3 and 4 can be replaced by the following weaker demand: there exist a profile $\tilde{\theta}_{-\hat{i}-\hat{j}}$, two alternatives a and b, $\theta_{\hat{i}}^1$, $\theta_{\hat{i}}^2 \in \Theta_{\hat{i}}$, and $\theta_{\hat{j}}^1$, $\theta_{\hat{j}}^2 \in \Theta_{\hat{j}}$, such that $q\left(\theta_{\hat{i}}^1, \theta_{\hat{j}}^1, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = q\left(\theta_{\hat{i}}^1, \theta_{\hat{j}}^2, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = a$ and $q\left(\theta_{\hat{i}}^2, \theta_{\hat{j}}^1, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = q\left(\theta_{\hat{i}}^2, \theta_{\hat{j}}^2, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = b$ and $v_{\hat{j}}\left(a, \theta_{\hat{j}}^1\right) - v_{\hat{j}}\left(a, \theta_{\hat{j}}^2\right) \neq v_{\hat{j}}\left(b, \theta_{\hat{j}}^1\right) - v_{\hat{j}}\left(b, \theta_{\hat{j}}^2\right)$.

Property 2: Θ_i is a convex subset of a finite dimensional Euclidean space, and $v_i(a, \theta_i)$ is a convex function of θ_i for every $i \in I$.

Property 3: There exist a profile $\tilde{\theta}_{-\hat{\imath}}$, two alternatives a and b, and $\theta^1_{\hat{\imath}}$, $\theta^2_{\hat{\imath}} \in \Theta_{\hat{\imath}}$, such that $q\left(\theta^1_{\hat{\imath}}, \tilde{\theta}_{-\hat{\imath}}\right) = a$, $q\left(\theta^2_{\hat{\imath}}, \tilde{\theta}_{-\hat{\imath}}\right) = b$, and there exists $r > 0$ such that for every $\theta_{-i} \in B\left(\left(\tilde{\theta}_{-\hat{\imath}}\right), r\right)$, $q(\theta^1_{\hat{\imath}}, \theta_{-i}) = a$ and $q(\theta^2_{\hat{\imath}}, \theta_{-i}) = b$.[12] In words, agent $\hat{\imath}$ is pivotal between alternatives a and b in some neighborhood of other agents report profiles.

Property 4: For a and b that satisfy property 3 there exist $\theta^1_{\hat{\jmath}}$, $\theta^2_{\hat{\jmath}}$ in every neighborhood in $\Theta_{\hat{\jmath}}$ such that $v_{\hat{\jmath}}\left(a, \theta^1_{\hat{\jmath}}\right) - v_{\hat{\jmath}}\left(b, \theta^1_{\hat{\jmath}}\right) > v_{\hat{\jmath}}\left(a, \theta^2_{\hat{\jmath}}\right) - v_{\hat{\jmath}}\left(b, \theta^2_{\hat{\jmath}}\right)$. In words, in every neighborhood of signals there exists a pair of signals $\theta^1_{\hat{\jmath}}$ and $\theta^2_{\hat{\jmath}}$ such that agent $\hat{\jmath}$'s valuation for moving from alternative b to alternative a is different given each signal.

To get a sense of the strength of this impossibility result consider the following widely used and analyzed setting. There is a finite set of alternatives A, a bijection function μ from A to $\{1, ..., |A|\}$, a convex signal space $\Theta_i \subseteq \mathbb{R}^{|A|}$ for every $i \in I$, and valuation functions $v_i(a, \theta_i) = \theta_i^{\mu(a)}$. Assuming that $\Theta_{\hat{\jmath}}$ contains an interval in some $\mu(a)$ axis and that the decision rule q is an affine maximizer,[13] we get that if agent $\hat{\imath}$ is pivotal between alternatives a and b, then properties 2, 3, and 4 hold. Therefore if property 1 also holds implementation is impossible. Settings of this kind are used in the analysis of many important economic scenarios, such as efficient auctions, efficient provision of public goods, efficient trading, and more. Moreover, Roberts (1979), Lavi et al. (2003), and Mishra and Sen (2012) characterize further conditions on the signal space such that if these conditions hold every decision rule that is dominant strategy implementable in PV is also an affine maximizer.[14] Lemma 1 then implies that if these conditions are met every decision rule that is ex-post implementable in IP is also an affine maximizer. Hence, in settings that satisfy both property 1 and the conditions

[12] $B\left(\left(\tilde{\theta}_{-\hat{\imath}}\right), r\right) \equiv \left\{\theta_{-i} \in \Theta_{-i} : d\left(\left(\theta_{-i}, \tilde{\theta}_{-\hat{\imath}}\right) \leq r\right)\right\}$

[13] A decision rule, q, is an *affine maximizer* if there exist an n-tuple $\lambda_{1,...,}\lambda_n$ not equal to zero and a function $\kappa : A \to \mathbb{R}$, and $q(\theta) = a$ if $\sum_{i \in I} \lambda_i \theta_i^{\mu(a)} + \kappa(a) > \sum_{i \in I} \lambda_i \theta_i^{\mu(a')} + \kappa(a')$ for every $a' \in A \setminus \{a\}$.

[14] Lavi et al. (2003) and Mishra and Sen (2012) consider decision rules that satisfy a certain property.

that appear in these papers our impossibility result applies to any decision rule with the property that agent $\hat{i}$ is a pivotal agent. We now present the impossibility result formally.

Theorem 2. *If in a profile Θ, $(v_i)_{i \in I}$, $IP \in P$, and $q(\theta)$, properties 1 to 4 hold, then $q(\theta)$ is not ex-post implementable in IP.*

The argument in the proof can be demonstrated by considering a model with two agents. Agent 1's utility may depend on the payoff of agent 2, i.e. $P_1 = \{2\}$, while agent 2's utility coincides with her own personal payoff, i.e. $P_2 = \emptyset$. Consider a decision rule $q(\theta)$ and a profile of valuation functions v_1 and v_2 that satisfy properties 2, 3 and 4. Consider some θ_2. The payoff of agent 2 given θ_2 as a function of agent 1's report, $\left(\hat{\theta}_1, \hat{\delta}_1\right)$, is $\Pi_2\left(\hat{\theta}_1, \hat{\delta}_1, \theta_2\right) = v_2\left(q\left(\hat{\theta}_1, \theta_2\right), \theta_2\right) + t_2\left(\hat{\theta}_1, \hat{\delta}_1, \theta_2\right)$. The transfer of agent 1 given θ_2 as a function of agent 1's report is $t_1\left(\hat{\theta}_1, \hat{\delta}_1, \theta_2\right)$. Agent 1's utility function given θ_2 is $v_1\left(q\left(\hat{\theta}_1, \theta_2\right), \theta_1\right) + \delta_1 \Pi_2\left(\hat{\theta}_1, \hat{\delta}_1, \theta_2\right) + t_1\left(\hat{\theta}_1, \hat{\delta}_1, \theta_2\right)$. Now assume that agent 1 reports δ_1 truthfully. Ex-post implementability implies that she must report θ_1 truthfully. The problem is therefore to incentivize agent 1 to report θ_1 truthfully when her utility function is $v_1\left(q\left(\hat{\theta}_1, \theta_2\right), \theta_1\right) + \delta_1 \Pi_2\left(\hat{\theta}_1, \delta_1, \theta_2\right) + t_1\left(\hat{\theta}_1, \delta_1, \theta_2\right)$. This problem is equivalent to the problem of incentivizing her to report truthfully in the PV setup.[15] Property 2 implies that in the PV setup revenue equivalence holds, i.e., the transfer to agent 1 given θ_2 in any transfer scheme that implements $q(\theta)$ is unique up to a constant.[16] Hence a truthful report of θ_1 implies that for every $\delta_1 \in \mathcal{D}_1$ and $\theta_1 \in \Theta_1$ we have

$$\delta_1 \Pi_2\left(\theta_1, \delta_1, \theta_2\right) + t_1\left(\theta_1, \delta_1, \theta_2\right) = \varphi\left(\theta_1, \theta_2\right) + \sigma\left(\delta_1, \theta_2\right) \tag{1}$$

where $\varphi : \Theta_1 \times \Theta_2 \to \mathbb{R}$ and $\sigma : \mathcal{D}_1 \times \Theta_2 \to \mathbb{R}$.[17]

[15] Define $\tilde{t}_1^{\delta_1}\left(\hat{\theta}_1, \theta_2\right) = \delta_1 \Pi_2\left(\hat{\theta}_1, \delta_1, \theta_2\right) + t_1\left(\hat{\theta}_1, \delta_1, \theta_2\right)$ and the problem is to incentivize agent 1 to report θ_1 truthfully given that her utility is $v_1\left(q\left(\hat{\theta}_1, \theta_2\right), \theta_1\right) + \tilde{t}_1^{\delta_1}\left(\hat{\theta}_1, \theta_2\right)$.

[16] See Krishna and Maenner (2001).

[17] Revenue equivalence means that $\tilde{t}_1^{\delta_1}\left(\hat{\theta}_1, \theta_2\right)$ equals some function that depends on θ_1 which we denote by $\varphi\left(\theta_1, \theta_2\right)$ plus a constant which we denote by $\sigma\left(\delta_1, \theta_2\right)$.

On the other hand assume agent 1 reports θ_1 truthfully. Ex-post implementability implies that she must report δ_1 truthfully, i.e, for every $\theta_1 \in \Theta_1$ and $\delta_1 \in \mathcal{D}_1$ we have

$$v_1(q(\theta_1,\theta_2),\theta_1)+\delta_1\Pi_2(\theta_1,\delta_1,\theta_2)+t_1(\theta_1,\delta_1,\theta_2) \geq v_1(q(\theta_1,\theta_2),\theta_1)+\delta_1\Pi_2\left(\theta_1,\hat{\delta}_1,\theta_2\right)+t_1\left(\theta_1,\hat{\delta}_1,\theta_2\right)$$

for every $\hat{\delta}_1 \in \mathcal{D}_1$. Subtracting $v_1(q(\theta_1,\theta_2),\theta_1)$ from both sides of the inequality we have

$$\delta_1\Pi_2(\theta_1,\delta_1,\theta_2)+t_1(\theta_1,\delta_1,\theta_2) \geq \delta_1\Pi_2\left(\theta_1,\hat{\delta}_1,\theta_2\right)+t_1\left(\theta_1,\hat{\delta}_1,\theta_2\right)$$

for every $\hat{\delta}_1 \in \mathcal{D}_1$. This implies that[18]

$$(2)\quad \delta_1\Pi_2(\theta_1,\delta_1,\theta_2)+t_1(\theta_1,\delta_1,\theta_2)=\underline{\delta_1}\Pi_2\left(\theta_1,\underline{\delta_1},\theta_2\right)+t_1\left(\theta_1,\underline{\delta_1},\theta_2\right)+\int_{\underline{\delta_1}}^{\delta_1}\Pi_2(\theta_1,s,\theta_2)\,ds$$

Combining equations (1) and (2) yields that for every $\delta_1 \in \mathcal{D}_1$ and every $\theta_1 \in \Theta_1$, $\int_{\underline{\delta_1}}^{\delta_1}\Pi_2(\theta_1,s,\theta_2)\,ds=\sigma(\delta_1,\theta_2)-\sigma\left(\underline{\delta_1},\theta_2\right)$. This implies that that for every $\theta_1,\theta_1^{'} \in \Theta_1$, $\Pi_2(\theta_1,\cdot,\theta_2) \overset{a.e}{=} \Pi_2\left(\theta_1^{'},\cdot,\theta_2\right)$. Now due to properties 3 and 4 we can find signals θ_1, $\theta_1^{'}$, θ_2 and $\theta_2^{'}$ such that $q(\theta_1,\theta_2)=q\left(\theta_1,\theta_2^{'}\right)=a$, $q\left(\theta_1^{'},\theta_2\right)=q\left(\theta_1^{'},\theta_2^{'}\right)=b$, and $v_2(a,\theta_2)-v_2(b,\theta_2)\neq v_2\left(a,\theta_2^{'}\right)-v_2\left(b,\theta_2^{'}\right)$. In addition, we can find a signal δ_1 such that $\Pi_2(\theta_1,\delta_1,\theta_2)=\Pi_2\left(\theta_1^{'},\delta_1,\theta_2\right)$ and $\Pi_2\left(\theta_1,\delta_1,\theta_2^{'}\right)=\Pi_2\left(\theta_1^{'},\delta_1,\theta_2^{'}\right)$. This yields that

$$t_2(\theta_1,\delta_1,\theta_2)-t_2\left(\theta_1^{'},\delta_1,\theta_2\right)\neq t_2\left(\theta_1,\delta_1,\theta_2^{'}\right)-t_2\left(\theta_1^{'},\delta_1,\theta_2^{'}\right)$$

However, for agent 2 to report truthfully the function t_2 must assign the same transfer to signals that map the same alternative for a given report of agent 1. This implies that

$$t_2(\theta_1,\delta_1,\theta_2)-t_2\left(\theta_1^{'},\delta_1,\theta_2\right)=t_2\left(\theta_1,\delta_1,\theta_2^{'}\right)-t_2\left(\theta_1^{'},\delta_1,\theta_2^{'}\right)$$

a contradiction.

The impossibility of ex-post implementation in the interdependent payoffs model is yet another example of the difficulty of implementing decision rules by robust solution

[18]This stems from the following result. Let $u(\delta,\hat{\delta})=\delta\cdot q\left(\hat{\delta}\right)+t\left(\hat{\delta}\right)$. If for every $\delta \in [\underline{\delta},\overline{\delta}]$, $\delta \in \underset{\hat{\delta}\in[\underline{\delta},\overline{\delta}]}{\arg\max}\, u(\delta,\hat{\delta})$ then for every $\delta \in \left[\underline{\delta},\overline{\delta}\right]$, $t(\delta)+\delta q(\delta)=t(\underline{\delta})+\underline{\delta}\cdot q(\underline{\delta})+\int_{\underline{\delta}}^{\delta} q(s)\,ds$.

concepts in economic environments that diverge from the independent private values model. Jehiel et al. (2006) present an impossibility result on ex-post implementation in environments with interdependent values. They show that for generic valuation functions the only deterministic decision rules that are ex-post implementable are constant. As mentioned in the Introduction, our model is different from their model in the following way. In the interdependent values model agent i's report affects his utility through the decision rule q and his personal transfer t_i. In the interdependent payoffs model, however, agent i's report affects his utility through the decision rule q, his personal transfer t_i, and the personal transfers of the agents whose payoffs affect agent i's utility[19] $(t_j)_{j \in P_i}$. That is, in the interdependent payoffs model mechanisms affect agents' incentives in a more complex way, compared to in the interdependent values model. On the one hand, since an agent's utility is affected by other agents' transfers, mechanisms provide more freedom to align agents' preferences with social preferences and to achieve implementation. On the other hand, since each agent's transfer also affects the incentives of the other agents, mechanisms also impose further restrictions on achieving implementation. In general environments, the extra degrees of freedom that mechanisms provide in settings with interdependent payoffs are offset by the restrictions they impose, and ex-post implementation is impossible. Nonetheless, in particular setups that do not satisfy the conditions of Theorem 2, the freedom that mechanisms provide in settings with interdependent payoffs allows for ex-post implementation of non-constant decision rules. We illustrate this point in the following example that presents a setup in which non-constant decision rules are ex-post implementable in the interdependent payoffs model but are not ex-post implementable in the interdependent values model.

Example 3. Consider the following setup. There are two agents $I = \{1, 2\}$ and two alternatives $A = \{a, b\}$. Each agent $i \in I$ receives two private signals $\theta_i \in [0, 1]$ and $\delta_i \in [0, 1]$. The first signal affects his valuation while the second signal affects the dependency of his utility on the other agent's payoff\valuation. Agent i's valuation, $i \in I$, if alternative a is chosen is $v_i(a, \theta_i) = \theta_i + c$, and his valuation if alternative b

[19]Note that while the effect of the agent's personal transfer on his utility is independent of the realization on signals, the effect of other agents' transfers on his utility depends on the realization of signals.

is chosen is $v_i(b,\theta_i)=\theta_i$. We now analyze the possibility to implement decision rules that depend only on information about agents' payoffs both in the interdependent payoffs model and in the interdependent values model.
We first adapt this setup to the interdependent payoffs model. In this case agent i's utility is $v_i(q,\theta_i)+\delta_i\cdot(v_j(q,\theta_j)+t_j)+t_i$ where $q\in A$. We now show that every decision rule is ex-post implementable in this model. Consider an arbitrary decision rule $q(\theta)$. For every $i\in\{1,2\}$ we define the following transfer function

$$t_i(\theta_i,\delta_i,\theta_j,\delta_j)=\begin{cases}-c & \text{if } q(\theta_i,\theta_j)=a\\ 0 & \text{if } q(\theta_i,\theta_j)=b\end{cases}$$

Under these transfer functions any type (θ_i,δ_i) of agent i receive the same utility, $\theta_i+\delta_i\cdot\theta_j$, irrespective of his report. Therefore, the decision rule is ex-post implementable.
We now adapt the above setup to the interdependent values model. In this case agent i's utility is $v_i(q,\theta_i)+\delta_i\cdot v_j(q,\theta_j)+t_i$ where $q\in A$. We now show that it is impossible to implement ex-post non-constant decision rules in this model. Consider an arbitrary type $\left(\tilde{\theta}_j,\tilde{\delta}_j\right)$ of agent j, $j\neq i$. Ex-post implementability implies that for every $(\theta_i,\delta_i),\left(\theta_i^{'},\delta_i^{'}\right)\in[0,1]^2$ such that $q\left(\theta_i,\tilde{\theta}_j\right)=q\left(\theta_i^{'},\tilde{\theta}_j\right)$ we have $t_i\left(\theta_i,\delta_i,\tilde{\theta}_j,\tilde{\delta}_j\right)=t_i\left(\theta_i^{'},\delta_i^{'},\tilde{\theta}_j,\tilde{\delta}_j\right)$.[20] That is, agent i's transfer function depends only on the chosen alternative, hence, we denote $t_i\left(\theta_i,\delta_i,\tilde{\theta}_j,\tilde{\delta}_j\right):=t_i\left(q(\theta_i,\theta_j),\tilde{\theta}_j,\tilde{\delta}_j\right)$. Consider a non-constant decision rule $q(\theta)$. Look at a type $\left(\tilde{\theta}_j,\tilde{\delta}_j\right)$ of agent j for which agent i is pivotal. This means that there exist two signals $\theta_i^{'}$ and $\theta_i^{''}$ such that $q\left(\theta_i^{'},\tilde{\theta}_j\right)=a$ and $q\left(\theta_i^{''},\tilde{\theta}_j\right)=b$. Now, ex-post implementability implies that for every $\delta_i\in[0,1]$ we have that

$$\theta_i^{'}+c+\delta_i\cdot\left(\tilde{\theta}_j+c\right)+t_i\left(a,\tilde{\theta}_j,\tilde{\delta}_j\right)\geq\theta_i^{'}+\delta_i\cdot\tilde{\theta}_j+t_i\left(b,\tilde{\theta}_j,\tilde{\delta}_j\right)$$

and

$$\theta_i^{''}+c+\delta_i\cdot\left(\tilde{\theta}_j+c\right)+t_i\left(a,\tilde{\theta}_j,\tilde{\delta}_j\right)\leq\theta_i^{''}+\delta_i\cdot\tilde{\theta}_j+t_i\left(b,\tilde{\theta}_j,\tilde{\delta}_j\right)$$

[20] Assume $t_i\left(\theta_i,\delta_i,\tilde{\theta}_j,\tilde{\delta}_j\right)>t_i\left(\theta_i^{'},\delta_i^{'},\tilde{\theta}_j,\tilde{\delta}_j\right)$, then agent i of type $\left(\theta_i^{'},\delta_i^{'}\right)$ will have a profitable deviation to (θ_i,δ_i)

hence we get that for every $\delta_i \in [0,1]$

$$c \cdot (1+\delta_i) = t_i\left(b, \tilde{\theta}_j, \tilde{\delta}_j\right) - t_i\left(a, \tilde{\theta}_j, \tilde{\delta}_j\right)$$

Since the left hand side of the equation varies with δ_i and the right hand side of the equation is constant we reach a contradiction.

4 Bayesian Implementation

We start with a definition of Bayesian implementation in the context of our model. Bayes–Nash equilibrium requires that the strategy of each agent i be optimal in expectation with respect to the strategies of the other agents given agent i's knowledge on the distributions of the other agents' signals. By the revelation principle, we can restrict our analysis to direct mechanisms. Consider a given profile of Θ, $(v_i)_{i\in I}$, $(f_i)_{i\in I}$, and an IP setup. We say that a social choice function $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ is *Bayesian implementable in IP* if for every $i \in I$, $\theta_i \in \Theta_i$ and $\delta_i \in \mathcal{D}_i$ we have,

$$(\theta_i, \delta_i) \in \underset{(\hat{\theta}_i, \hat{\delta}_i) \in \Theta_i \times \mathcal{D}_i}{\arg\max} E_{\theta_{-i},\delta_{-i}} \Big[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right) +$$

$$\sum_{k\in P_i}\left[\delta_i^k \left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right)\right)\right]\Big]$$

A decision rule $q(\theta)$ is *Bayesian implementable in IP* if there exists a profile of real valued functions $(t_1(\theta,\delta), ..., t_n(\theta,\delta))$ such that $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ is Bayesian implementable in IP.

Consider a given profile of Θ, $(v_i)_{i\in I}$, $(f_i)_{i\in I}$, and the PV setup. We say that a social choice function of the form $(q(\theta), t_1(\theta), ..., t_n(\theta))$ is *Bayesian implementable in PV* if for every $i \in I$ and $\theta \in \Theta$ we have,

$$\theta_i \in \underset{\hat{\theta}_i \in \Theta_i}{\arg\max} E_{\theta_{-i}} \left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right)\right)\right]$$

A decision rule $q(\theta)$ is *Bayesian implementable in PV* if there exists a profile of real valued functions $(t_1(\theta), ..., t_n(\theta))$ such that $(q(\theta), t_1(\theta), ..., t_n(\theta))$ is Bayesian implementable in PV. Bayesian implementability in IP implies Bayesian implementability in PV:

Proposition 4. *Consider a given profile of* Θ*,* $(v_i)_{i\in I}$*,* $(f_i)_{i\in I}$*, and an IP setup. If a decision rule* $q(\theta)$ *is Bayesian implementable in IP then it is Bayesian implementable in PV.*

We show that the converse is also true: every decision rule that is Bayesian implementable in PV is Bayesian implementable in every IP. In addition, there exists a transfer scheme that implements the decision rule for every element in R, namely, for every IP setup and for the PV setup. Moreover, if the decision rule is dominant strategy implementable in PV, then there exists a transfer scheme that implements the decision rule in dominant strategy in PV and in a Bayes–Nash equilibrium in every IP setup. We achieve these results by constructing a transfer scheme that eliminates the effect of agent i's report on the expected payoff of any agent but agent i. At the same time, this transfer scheme incentivizes agent i to report truthfully when she is interested in maximizing her own personal payoff. Therefore, this transfer scheme incentivizes truth telling in every setup.[21] We now show the result formally.

Theorem 5. *Consider a profile* Θ*,* $(v_i)_{i\in I}$*, and* $(f_i)_{i\in I}$*. Let* $(q(\theta), t_1(\theta), ..., t_n(\theta))$ *be Bayesian implementable in PV; then there exists a social choice function* $\left(q(\theta), t_1^{'}(\theta), ..., t_n^{'}(\theta)\right)$ *that is Bayesian implementable in IP for every* $IP \in P$ *with the following properties:*

1. $\left(q(\theta), t_1^{'}(\theta), ..., t_n^{'}(\theta)\right)$ *is Bayesian implementable in PV.*
2. *If* $(q(\theta), t_1(\theta), ..., t_n(\theta))$ *is dominant strategy implementable in PV, then* $\left(q(\theta), t_1^{'}(\theta), ..., t_n^{'}(\theta)\right)$ *is dominant strategy implementable in PV.*
3. $E_\theta[t_i(\theta)] = E_\theta\left[t_i^{'}(\theta)\right]$ *for every* $i \in I$.

The structure of the proof is as follows. Given a transfer scheme $(t_i(\theta))_{i\in I}$ that implements $q(\theta)$ in PV, we define $\left(t_i^{'}(\theta)\right)_{i\in I}$ to be

$$t_i^{'}(\theta) = t_i(\theta) - \sum_{j\in I\setminus\{i\}} E_{\tilde{\theta}_{-j}}\left[v_i\left(q\left(\theta_j, \tilde{\theta}_{-j}\right), \tilde{\theta}_i\right) + t_i\left(\theta_j, \tilde{\theta}_{-j}\right)\right] + \sum_{j\in I\setminus\{i\}} E_{\tilde{\theta}}\left[v_i\left(q\left(\tilde{\theta}\right), \tilde{\theta}_i\right) + t_i\left(\tilde{\theta}\right)\right]$$

[21] Similar approach is used in Bierbrauer and Netzer (2016).

Consider an IP setup. Let $i \in I$ and $k \in P_i$. The i-th element in the second additive term (the sum over $E_{\tilde{\theta}_{-j}}$) in the transfer function $t_k^{'}(\theta)$ eliminates the effect of the report of agent i, $\hat{\theta}_i$, on the expected payoff of agent k from agent i's perspective. In other words, from agent i's perspective, the report $\hat{\theta}_i$ does not affect the expected payoff of agent k, for every $k \in P_i$. It is therefore sufficient to show that $\left(t_i^{'}(\theta)\right)_{i \in I}$ Bayesian implements $q(\theta)$ in PV. This follows from the fact that $t_i^{'}(\theta)$ equals $t_i(\theta)$ plus additive terms that do not depend on $\hat{\theta}_i$ and that $(t_i(\theta))_{i \in I}$ Bayesian implements $q(\theta)$ in PV. For the same reasons we get that if $(t_i(\theta))_{i \in I}$ implements $q(\theta)$ in dominant strategy in PV then so does $\left(t_i^{'}(\theta)\right)_{i \in I}$.

5 Interdependent Utilities

Our model considers agents who posses preferences of interdependent payoffs. However, this model can be suitable for the case where agents have interdependent utilities. Consider the following two-agents case. Each agent i benefits from her own welfare, Π_i, and from observing the other agent's utility. The assumption is that the utility function of agent i is $U_i = \Pi_i + \delta_i U_j$, where $\delta_i \in \left[0, \overline{\delta_i}\right]$ with $\overline{\delta_i} < 1$.[22] Agent i's welfare is $\Pi_i = v_i(a, \theta_i) + t_i$ where $a \in A$ is the chosen alternative, $\theta_i \in \Theta_i$ is agent i's signal, and t_i is agent i's monetary transfer. We assume that δ_i and θ_i hold the same properties which are detailed in section 2.[23] Solving for U_1 and U_2 we get that $U_i = \left(\frac{1}{1-\delta_1\delta_2}\right)\Pi_i + \left(\frac{\delta_i}{1-\delta_1\delta_2}\right)\Pi_j$. In such a setup the standard utilitarian social choice function that sums the agents' utilities, namely, $U_1 + U_2 = \frac{(1+\delta_2)\cdot\Pi_1 + (1+\delta_1)\cdot\Pi_2}{1-\delta_1\delta_2}$, has been criticized for rewarding the more selfish agent by assigning her a greater weight.[24] The way to deal with this ethical criticism is to remove anti-social preferences from consideration.[25] Therefore, the appropriate social choice function is the sum of welfares, namely, the objective function of the designer is $\Pi_1 + \Pi_2$. Under this objective function the optimal decision rule depends on θ alone (and not on δ). We show in the Appendix that the problem of implementing the optimal decision rule in this case

[22]This setup appears in Bergstrom (1989), (1999).
[23]This assumption seems natural in this scenario.
[24]See Blanchet and Fleurbaey (2006).
[25]See Goodin (1986) and Harsanyi (1977).

can be solved by using the results on implementation that we have developed for the interdependent payoffs model.

6 Concluding Remarks

We have considered the problem of implementation in a model with agents who have interdependent payoffs. We have considered both ex-post and Bayesian implementation. We have shown that ex-post implementation is impossible, while Bayesian implementation allows for the implementation of every decision rule that is implementable in the independent private values model. These results suggest that the less knowledge there is of the economic environment, the harder it is to acquire information in the presence of personal interests. Our impossibility result highlights the question whether ex-post implementation is possible in other environments with social preferences. The environment we have considered joins several other environments in which it has been shown that implementation in robust solution concepts is impossible. This presents yet another example of the difficulty of carrying out robust implementation.

A Appendix

A.1 Interdependent Utilities

In this subsection we show that the problem of implementing the optimal decision rule in the interdependent utilities setup can be solved by using the results on implementation that we have developed for the interdependent payoffs model. We start by showing that ex-post implementation of the optimal decision rule is impossible in the interdependent utilities setup. Consider the optimal decision rule in the interdependent utilities setup

$$q(\theta) \in \arg\max_{a \in A} v_1(a, \theta_1) + v_2(a, \theta_2)$$

The optimal decision rule is ex-post implementable in the interdependent utilities setup if for every $i \in I$, $\theta \in \Theta$ and $\delta \in \mathcal{D}$ we have

$$\frac{1}{1-\delta_1\delta_2}\left[v_i\left(q\left(\theta_i,\theta_{-i}\right),\theta_i\right)+t_i\left(\left(\theta_i,\theta_{-i}\right),\left(\delta_i,\delta_{-i}\right)\right)+\delta_i\left(v_j\left(q\left(\theta_i,\theta_j\right),\theta_j\right)+t_j\left(\left(\theta_i,\theta_j\right),\left(\delta_i,\delta_j\right)\right)\right)\right]\geq$$

$$\frac{1}{1-\delta_1\delta_2}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)+t_i\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\hat{\delta}_i,\delta_{-i}\right)\right)+\delta_i\left(v_j\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_j\right)+t_j\left(\left(\hat{\theta}_i,\theta_j\right),\left(\hat{\delta}_i,\delta_j\right)\right)\right)\right]$$

this inequality holds if and only if for every $i \in I$, $\theta \in \Theta$ and $\delta \in \mathcal{D}$ we have

$$v_i\left(q\left(\theta_i,\theta_{-i}\right),\theta_i\right)+t_i\left(\left(\theta_i,\theta_{-i}\right),\left(\delta_i,\delta_{-i}\right)\right)+\delta_i\left(v_j\left(q\left(\theta_i,\theta_j\right),\theta_j\right)+t_j\left(\left(\theta_i,\theta_j\right),\left(\delta_i,\delta_j\right)\right)\right)$$

$$\geq v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)+t_i\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\hat{\delta}_i,\delta_{-i}\right)\right)+\delta_i\left(v_j\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_j\right)+t_j\left(\left(\hat{\theta}_i,\theta_j\right),\left(\hat{\delta}_i,\delta_j\right)\right)\right)$$

and this inequality holds if and only if $q(\theta)$ is ex-post implementable in IP. We showed in Section 4 that if properties 1 to 4 hold, the optimal decision rule is not ex-post implementable in IP. Therefore, if these properties hold in the interdependent utilities setup, the optimal decision rule is not ex-post implementable in the interdependent utilities setup.

We now show, using the result on Bayesian implementation in the interdependent payoffs model, that the optimal decision rule is Bayesian implementable in the interdependent utilities setup.

The optimal decision rule is dominant strategy implementable in PV. Therefore, by Theorem 4, there exists a transfer scheme $(t_1(\theta), t_2(\theta))$ the implements it in a Bayes-Nash equilibrium in any IP, i.e., for every δ_i and θ_i

$$\theta_i \in \underset{\hat{\theta}_i\in\Theta_i}{\arg\max} E_{\theta_j,\delta_j}\left[v_i\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_i\right)+t_i\left(\hat{\theta}_i,\theta_j\right)+\delta_i\left(v_j\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_j\right)+t_j\left(\hat{\theta}_i,\theta_j\right)\right)\right]$$

given this transfer scheme we get that for every δ_i and θ_i

$$\theta_i \in \underset{\hat{\theta}_i\in\Theta_i}{\arg\max} E_{\theta_j,\delta_j}\left\{\frac{1}{1-\delta_1\delta_2}\left[v_i\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_i\right)+t_i\left(\hat{\theta}_i,\theta_j\right)+\delta_i\left(v_j\left(q\left(\hat{\theta}_i,\theta_j\right),\theta_j\right)+t_j\left(\hat{\theta}_i,\theta_j\right)\right)\right]\right\}$$

i.e., the optimal decision rule is Bayesian implementable in the interdependent utilities setup.

A.2 Proofs

A.2.1 Proof of Lemma 1

Assume that $q(\theta)$ is ex-post implementable in IP; then there exists a social choice function $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ such that for every $i \in I$, $\theta \in \Theta$, and $\delta \in \mathcal{D}$, we have

$$(\theta_i, \delta_i) \in \underset{(\hat{\theta}_i, \hat{\delta}_i) \in \Theta_i \times \mathcal{D}_i}{\arg\max}\ v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right) +$$

$$\sum_{k \in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i, \delta_{-i}\right)\right)\right)\right]$$

Choose an arbitrary $\tilde{\delta} \in \mathcal{D}$; then for every $i \in I$, and $\theta \in \Theta$, we have

$$\theta_i \in \arg\max_{\hat{\theta}_i} v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \tilde{\delta}\right) + \sum_{k \in P_i}\left[\tilde{\delta}_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \tilde{\delta}\right)\right)\right]$$

For every $i \in I$, we define $t_i^{'}(\theta) := t_i\left(\theta, \tilde{\delta}\right) + \sum_{k \in P_i}\left[\tilde{\delta}_i^k\left(v_k(q(\theta), \theta_k) + t_k\left(\theta, \tilde{\delta}\right)\right)\right]$. Then for every $i \in I$, and $\theta \in \Theta$, we have $\theta_i \in \arg\max_{\hat{\theta}_i} v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i^{'}\left(\hat{\theta}_i, \theta_{-i}\right)$, namely, $\left(q(\theta), t_1^{'}(\theta), ..., t_n^{'}(\theta)\right)$ is ex-post implementable in PV, and hence $q(\theta)$ is ex-post implementable in PV ■

A.2.2 Proof of Theorem 2

We first present two lemmas that are needed to prove the main theorem and some notations.

Lemma 5. *Assume that property 2 holds; Then for every decision rule $q(\theta)$ that is ex-post implementable in PV, there exists a profile of functions $(\tilde{t}_1(\theta), ..., \tilde{t}_n(\theta))$ such that a social choice function $\left(q(\theta), t_1^{'}(\theta), ..., t_n^{'}(\theta)\right)$ is ex-post implementable in PV if and only if there exists a function $\tau_i : \Theta_{-i} \to \mathbb{R}$ such that $t_i^{'}(\theta) = \tilde{t}_i(\theta) + \tau_i(\theta_{-i})$ for every $i \in I$. Moreover, if for a given profile θ_{-i} there exist θ_i^1 and θ_i^2 such that $q(\theta_i^1, \theta_{-i}) = q(\theta_i^2, \theta_{-i})$, then $\tilde{t}_i(\theta_i^1, \theta_{-i}) = \tilde{t}_i(\theta_i^2, \theta_{-i})$.*

Proof: For the first part of the Lemma see Krishna and Maenner (2001). we proceed to the proof of the second part of the lemma. Let $i \in I$ and let θ_{-i} be a profile such that there exist θ_i^1 and θ_i^2 for which $q(\theta_i^1, \theta_{-i}) = q(\theta_i^2, \theta_{-i})$ and assume by negation that $\tilde{t}_i(\theta_i^1, \theta_{-i}) \neq \tilde{t}_i(\theta_i^2, \theta_{-i})$. Assume w.l.o.g. that $\tilde{t}_i(\theta_i^1, \theta_{-i}) > \tilde{t}_i(\theta_i^2, \theta_{-i})$, then

$v_i\left(q\left(\theta_i^1,\theta_{-i}\right),\theta_i^2\right)+\tilde{t}_i\left(\theta_i^1,\theta_{-i}\right)+\tau_i\left(\theta_{-i}\right)>v_i\left(q\left(\theta_i^2,\theta_{-i}\right),\theta_i^2\right)+\tilde{t}_i\left(\theta_i^2,\theta_{-i}\right)+\tau_i\left(\theta_{-i}\right)$, in contradiction to the fact that $\left(q\left(\theta\right),\tilde{t}_1\left(\theta\right),...,\tilde{t}_n\left(\theta\right)\right)$ is ex-post implementable in PV ■

Lemma 6: *Let* $u:\left[\underline{\delta},\overline{\delta}\right]^2\rightarrow\mathbb{R}$ *be* $u(\delta,\hat{\delta})=\delta\cdot q\left(\hat{\delta}\right)+t\left(\hat{\delta}\right)$. *If for every* $\delta\in\left[\underline{\delta},\overline{\delta}\right]$, $\delta\in\arg\max\limits_{\hat{\delta}\in\left[\underline{\delta},\overline{\delta}\right]}u(\delta,\hat{\delta})$, *then:*

1. $q\left(\cdot\right)$*is non-decreasing*

2. for every $\delta\in\left[\underline{\delta},\overline{\delta}\right]$, $\qquad t\left(\delta\right)+\delta q\left(\delta\right)=t\left(\underline{\delta}\right)+\underline{\delta}\cdot q\left(\underline{\delta}\right)+\int_{\underline{\delta}}^{\delta}q(s)\,ds$.

Proof: Assume that $\delta\in\arg\max\limits_{\hat{\delta}}u(\delta,\hat{\delta})$; then for every $\tilde{\delta}>\delta$ we have $u(\delta,\delta)\geq\delta\cdot q\left(\tilde{\delta}\right)+t\left(\tilde{\delta}\right)=u(\tilde{\delta},\tilde{\delta})+\left(\delta-\tilde{\delta}\right)\cdot q\left(\tilde{\delta}\right)$ and $u(\tilde{\delta},\tilde{\delta})\geq\tilde{\delta}\cdot q\left(\delta\right)+t\left(\delta\right)=u(\delta,\delta)+\left(\tilde{\delta}-\delta\right)\cdot q\left(\delta\right)$. We let $V(\delta):=u(\delta,\delta)$. Then we get $q\left(\tilde{\delta}\right)\geq\frac{V(\tilde{\delta})-V(\delta)}{\tilde{\delta}-\delta}\geq q(\delta)$, i.e., $q\left(\cdot\right)$ is non-decreasing. In addition if $V\left(\cdot\right)$ is differentiable in δ, then $V'\left(\delta\right)=q\left(\delta\right)$. Now $V(\delta)=\max\limits_{\hat{\delta}\in\left[\underline{\delta},\overline{\delta}\right]}u(\delta,\hat{\delta})$, namely, it is a maximum of affine functions therefore it is convex and thus absolutely continuous so $V(\delta)=V(\underline{\delta})+\int_{\underline{\delta}}^{\delta}q(s)\,ds$, namely $t\left(\delta\right)+\delta q\left(\delta\right)=t\left(\underline{\delta}\right)+\underline{\delta}\cdot q\left(\underline{\delta}\right)+\int_{\underline{\delta}}^{\delta}q(s)\,ds$ ■

Notations: We denote $\Theta_{-i-j}:=\underset{k\in I\setminus\{i,j\}}{\times}\Theta_k$ with generic element θ_{-i-j}. We denote $D_{-i}^{-j}:=\mathcal{D}\setminus\left[\underline{\delta_i^j},\overline{\delta_i^j}\right]$ with generic element δ_{-i}^{-j}.

Proof of the Theorem: Assume by negation that $q\left(\theta\right)$ is ex-post implementable in IP, then there exists a social choice function $\left(q\left(\theta\right),t_1\left(\theta,\delta\right),...,t_n\left(\theta,\delta\right)\right)$ that is ex-post implementable in IP. According to the proof of Lemma 1, given some arbitrary profile $\tilde{\delta}\in\mathcal{D}$, if we define $\mu_i\left(\theta\right):=t_i\left(\theta,\tilde{\delta}\right)+\sum\limits_{k\in P_i}\left[\tilde{\delta}_i^k\left(v_k\left(q\left(\theta\right),\theta_k\right)+t_k\left(\theta,\tilde{\delta}\right)\right)\right]$, then $\left(q\left(\theta\right),\mu_1\left(\theta\right),...,\mu_n\left(\theta\right)\right)$ is ex-post implementable in PV. Therefore, according to Lemma 5, there exists a profile of functions $\tilde{t}_1\left(\theta\right),...,\tilde{t}_n\left(\theta\right)$ and a profile of functions $\tau_1\left(\theta_{-1},\tilde{\delta}\right),...,\tau_n\left(\theta_{-n},\tilde{\delta}\right)$ such that $\mu_i\left(\theta\right)=\tilde{t}_i\left(\theta\right)+\tau_i\left(\theta_{-i},\tilde{\delta}\right)$, and so we get $t_i\left(\theta,\tilde{\delta}\right)+\sum\limits_{k\in P_i}\left[\tilde{\delta}_i^k\left(v_k\left(q\left(\theta\right),\theta_k\right)+t_k\left(\theta,\tilde{\delta}\right)\right)\right]=\tilde{t}_i\left(\theta\right)+\tau_i\left(\theta_{-i},\tilde{\delta}\right)$. This is true for every $\delta\in\mathcal{D}$ and therefore we can present the following general expression. For every $\delta\in\mathcal{D}$,

$$(1)\qquad t_i\left(\theta,\delta\right)+\sum_{k\in P_i}\left[\delta_i^k\left(v_k\left(q\left(\theta\right),\theta_k\right)+t_k\left(\theta,\delta\right)\right)\right]=\tilde{t}_i\left(\theta\right)+\tau_i\left(\theta_{-i},\delta\right)$$

Let $i\in I$ and $j\in P_i$. Consider an arbitrary profile $\left(\theta_{-i},\delta_{-i}^{-j}\right)\in\Theta_{-i}\times\mathcal{D}_{-i}^{-j}$ (i.e., a profile of all the signals except for θ_i and δ_i^j). Moreover assume that $\left(\hat{\delta}_i^k\right)_{k\in P_i\setminus\{j\}} = \left(\delta_i^k\right)_{k\in P_i\setminus\{j\}}$ (i.e., agent i truthfully reports all her signals except for θ_i and δ_i^j). The utility function of agent i given this profile as a function of her true signal $\left(\theta_i, \delta_i^j\right)$ and her report $\left(\hat{\theta}_i, \hat{\delta}_i^j\right)$ is

$$(2) \quad v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i^j, \delta_{-i}^{-j}\right)\right) + \sum_{k\in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i^j, \delta_{-i}^{-j}\right)\right)\right)\right]$$

We present the following notations:

$$\bar{t}_i\left(\hat{\theta}_i, \hat{\delta}_i^j\right) := t_i\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i^j, \delta_{-i}^{-j}\right)\right) + \sum_{k\in P_i\setminus\{j\}}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i^j, \delta_{-i}^{-j}\right)\right)\right)\right]$$

$$\Pi_j\left(\hat{\theta}_i, \hat{\delta}_i^j\right) := v_j\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_j\right) + t_j\left(\left(\hat{\theta}_i, \theta_{-i}\right), \left(\hat{\delta}_i^j, \delta_{-i}^{-j}\right)\right)$$

$$\hat{t}_i\left(\theta_i\right) := \tilde{t}_i\left(\theta_i, \theta_{-i}\right)$$

$$C\left(\delta_i^j\right) := \tau_i\left(\theta_{-i}, \left(\delta_i^j, \delta_{-i}^{-j}\right)\right)$$

We write (2) using the above notations:

$$(2') \qquad v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + \bar{t}_i\left(\hat{\theta}_i, \hat{\delta}_i^j\right) + \delta_i^j \cdot \Pi_j\left(\hat{\theta}_i, \hat{\delta}_i^j\right)$$

By equation (1), ex-post implementation implies that

$$(3) \qquad \bar{t}_i\left(\theta_i, \delta_i^j\right) + \delta_i^j \cdot \Pi_j\left(\theta_i, \delta_i^j\right) = \hat{t}_i\left(\theta_i\right) + C\left(\delta_i^j\right)$$

for every $\delta_i^j \in \left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$ and, in particular, for $\underline{\delta_i^j}$

$$(3') \qquad \bar{t}_i\left(\theta_i, \underline{\delta_i^j}\right) + \underline{\delta_i^j} \cdot \Pi_j\left(\theta_i, \underline{\delta_i^j}\right) = \hat{t}_i\left(\theta_i\right) + C\left(\underline{\delta_i^j}\right)$$

Now ex-post implementation implies that for every $\theta_i \in \Theta_i$ and for every $\hat{\delta}_i^j \in \left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$

$$(4) \qquad v_i\left(q\left(\theta\right), \theta_i\right) + \bar{t}_i\left(\theta_i, \delta_i^j\right) + \delta_i^j \cdot \Pi_j\left(\theta_i, \delta_i^j\right) \geq v_i\left(q\left(\theta\right), \theta_i\right) + \bar{t}_i\left(\theta_i, \hat{\delta}_i^j\right) + \delta_i^j \cdot \Pi_j\left(\theta_i, \hat{\delta}_i^j\right)$$

Subtracting $v_i\left(q\left(\theta\right), \theta_i\right)$ from both sides of the inequality we have

$$\delta_i^j \cdot \Pi_j\left(\theta_i, \delta_i^j\right) + \bar{t}_i\left(\theta_i, \delta_i^j\right) \geq \delta_i^j \cdot \Pi_j\left(\theta_i, \hat{\delta}_i^j\right) + \bar{t}_i\left(\theta_i, \hat{\delta}_i^j\right)$$

By lemma 6, this implies that

$$(5) \qquad \bar{t}_i\left(\theta_i, \delta_i^j\right) + \delta_i^j \cdot \Pi_j\left(\theta_i, \delta_i^j\right) = \bar{t}_i\left(\theta_i, \underline{\delta_i^j}\right) + \underline{\delta_i^j} \cdot \Pi_j\left(\theta_i, \underline{\delta_i^j}\right) + \int_{\underline{\delta_i^j}}^{\delta_i^j} \Pi_j\left(\theta_i, s\right)\, ds$$

Plugging equations (3) and (3') into equation (5) yields

$$(6) \qquad \hat{t}_i\left(\theta_i\right) + C\left(\delta_i^j\right) = \hat{t}_i\left(\theta_i\right) + C\left(\underline{\delta_i^j}\right) + \int_{\underline{\delta_i^j}}^{\delta_i^j} \Pi_j\left(\theta_i, s\right)\, ds$$

Subtracting $\hat{t}_i\left(\theta_i\right)$ from both sides of the equality we have for every $\delta_i^j \in \left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$

$$(6') \qquad \int_{\underline{\delta_i^j}}^{\delta_i^j} \Pi_j\left(\theta_i, s\right)\, ds = C\left(\delta_i^j\right) - C\left(\underline{\delta_i^j}\right)$$

for every $\theta_i \in \Theta_i$.
Therefore, we get that for every $\theta_i, \theta_i^{'} \in \Theta_i$, $\Pi_j\left(\theta_i, \cdot\right) \overset{a.e}{=} \Pi_j\left(\theta_i^{'}, \cdot\right)$. This means that the functions can receive different values only on a subset of $\left[\underline{\delta_i^j}, \overline{\delta_i^j}\right]$ that has a measure of zero. We conclude that given arbitrary θ_{-i} and δ_{-i}^{-j} it must be the case that

$$(7) \quad v_j\left(q\left(\theta_i, \theta_{-i}\right), \theta_j\right) + t_j\left(\left(\theta_i, \theta_{-i}\right), \left(\cdot, \delta_{-i}^{-j}\right)\right) \overset{a.e}{=} v_j\left(q\left(\theta_i^{'}, \theta_{-i}\right), \theta_j\right) + t_j\left(\left(\theta_i^{'}, \theta_{-i}\right), \left(\cdot, \delta_{-i}^{-j}\right)\right)$$

for every $\theta_i, \theta_i^{'} \in \Theta_i$.

- According to property 1 there exists an agent $\hat{i}$ and $\hat{j} \in P_{\hat{i}}$ such that for every $k \in P_{\hat{j}}$, $0 \in \left[\underline{\delta_{\hat{j}}^k}, \overline{\delta_{\hat{j}}^k}\right]$.

- According to property 3 there exist a profile $\tilde{\theta}_{-\hat{i}}$, two alternatives a and b and two types $\theta_{\hat{i}}^1$, $\theta_{\hat{i}}^2 \in \Theta_{\hat{i}}$ such that $q\left(\theta_{\hat{i}}^1, \tilde{\theta}_{-\hat{i}}\right) = a$ and $q\left(\theta_{\hat{i}}^2, \tilde{\theta}_{-\hat{i}}\right) = b$. In addition there exists $r > 0$ such that for every $\theta_{-i} \in B\left(\left(\tilde{\theta}_{-\hat{i}}\right), r\right)$, $q(\theta_{\hat{i}}^1, \theta_{-i}) = a$ and $q(\theta_{\hat{i}}^2, \theta_{-i}) = b$.

- According to property 4 there exist $\theta_{\hat{j}}^1$, $\theta_{\hat{j}}^2 \in \Theta_{\hat{j}}$ such that, $v_{\hat{j}}\left(a, \theta_{\hat{j}}^1\right) - v_{\hat{j}}\left(b, \theta_{\hat{j}}^1\right) \neq v_{\hat{j}}\left(a, \theta_{\hat{j}}^2\right) - v_{\hat{j}}\left(b, \theta_{\hat{j}}^2\right)$ and in addition $\left(\theta_{\hat{j}}^1, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) \in B\left(\left(\tilde{\theta}_{-\hat{i}}\right), r\right)$ and $\left(\theta_{\hat{j}}^2, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) \in B\left(\left(\tilde{\theta}_{-\hat{i}}\right), r\right)$.

Therefore,

$$q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = a \tag{8}$$

$$q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = b$$

By Lemma 5, the equations in (8) imply that

$$\tilde{t}_{\hat{j}}\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = \tilde{t}_{\hat{j}}\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = \alpha \tag{9}$$

$$\tilde{t}_{\hat{j}}\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = \tilde{t}_{\hat{j}}\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right) = \beta$$

Consider some profile in $\mathcal{D}_{-\hat{i}}^{-\hat{j}}$ in which $\delta_{\hat{j}}^{k} = 0$ for every $k \in P_{\hat{j}}$, and denote it by $\tilde{\delta}_{-\hat{i}}^{-\hat{j}}$.
Consider the profile $\left(\theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)$. By equation (7) we have that

$$v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{1}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \left(\delta_{\hat{i}}^{\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)\right) = v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{1}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \left(\delta_{\hat{i}}^{\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)\right) \tag{10}$$

for every $\delta_{\hat{i}}^{\hat{j}} \in M_1$, where $M_1 = \left[\underline{\delta_{\hat{i}}^{\hat{j}}}, \overline{\delta_{\hat{i}}^{\hat{j}}}\right] \setminus K_1$, where K_1 is a set with a measure zero that is contained in $\left[\underline{\delta_{\hat{i}}^{\hat{j}}}, \overline{\delta_{\hat{i}}^{\hat{j}}}\right]$.
Consider the profile $\left(\theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)$. By equation (7) we have that

$$v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{2}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \left(\delta_{\hat{i}}^{\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)\right) = v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{2}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \left(\delta_{\hat{i}}^{\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)\right) \tag{11}$$

for every $\delta_{\hat{i}}^{\hat{j}} \in M_2$, where $M_2 = \left[\underline{\delta_{\hat{i}}^{\hat{j}}}, \overline{\delta_{\hat{i}}^{\hat{j}}}\right] \setminus K_2$, where K_2 is a set with a measure zero which is contained in $\left[\underline{\delta_{\hat{i}}^{\hat{j}}}, \overline{\delta_{\hat{i}}^{\hat{j}}}\right]$.
Now $M_1 \cap M_2 = \left[\underline{\delta_{\hat{i}}^{\hat{j}}}, \overline{\delta_{\hat{i}}^{\hat{j}}}\right] \setminus K_1 \cup K_2 \neq \emptyset$ and therefore there exists $\tilde{\delta}_{\hat{i}}^{\hat{j}} \in M_1 \cap M_2$. We denote this profile by $\tilde{\delta} = \left(\tilde{\delta}_{\hat{i}}^{\hat{j}}, \tilde{\delta}_{-\hat{i}}^{-\hat{j}}\right)$. Equations (10) and (11), then imply that

$$\begin{aligned}
&\left\{v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{1}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \tilde{\delta}\right) - \left[v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{1}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{1}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \tilde{\delta}\right)\right]\right\} \\
&- \left\{v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{2}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{1}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \tilde{\delta}\right) - \left[v_{\hat{j}}\left(q\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \theta_{\hat{j}}^{2}\right) + t_{\hat{j}}\left(\left(\theta_{\hat{i}}^{2}, \theta_{\hat{j}}^{2}, \tilde{\theta}_{-\hat{i}-\hat{j}}\right), \tilde{\delta}\right)\right]\right\} = 0
\end{aligned} \tag{12}$$

since the expression in each wavy brackets is equal to zero. In addition, under the profile $\tilde{\delta}$ we have that $\tilde{\delta}_{\hat{j}}^{k} = 0$ for every $k \in P_{\hat{j}}$. Therefore, by equation (1) we have that $t_{\hat{j}}\left(\theta, \tilde{\delta}\right) = \tilde{t}_{\hat{j}}(\theta) + \tau_{\hat{j}}\left(\theta_{-\hat{j}}, \tilde{\delta}\right)$. Plugging this identity and the equations in (8) and (9) to equation (12) yields

$$(12)\qquad \left\{v_{\hat{j}}\left(a,\theta_{\hat{j}}^{1}\right)+\alpha+\tau_{\hat{j}}\left(\theta_{\hat{i}}^{1},\tilde{\theta}_{-\hat{i}-\hat{j}},\tilde{\delta}\right)-\left[v_{\hat{j}}\left(b,\theta_{\hat{j}}^{1}\right)+\beta+\tau_{\hat{j}}\left(\theta_{\hat{i}}^{2},\tilde{\theta}_{-\hat{i}-\hat{j}},\tilde{\delta}\right)\right]\right\}$$

$$-\left\{v_{\hat{j}}\left(a,\theta_{\hat{j}}^{2}\right)+\alpha+\tau_{\hat{j}}\left(\theta_{\hat{i}}^{1},\tilde{\theta}_{-\hat{i}-\hat{j}},\tilde{\delta}\right)-\left[v_{\hat{j}}\left(b,\theta_{\hat{j}}^{2}\right)+\beta+\tau_{\hat{j}}\left(\theta_{\hat{i}}^{2},\tilde{\theta}_{-\hat{i}-\hat{j}},\tilde{\delta}\right)\right]\right\}=$$

$$=v_{\hat{j}}\left(a,\theta_{\hat{j}}^{1}\right)-v_{\hat{j}}\left(b,\theta_{\hat{j}}^{1}\right)-\left(v_{\hat{j}}\left(a,\theta_{\hat{j}}^{2}\right)-v_{\hat{j}}\left(b,\theta_{\hat{j}}^{2}\right)\right)\neq 0$$

a contradiction. ■

A.2.3 Proof of Proposition 4

Assume that $q(\theta)$ is Bayesian implementable in IP. Then there exists a social choice function $(q(\theta), t_1(\theta,\delta), ..., t_n(\theta,\delta))$ such that for every $i \in I$, $\theta_i \in \Theta_i$, and $\delta_i \in \mathcal{D}_i$ we have

$$(1)\quad (\theta_i,\delta_i)\in \underset{\left(\hat{\theta}_i,\hat{\delta}_i\right)\in\Theta_i\times\mathcal{D}_i}{\arg\max}\, E_{\theta_{-i},\delta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)+t_i\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\hat{\delta}_i,\delta_{-i}\right)\right)+\right.$$

$$\left.\sum_{k\in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_k\right)+t_k\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\hat{\delta}_i,\delta_{-i}\right)\right)\right)\right]\right]$$

Consider an arbitrary profile $\tilde{\delta}_i$ and assume that $\hat{\delta}_i=\tilde{\delta}_i$; then (1) implies that for every $\theta_i\in\Theta_i$ we have

$$\theta_i\in\underset{\hat{\theta}_i\in\Theta_i}{\arg\max}E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)\right]+E_{\theta_{-i},\delta_{-i}}\left[\sum_{k\in P_i}\left[\tilde{\delta}_i^k\left(v_k\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_k\right)+t_k\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\tilde{\delta}_i,\delta_{-i}\right)\right)\right)\right]\right]$$

Define $\mu_i\left(\hat{\theta}_i\right):=E_{\theta_{-i},\delta_{-i}}\left[\sum_{k\in P_i}\left[\tilde{\delta}_i^k\left(v_k\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_k\right)+t_k\left(\left(\hat{\theta}_i,\theta_{-i}\right),\left(\tilde{\delta}_i,\delta_{-i}\right)\right)\right)\right]\right]$. Thus we get
$\theta_i\in\underset{\hat{\theta}_i\in\Theta_i}{\arg\max}\,E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)\right]+\mu_i\left(\hat{\theta}_i\right)$. For every $i\in I$, let $t_i^{'}(\theta)=\mu_i(\theta_i)$; then $\theta_i\in\underset{\hat{\theta}_i\in\Theta_i}{\arg\max}\,E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)+t_i^{'}\left(\hat{\theta}_i,\theta_{-i}\right)\right]$ for every $\theta_i\in\Theta_i$. Therefore $\left(q(\theta),t_1^{'}(\theta),...,t_n^{'}(\theta)\right)$ is Bayesian implementable in PV, and hence $q(\theta)$ is Bayesian implementable in PV. ■

A.2.4 Proof of Theorem 5

Let $(q(\theta), t_1(\theta), ..., t_n(\theta))$ be a Bayesian implementable social choice function in PV and let $IP \in P$. Define for every $i \in I$

$$(1)\quad t_i^{'}(\theta) \coloneqq t_i(\theta) - \sum_{j \in I \setminus \{i\}} E_{\tilde{\theta}_{-j}}\left[v_i\left(q\left(\theta_j, \tilde{\theta}_{-j}\right), \tilde{\theta}_i\right) + t_i\left(\theta_j, \tilde{\theta}_{-j}\right)\right] + \sum_{j \in I \setminus \{i\}} E_{\tilde{\theta}}\left[v_i\left(q\left(\tilde{\theta}\right), \tilde{\theta}_i\right) + t_i\left(\tilde{\theta}\right)\right]$$

In a IP setup, given the profile of functions $\left(t_i^{'}\right)_{i \in I}$, the expected utility of agent i as a function of his report is

$$(2)\quad E_{\theta_{-i}, \delta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i^{'}\left(\hat{\theta}_i, \theta_{-i}\right) + \sum_{k \in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k^{'}\left(\hat{\theta}_i, \theta_{-i}\right)\right)\right]\right] =$$

$$= E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i^{'}\left(\hat{\theta}_i, \theta_{-i}\right) + \sum_{k \in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k^{'}\left(\hat{\theta}_i, \theta_{-i}\right)\right)\right]\right]$$

According to the definition in (1) we get that for every $k \in P_i$,

$$E_{\theta_{-i}}\left[\left(v_k\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_k\right) + t_k^{'}\left(\hat{\theta}_i, \theta_{-i}\right)\right)\right] = E_{\tilde{\theta}}\left[v_k\left(q\left(\tilde{\theta}\right), \tilde{\theta}_i\right) + t_k\left(\tilde{\theta}\right)\right]$$

which is constant and equals the expected payoff of agent k in the PV setup when the profile of transfer functions is $(t_i(\theta))_{i \in I}$. In addition by the definition of $t_i^{'}$ we get that

$$E_{\theta_{-i}}\left[t_i^{'}\left(\hat{\theta}_i, \theta_{-i}\right)\right] = E_{\theta_{-i}}\left[t_i\left(\hat{\theta}_i, \theta_{-i}\right)\right]$$

Using these results we get that (2), i.e., the expected utility of agent i as a function of his report, equals

$$E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\hat{\theta}_i, \theta_{-i}\right)\right] + \sum_{k \in P_i} \delta_i^k \cdot E_{\tilde{\theta}}\left[v_k\left(q\left(\tilde{\theta}\right), \tilde{\theta}_i\right) + t_k\left(\tilde{\theta}\right)\right]$$

Since $(q(\theta), t_1(\theta), ..., t_n(\theta))$ is Bayesian implementable in PV,

$$\theta_i \in \arg\max_{\hat{\theta}_i \in \Theta_i} E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\hat{\theta}_i, \theta_{-i}\right)\right]$$

which implies that

$$\theta_i \in \arg\max_{\hat{\theta}_i \in \Theta_i} E_{\theta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) + t_i\left(\hat{\theta}_i, \theta_{-i}\right)\right] + \sum_{k \in P_i} \delta_i^k \cdot E_{\tilde{\theta}}\left[v_k\left(q\left(\tilde{\theta}\right), \tilde{\theta}_i\right) + t_k\left(\tilde{\theta}\right)\right]$$

and therefore

$$(\theta_i,\delta_i)\in\underset{(\hat{\theta}_i,\hat{\delta}_i)\in\Theta_i\times\mathcal{D}_i}{\arg\max} E_{\theta_{-i},\delta_{-i}}\left[v_i\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)+t_i^{'}\left(\hat{\theta}_i,\theta_{-i}\right)+\sum_{k\in P_i}\left[\delta_i^k\left(v_k\left(q\left(\hat{\theta}_i,\theta_{-i}\right),\theta_k\right)+t_k^{'}\left(\hat{\theta}_i,\theta_{-i}\right)\right)\right]\right]$$

Therefore, $\left(q(\theta),t_1^{'}(\theta),...,t_n^{'}(\theta)\right)$ is Bayesian implementable in IP. Now, for every $i\in I$, $t_i^{'}(\theta)$ equals $t_i(\theta)$ plus a function that does not depend on θ_i. Therefore, if $(t_i(\theta))_{i\in I}$ Bayesian implements $q(\theta)$ in PV then so does $\left(t_i^{'}(\theta)\right)_{i\in I}$, and if $(t_i(\theta))_{i\in I}$ implements $q(\theta)$ in dominant strategy in PV, then so does $\left(t_i^{'}(\theta)\right)_{i\in I}$. Moreover for every $i\in I$, $E_{\theta_{-i}}\left[t_i^{'}\left(\hat{\theta}_i,\theta_{-i}\right)\right]=E_{\theta_{-i}}\left[t_i\left(\hat{\theta}_i,\theta_{-i}\right)\right]$ and therefore $E_\theta[t_i(\theta)]=E_\theta\left[t_i^{'}(\theta)\right]$ for every $i\in I$ ■

References

[1] Bartling, B., and Netzer, N. An externality-robust auction: Theory and experimental evidence. *Games and Economic Behavior 97* (2016), 186–204.

[2] Bergstrom, T. Puzzles: Love and spaghetti, the opportunity cost of virtue. *The Journal of Economic Perspectives 3*, 2 (1989), 165–173.

[3] Bergstrom, T. C. Systems of benevolent utility functions. *Journal of Public Economic Theory 1*, 1 (1999), 71–100.

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

[5] Bikhchandani, S. Ex post implementation in environments with private goods. *Theoretical Economics 1* (2006), 369–393.

[6] Blanchet, D., and Fleurbaey, M. Selfishness, altruism and normative principles in the economic analysis of social transfers. *Handbook of the economics of giving, altruism and reciprocity 2* (2006), 1465–1503.

[7] Cooper, D., and Kagel, J. H. Other regarding preferences: a selective survey of experimental results. *Handbook of experimental economics 2* (2009).

[8] Dasgupta, P., and Maskin, E. Efficient auctions. *Quarterly Journal of Economics* (2000), 341–388.

[9] Desiraju, R., and Sappington, D. E. Equity and adverse selection. *Journal of Economics & Management Strategy 16*, 2 (2007), 285–318.

[10] Gershkov, A., Goeree, J. K., Kushnir, A., Moldovanu, B., and Shi, X. On the equivalence of bayesian and dominant strategy implementation. *Econometrica 81*, 1 (2013), 197–220.

[11] Gibbard, A. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society* (1973), 587–601.

[12] Goodin, R. E. Laundering preferences. *Foundations of social choice theory 75* (1986), 81–86.

[13] Groves, T. Incentives in teams. *Econometrica: Journal of the Econometric Society* (1973), 617–631.

[14] Harsanyi, J. C. Morality and the theory of rational behavior. *Social Research* (1977), 623–656.

[15] Heydenreich, B., Müller, R., Uetz, M., and Vohra, R. V. Characterization of revenue equivalence. *Econometrica 77*, 1 (2009), 307–316.

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

[17] Jehiel, P., and Moldovanu, B. Efficient design with interdependent valuations. *Econometrica 69*, 5 (2001), 1237–1259.

[18] Krishna, V., and Maenner, E. Convex potentials with an application to mechanism design. *Econometrica 69*, 4 (2001), 1113–1119.

[19] Lavi, R., Mu'Alem, A., and Nisan, N. Towards a characterization of truthful combinatorial auctions. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on* (2003), IEEE, pp. 574–583.

[20] Mishra, D., and Sen, A. Robert's theorem with neutrality: A social welfare ordering approach. *Games and Economic Behavior 75*, 1 (2012), 283–298.

[21] Perry, M., and Reny, P. J. An efficient auction. *Econometrica 70*, 3 (2002), 1199–1212.

[22] Roberts, K. The characterization of implementable choice rules. *Aggregation and revelation of preferences 12*, 2 (1979), 321–348.

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

[24] von Siemens, F. A. Heterogeneous social preferences, screening, and employment contracts. *Oxford Economic Papers* (2010), 499–522.

[25] Wilson, R. Game-theoretic analysis of trading processes. *Advances in Economic Theory Cambridge University Press* (1987), 33–70.