ROBUST MECHANISM DESIGN

By

Dirk Bergemann and Stephen Morris

April 2004

COWLES FOUNDATION DISCUSSION PAPER NO. 1421R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Robust Mechanism Design*

Dirk Bergemann† Stephen Morris‡

First Version: December 2000
This Version: April 2004

Abstract

The mechanism design literature assumes too much common knowledge of the environment among the players and planner. We relax this assumption by studying implementation on richer type spaces.

We ask when ex post implementation is equivalent to interim (or Bayesian) implementation for all possible type spaces. The equivalence holds in the case of *separable* environments; examples of separable environments arise (1) when the planner is implementing a social choice function (not correspondence); and (2) in a quasilinear environment with no restrictions on transfers. The equivalence fails in general, including in some quasilinear environments with budget balance.

In private value environments, ex post implementation is equivalent to dominant strategies implementation. The private value versions of our results offer new insights into the relation between dominant strategy implementation and Bayesian implementation.

Keywords: Mechanism Design, Common Knowledge, Universal Type Space, Interim Equilibrium, Ex-Post Equilibrium, Dominant Strategies.

Jel Classification: C79, D82

*This research is supported by NSF Grant #SES-0095321. We would like to thank the co-editor, three anonymous referees and seminar participants at many institutions for helpful comments. We thank Bob Evans for pointing out errors in earlier examples and Sandeep Baliga, Matt Jackson, Jon Levin, Bart Lipman, Eric Maskin, Zvika Neeman, Andrew Postlewaite, Ilya Segal and Tomas Sjöström for valuable discussions.

†Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511, dirk.bergemann@yale.edu.

‡Department of Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06511, stephen.morris@yale.edu.

"Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent it assumes other features to be common knowledge, such as one player's probability assessment about another's preferences or information.

I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems. Only by repeated weakening of common knowledge assumptions will the theory approximate reality." Wilson (1987)

1 Introduction

The theory of mechanism design helps us understand institutions ranging from simple trading rules to political constitutions. We can understand institutions as the solution to a well defined planner's problem of achieving some objective or maximizing some utility function subject to incentive constraints. But a common criticism of mechanism design theory is that the optimal mechanisms solving the well defined planner's problem seem unreasonably complicated. Researchers have often therefore restricted attention to mechanisms that are "more robust", or less sensitive to the assumed structure of the environment.[1] However, if the optimal solution to the planner's problem is too complicated or sensitive to be used in practice, it is presumably because the original description of the planner's problem was itself flawed. We would like to see if improved modelling of the planner's problem endogenously generates the "robust" features of mechanisms that researchers have been tempted to assume.

As suggested by Robert Wilson in the above quote, the problem is that we make too many implicit common knowledge assumptions in our description of the planner's problem.[2] The modelling strategy must be to first make explicit the implicit common knowledge assumptions, and then weaken them. The approach to modelling incomplete information introduced by Harsanyi (1967/1968) and formalized by Mertens and Zamir (1985) is ideally

[1] Discussions of this issue are an old theme in the mechanism design literature. Hurwicz (1972) discussed the need for "nonparametric" mechanisms (independent of parameters of the model). Wilson (1985) states that a desirable property of a trading rule is that it "does not rely on features of the agents' common knowledge, such as their probability assessments." Dasgupta and Maskin (2000) "seek auction rules that are independent of the details - such as functional forms or distribution of signals - of any particular application and that work well in a broad range of circumstances".

[2] An important paper of Neeman (2001) shows how rich type spaces can be used to relax implicit common knowledge assumptions in a mechanism design context. For other approaches to formalizing robust mechanism design, see Chung and Ely (2003), Duggan and Roberts (1997), Eliaz (2002), Hagerty and Rogerson (1987) and Lopomo (1998, 2000).

suited to this task. In fact, Harsanyi's work was intended to address the then prevailing criticism of game theory that the very description of a game embodied common knowledge assumptions that could never prevail in practise. Harsanyi argued that by allowing an agent's type to include his beliefs about the strategic environment, his beliefs about other agents' beliefs, and so on, any environment of incomplete information could be captured by a type space. With this sufficiently large type space (including all possible beliefs and higher order beliefs), it is true (tautologically) that there is common knowledge among the agents of each agent's set of possible types and each type's beliefs over the types of other agents. However, as a practical matter, applied economic analysis tends to assume much smaller type spaces than the universal type space, *and yet maintain the assumption that there is common knowledge among the agents of each agent's type space and each type's beliefs over the types of other agents.* In the small type space case, this is a very substantive restriction. There has been remarkably little work since Harsanyi checking whether analysis of incomplete information games in economics is robust to the implicit common knowledge assumptions built into small type spaces.[3] We will investigate the importance of these implicit common knowledge assumptions in the context of mechanism design.[4]

Formally, we fix a payoff environment, specifying a set of payoff types for each agent, a set of outcomes, utility functions for each agent and a social choice correspondence (SCC) mapping payoff type profiles into sets of acceptable outcomes. The planner *(partially) implements*[5] the social choice correspondence if there exists a mechanism and an equilibrium strategy profile of that mechanism such that equilibrium outcomes for every payoff type profile are acceptable according to the SCC.[6] This is sometimes referred to as Bayesian implementation, but since we do not have a common prior, we will call it interim implementation.

While holding fixed this environment, we can construct many type spaces, where an agent's type specifies both his payoff type and his belief about other agents' types. Crucially, there may be many types of an agent with the same payoff type. The larger the type space, the harder it will be to implement the social choice correspondence, and so the more "robust" the resulting mechanism will be. The smallest type space we can work with is the "payoff type space," where we set the possible types of each agent equal to the set

[3] Battigalli and Siniscalchi (2003), Morris and Shin (2003).

[4] Neeman (2001) argued that small type space assumptions are especially important in the full surplus extraction results of Cremer and McLean (1985).

[5] "Partial implementation" is sometimes called "truthful implementation" or "incentive compatible implementation." Since we look exclusively at partial implementation in this paper, we will write "implement" instead of "partially implement".

[6] In a companion paper, Bergemann and Morris (2004), we use the framework of this paper to look at full implementation, i.e., requiring that every equilibrium delivers an outcome consistent with the social choice correspondence.

of payoff types, and assume a common knowledge prior over this type space. This is the usual exercise performed in the mechanism design literature. The largest type space we can work with is the union of all possible type spaces that could have arisen from the payoff environment. This is equivalent to working with a "universal type space," in the sense of Mertens and Zamir (1985). There are many type spaces in between the payoff type space and the universal type space that are also interesting to study. For example, we can look at all payoff type spaces (so that the agents have common knowledge of a prior over payoff types but the mechanism designer does not); and we can look at type spaces where the common prior assumption holds.

In the face of a planner who does not know about agents' beliefs about others' payoff types, a recent literature has looked at mechanisms that implement the SCC in *ex post equilibrium* (see references in footnote 9). This requires that in a payoff type direct mechanism - where each agent is asked to report his payoff type - each agent has an incentive to tell the truth if he expects others to tell the truth, whatever their types turn out to be. In the special case of private values, ex post implementation is equivalent to dominant strategies implementation. If an SCC is ex post implementable, then it is clearly interim implementable on *every* type space, since the payoff type direct mechanism can be used to implement the SCC.

The converse is not always true. In Examples 1 and 2, ex post implementation is impossible. Nonetheless, interim implementation is possible on every type space. The gap arises because the planner may have the equilibrium outcome depend on the agents' higher order belief types, as well as their realized payoff type. The planner has no intrinsic interest in conditioning on non-payoff-relevant aspects of agents' types, but he is able to introduce slack in incentive constraints by doing so.

The main question we address in this paper is when the converse is true. A payoff environment is *separable* if the outcome space has a common component and a private value component for each agent. Each agent cares only about the common component and his own private component. The social choice correspondence picks a unique element from the common component and has a product structure over all components. In separable environments, interim implementation on all common prior payoff type spaces implies ex post implementation.[7] Whenever the social choice correspondence is a function, the environment has a separable representation (since we can make private value components degenerate). The other leading example of a separable environment is the problem of choosing an allocation when arbitrary transfers are allowed and agents have quasi-linear utility. If the allocation choice is a function but the planner does not care about the level and distribu-

[7] This result extends to all common prior full support type spaces in the quasilinear case and when the environment is compact.

tion of transfers, then we have a separable environment.

This result provides a strong foundation for using ex post equilibrium as a solution concept in separable environments. Since ex post implementation implies interim implementation on all type spaces (with or without the common prior or the payoff type restrictions), we also have equivalence between ex post implementation and interim implementation on all type spaces.

Thus for separable environments, the restriction to payoff type spaces is not important. But this is not true in general. In Example 3, we report a two agent quasi-linear environment where we add the balanced budget requirement: transfers must add up to zero. In this example, ex post implementation and interim implementation on all type spaces are both impossible, but implementation on all payoff type spaces is possible. As a leading example of an important economic non-separable environment, we look more generally at quasi-linear environments with budget balance. With two agents, there is an equivalence between ex post implementation and interim implementation on all type spaces. With at most two payoff types for each agent, there is the stronger equivalence between ex post implementation and interim implementation on all payoff type spaces. But with three or more agents with three or more types, equivalence between ex post implementation and interim implementation on all type spaces breaks down.

In private values environments, ex post implementation is equivalent to dominant strategies implementation. Our positive and negative results all have counterparts in private values environments. In particular, we (1) identify conditions when Bayesian implementation on all type spaces is equivalent to dominant strategies implementation, (2) give examples where the equivalence does not hold and (3) show how and when the equivalence may depend on type spaces richer than the payoff type space. While related questions have long been discussed in the implementation literature (e.g., Ledyard (1978) and Dasgupta, Hammond and Maskin (1979)) - we discuss the relation in detail in the concluding Section 6 - our questions have not been addressed even under private values.

The paper is organized as follows. Section 2 provides the setup, introduces the type spaces and provides the equilibrium notions. In Section 3 we present in some detail three examples which illustrate the role of type spaces in the implementation problem and point to the complex relationship between ex post implementation on the payoff type space and interim implementation on larger type spaces. In Section 4 we present equivalence results for separable social choice environments. The separable environment includes as special cases all social choice *functions* and the quasi-linear environment without a balanced budget requirement. Section 5 investigates the quasi-linear environment with a balanced budget requirement. We conclude with a discussion of further issues in Section 6.

2 Setup

2.1 Payoff Environment

We consider a finite set of agents $1, 2, ..., I$. Agent i's *payoff type* is $\theta_i \in \Theta_i$, where Θ_i is a finite set. We write $\theta \in \Theta = \Theta_1 \times ... \times \Theta_I$. There is a set of outcomes Y. Each agent has utility function $u_i : Y \times \Theta \to \mathbb{R}$. A social correspondence is a mapping $F : \Theta \to 2^Y \backslash \emptyset$. If the true payoff type profile is θ, the planner would like the outcome to be an element of $F(\theta)$.

An important special case - studied in some of our examples and results - is a *quasi-linear environment* where the set of outcomes Y has the product structure $Y = Y_0 \times Y_1 \times \cdots \times Y_I$, where $Y_1 = Y_2 = .. = Y_I = \mathbb{R}$, and a utility function:

$$u_i(y, \theta) = u_i(y_0, y_1, ..., y_I, \theta) \triangleq v_i(y_0, \theta) + y_i$$

which is linear in y_i for every agent i. The planner is concerned only about choosing an "allocation" $y_0 \in Y_0$ and does not care about transfers. Thus there is a function $f_0 : \Theta \to Y_0$ and

$$F(\theta) = \{(y_0, y_1, ..., y_I) \in Y : y_0 = f_0(\theta)\}.$$

Throughout the paper, this environment is fixed and informally understood to be common knowledge. We allow for interdependent types - one agent's payoff from a given outcome depends on other agents' payoff types. The payoff type profile is understood to contain all information that is relevant to whether the planner achieves his objective or not. For example, we do not allow the planner to trade off what happens in one state with what happens in another state. For the latter reason, this setup is somewhat restrictive. However, it incorporates many classic problems such as the efficient allocation of an object or the efficient provision of a public good.

2.2 Type Spaces

While maintaining that the above payoff environment is common knowledge, we want to allow for agents to have all possible beliefs and higher order beliefs about other agents' types. A flexible framework for modelling such beliefs and higher order beliefs are type spaces.

A type space is a collection

$$\mathcal{T} = \left(T_i, \widehat{\theta}_i, \widehat{\pi}_i\right)_{i=1}^{I}.$$

Agent i's *type* is $t_i \in T_i$. A type of agent i must include a description of his payoff type. Thus there is a function

$$\widehat{\theta}_i : T_i \to \Theta_i,$$

with $\widehat{\theta}_i(t_i)$ being agent i's *payoff type* when his type is t_i. A type of agent i must also include a description of his beliefs about the types of the other agent. Write $\Delta(Z)$ for the space of probability measures on the Borel field of a measurable space Z. The belief of type t_i of agent i is a function

$$\widehat{\pi}_i : T_i \to \Delta(T_{-i}),$$

with $\widehat{\pi}_i(t_i)$ being agent i's *belief type* when his type is t_i. Thus $\widehat{\pi}_i(t_i)[E]$ is the probability that type t_i of agent i assigns to other agents' types, t_{-i}, being an element of a measurable set $E \subseteq T_{-i}$. In the special case where each T_j is finite, we will abuse notation slightly by writing $\widehat{\pi}_i(t_i)[t_{-i}]$ for the probability that type t_i of agent i assigns to other agents having types t_{-i}.

2.3 Solution Concepts

Fix a payoff environment and a type space $\mathcal{T}$. A mechanism specifies a message set for each agent and a mapping from message profiles to outcomes. Social choice correspondence F is interim implementable if there exists a mechanism and an interim (or Bayesian) equilibrium of that mechanism such that outcomes are consistent with F. However, by the revelation principle, we can restrict attention to truth-telling equilibria of direct mechanisms.[8] A direct mechanism is a function $f : T \to Y$.

Definition 1 *A direct mechanism $f : T \to Y$ is interim incentive compatible on type space $\mathcal{T}$ if*

$$\int\limits_{t_{-i}\in T_{-i}} u_i\left(f(t_i,t_{-i}),\widehat{\theta}(t_i,t_{-i})\right)d\widehat{\pi}_i(t_i) \geq \int\limits_{t_{-i}\in T_{-i}} u_i\left(f(t_i',t_{-i}),\widehat{\theta}(t_i,t_{-i})\right)d\widehat{\pi}_i(t_i)$$

for all i, $t \in T$ and $t_i' \in T_i$.

The notion of interim incentive compatibility is often referred to as Bayesian incentive compatibility. We use the former terminology as there need not be a common prior on the type space.

Definition 2 *A direct mechanism $f : T \to Y$ on $\mathcal{T}$ achieves F if*

$$f(t) \in F\left(\widehat{\theta}(t)\right)$$

for all $t \in T$.

[8] See Myerson (1991), Chapter 6.

It should be emphasized that a direct mechanism f can prescribe varying allocations for a given payoff profile θ as different types, t and t', may have an identical payoff profile $\theta = \widehat{\theta}(t) = \widehat{\theta}(t')$.

Definition 3 *A social choice correspondence F is interim implementable on $\mathcal{T}$ if there exists $f : T \to Y$ such that f is interim incentive compatible on $\mathcal{T}$ and f achieves F.*

We will be interested in comparing interim implementation with the stronger solution concept of ex post implementation. Ex post implementation uses the stronger solution concept of ex post equilibrium for incomplete information games.[9] By the revelation principle, it is again enough to verify ex post incentive compatibility.

Definition 4 *A direct mechanism $f : \Theta \to Y$ is ex post incentive compatible if, for all i and $\theta \in \Theta$,*

$$u_i(f(\theta), \theta) \geq u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta\right),$$

for all $\theta_i' \in \Theta_i$.

The notion of ex post incentive compatibility requires agent i to prefer truthtelling at θ if all the other agents also report truthfully. In contrast the notion of dominant strategy implementation requires agent i to prefer truthtelling for all possible reports by the other agents, truthtelling or not.

Definition 5 *A direct mechanism $f : \Theta \to Y$ is dominant strategies incentive compatible if, for all i and $\theta \in \Theta$,*

$$u_i\left(f\left(\theta_i, \theta_{-i}'\right), \theta\right) \geq u_i\left(f\left(\theta'\right), \theta\right),$$

for all $\theta' \in \Theta$.

If there are private values (i.e., each $u_i(y, \theta)$ depends on θ only through θ_i), then ex post incentive compatibility is equivalent to dominant strategies incentive compatibility.

Definition 6 *A social choice correspondence F is ex post implementable if there exists $f : \Theta \to Y$ such that f is ex post incentive compatible and $f(\theta) \in F(\theta)$ for all $\theta \in \Theta$.*

[9] Ex post incentive compatibility was discussed as "uniform incentive compatibility" by Holmstrom and Myerson (1983). Ex post equilibrium is increasingly studied in game theory (see Kalai (2002)) and is often used in mechanism design as a more robust solution concept (Cremer and McLean (1985)). A recent literature on interdependent value environments has obtained positive and negative results using this solution concept: Dasgupta and Maskin (2000), Bergemann and Valimaki (2002), Perry and Reny (2002), Jehiel and Moldovanu (2001) and Jehiel et al. (2004).

2.4 Questions

We are interested in characterizing interim incentive compatibility on different type spaces. We first introduce some key properties of type spaces. A type space $\mathcal{T}$ is a *payoff type space* if each $T_i = \Theta_i$ and each $\widehat{\theta}_i$ is the identity map. Type space $\mathcal{T}$ is *finite* if each T_i is finite. Finite type space $\mathcal{T}$ has *full support* if $\widehat{\pi}_i(t_i)[t_{-i}] > 0$ for all i and t. Finite type space $\mathcal{T}$ satisfies the *common prior assumption* (with prior p) if there exists $p \in \Delta(T)$ such that

$$\sum_{t_{-i} \in T_{-i}} p(t_i, t_{-i}) > 0 \text{ for all } i \text{ and } t_i$$

and

$$\widehat{\pi}_i(t_i)[t_{-i}] = \frac{p(t_i, t_{-i})}{\sum\limits_{t'_{-i} \in T_{-i}} p\left(t_i, t'_{-i}\right)}.$$

The standard approach in the mechanism design literature is to restrict attention to a common prior payoff type space (perhaps with full support). Thus it is assumed that there is common knowledge among the agents of a common prior over the payoff types. A payoff type space can be thought of the smallest type space embedding the payoff environment described above. Restricting attention to a full support, common prior, payoff type space is *with* loss of generality. We want to relax the implicit common knowledge assumptions embodied in those restrictions by asking the following progressively tougher questions about interim implementability:

- Is F interim implementable on all full support common prior payoff type spaces?
- Is F interim implementable on all common prior payoff type spaces?
- Is F interim implementable on all common prior type spaces?
- Is F interim implementable on all type spaces?

By requiring that F be interim implementable on all type spaces, we are asking for a mechanism that can implement F with no common knowledge assumptions beyond those in the specification of the payoff environment. If we constructed a universal type space for the payoff environment, that universal type space would be an example of a type space and thus interim implementability on all type spaces would imply interim implementability on the universal type space. We discuss the relation between our approach and the universal type space in more detail in Section 6.

We will see that relaxing common knowledge assumptions makes a difference. In particular, we will show that while the common prior assumption is not important and the full support assumption does not play a big role,[10] the payoff type space restriction *is* important. In example 3 in the next section, it is possible to interim implement on any payoff type space (with or without the common prior) but not all type spaces. We are especially interested in the relation between the ex post implementability of F and interim implementability. In Sections 4 and 5, we provide sufficient conditions for ex post implementability to be equivalent to interim implementability on all type spaces. But Examples 1 and 2 in the next section show that it is possible to find social choice correspondences that are interim implementable on any type space but not ex post implementable.

3 Examples

This section presents three examples illustrating the relationship between interim implementation on different type spaces and ex post implementation.

The first two examples exhibit social choice correspondences that are interim implementable on all type spaces, but are not ex post implementable. The first example is very simple, but relies on (i) a restriction to deterministic allocations, (ii) a social choice correspondence that depends on only one agent's payoff type; and (iii) interdependent types. In the second example, we show how to dispense with all three features. Since this second example has private values, we thus have an example where dominant strategies implementation is impossible but interim implementation is possible on any type space.

The third example exhibits a social choice correspondence that is interim implementable on all payoff type spaces (with or without the common prior) but is not interim implementable on all type spaces. The social choice correspondence represents efficient allocations in a quasi-linear environment with a balanced budget requirement. As such it also illustrates some of the results presented in later in Section 5 on social choice problems with a balanced budget.

3.1 F is Interim Implementable on All Type Spaces but not Ex Post Implementable

Example 1 There are two agents. Each agent has two possible types: $\Theta_1 = \{\theta_1, \theta_1'\}$ and $\Theta_2 = \{\theta_2, \theta_2'\}$. There are three possible allocations: $Y = \{a, b, c\}$. The payoffs of the two agents are given by the following tables (each box describes agent 1's payoff, then agent 2's

[10] However, different type space assumptions will be important for different questions. The full support assumption is crucial when we look at full implementation (see Bergemann and Morris (2004)) and the common prior assumption is important when we look at revenue maximization (see Bergemann, Morris and Segal (2004)).

payoff):

a	θ_2	θ_2'
θ_1	$1,0$	$-1,2$
θ_1'	$0,0$	$0,0$

b	θ_2	θ_2'
θ_1	$-1,2$	$1,0$
θ_1'	$0,0$	$0,0$

c	θ_2	θ_2'
θ_1	$0,0$	$0,0$
θ_1'	$1,1$	$1,1$

The social choice correspondence is given by

F	θ_2	θ_2'
θ_1	$\{a,b\}$	$\{a,b\}$
θ_1'	$\{c\}$	$\{c\}$

These choices are maximizers of the sum of agents' utility. The key feature of this example is that the agents agree about the optimal choice when agent 1 is type θ_1'; when agent 1 is type θ_1, they agree that it is optimal to choose either a or b. But each agent has strict and opposite preferences over outcomes a and b: 1 strictly prefers a when 2's type is θ_2, while 2 strictly prefers a when his type is θ_2'.

We now show - by contradiction - that this correspondence is not ex post implementable. If F was implementable, we would have to have c chosen at profiles (θ_1', θ_2) and (θ_1', θ_2'); and either a or b chosen at profiles (θ_1, θ_2) and (θ_1, θ_2'). But in order for type θ_1 to have an incentive to tell the truth when he is sure that agent 2 is type θ_2, we must have a chosen at profile (θ_1, θ_2); and in order for type θ_1 to have incentive to tell the truth when he is sure that agent 2 is type θ_2', we must have b chosen at profile (θ_1, θ_2'). But if a is chosen at profile (θ_1, θ_2) and b is chosen at profile (θ_1, θ_2'), then both types of agent 2 will have an incentive to misreport their types when they are sure that agent 1 is type θ_1.

However, the correspondence is interim implementable on any type space using the very simple mechanism of letting agent 1 pick the outcome. There is always an equilibrium of this mechanism where agent 1 will pick outcome a if his type is θ_1 and he assigns probability at least $\frac{1}{2}$ to the other agent being type θ_2; agent 1 will pick outcome b if his type is θ_1 and he assigns probability less than $\frac{1}{2}$ to the other agent being type θ_2; and agent 1 will pick outcome c if his type is θ_1'. By allowing the mechanism to depend on agent 1's beliefs about agent 2's type (something the planner does not care about intrinsically), the planner is able to relax incentive constraints that he cares about.

The failure of ex post implementation in this example relied on the assumption that only pure outcomes were chosen. This restriction can easily be dropped at the expense of adding a third payoff type for agent 1, so that the binding ex post incentive constraint for agent 1 is with a different type and outcome depending on 2's type. Example 1 also had the social choice correspondence depending only on agent 1's payoff type and had interdependent values. We can mechanically change these two assumptions by letting the planner want different outcomes depending on agent 2's type. Now instead of having agent 1's utility depend on agent 2's type, it can depend on the planner's refined choice.

Example 2 There are two agents. Agent 1 has three possible types, $\Theta_1 = \{\theta_1, \theta_1', \theta_1''\}$, and agent 2 has two possible types, $\Theta_2 = \{\theta_2, \theta_2'\}$. There are eight possible pure allocations, $\{a, b, c, d, a', b', c', d'\}$, and lotteries are allowed, so $Y = \Delta(\{a, b, c, d, a', b', c', d'\})$. The private value payoffs of agent 1 are given by the following table:

u_1	a	b	c	d	a'	b'	c'	d'
θ_1	1	-1	$\frac{1}{2}$	-1	-1	1	-1	$\frac{1}{2}$
θ_1'	0	0	1	0	0	0	1	0
θ_1''	0	0	0	1	0	0	0	1

The private value payoffs of agent 2 are given by the following table:

u_2	a	b	c	d	a'	b'	c'	d'
θ_2	0	1	0	0	0	1	-1	-1
θ_2'	1	0	-1	-1	1	0	0	0

The social choice correspondence F is described by the following table.[11]

	θ_2	θ_2'
θ_1	$\{a, b\}$	$\{a', b'\}$
θ_1'	$\{c\}$	$\{c'\}$
θ_1''	$\{d\}$	$\{d'\}$

We now show - by contradiction - that this correspondence is not ex post implementable. Let q be the probability that a is chosen at profile (θ_1, θ_2) and let q' be the probability that a' is chosen at profile (θ_1, θ_2'). In order for type θ_1 to have an incentive to tell the truth (and not report himself to be type θ_1') when he is sure that agent 2 is type θ_2, we must have

$$q - (1 - q) \geq \frac{1}{2} \quad \Leftrightarrow \quad q \geq \frac{3}{4}. \tag{1}$$

In order for type θ_1 to have an incentive to tell the truth (and not report himself to be type θ_1'') when he is sure that agent 2 is type θ_2', we must have

$$-q' + (1 - q') \geq \frac{1}{2} \quad \Leftrightarrow \quad q' \leq \frac{1}{4}. \tag{2}$$

[11] The SCC F in this example is not ex post Pareto efficient at (θ_1, θ_2) and (θ_1, θ_2'), as b' and a, respectively, Pareto dominate b and a', respectively. We choose this example for the simplicity of its payoffs. Yet, we have constructed examples with the same number of agents, states and allocations such that the SCC F is ex ante Pareto efficient and interim implementable on all type spaces, but not ex post, and a fortiori, not dominant strategy implementable.

But in order for agent 2 to have an incentive to tell the truth when he is type θ_2 and he is sure that agent 1 is type θ_1, we must have

$$1 - q \geq 1 - q';$$

thus

$$q' \geq q. \tag{3}$$

However, (1), (2) and (3) generate a contradiction, so ex post implementation is not possible.

But it is straightforward to implement on any interim type space. Consider the following indirect mechanism for any arbitrary type space where individual 1 chooses a message $m_1 \in \{m_1^1, m_1^2, m_1^3, m_1^4\}$ and individual 2 chooses a message $m_2 \in \{m_2^1, m_2^2\}$ and let outcomes be chosen as follows:

	m_2^1	m_2^2
m_1^1	a	a'
m_1^2	b	b'
m_1^3	c	c'
m_1^4	d	d'

There is always an equilibrium where type θ_1 of agent 1 sends message m_1^1 if he believes agent 2 is type θ_2 with probability at least $\frac{1}{2}$ and message m_1^2 if he believes agent 2 is type θ_2 with probability less than $\frac{1}{2}$; type θ_1' always sends message m_1^3; and type θ_1'' always sends message m_1^4. Type θ_2 of agent 2 sends message m_2^1 and type θ_2' sends message m_2^2, and this strategy is a dominant strategy for agent 2.

This private values example has the feature that dominant strategies implementation is impossible but interim implementation is possible on any type space, and seems to be the first example in the literature noting this possibility.[12]

As we will see in the next section, a necessary feature of the example is that we have a social choice correspondence (not function) that we are trying to implement. In the example, it was further key that there were aspects of the allocation that the planner did not care about but the agents did. In the example, this may look a little contrived but note that this a natural feature of quasi-linear environments where the planner wants to maximize the total welfare of agents. We will next present a quasi-linear utility example that exploits this feature.

[12]It is often noted that in public good problems with budget balance, dominant strategies implementation is impossible while Bayesian implementation is possible. However, the positive Bayesian implementation results (d'Aspremont and Gerard-Varet (1979) and d'Aspremont, Cremer and Gerard-Varet (1995, 2002)) hold only for "generic" priors on a fixed type space, not for all type spaces in our sense. They provide examples showing that Bayesian implementation fails for some type spaces.

3.2 F is Interim Implementable on All Naive Type Spaces but not Interim Implementable on All Type Spaces

Example 3 This example has two agents, denoted by 1 and 2. Agent 1 has three possible payoff types, $\Theta_1 = \{\theta_1, \theta_1', \theta_1''\}$, and agent 2 has two possible payoff types, $\Theta_2 = \{\theta_2, \theta_2'\}$. The set of feasible "allocations" is given by:

$$Y_0 = \{a, b, c, d\}.$$

The agents' gross utilities from the allocations, $v_1(y_0, \theta)$ and $v_2(y_0, \theta)$, respectively, are given by:

a	θ_2	θ_2'
θ_1	$0, 2$	$0, 2$
θ_1'	$-4, 0$	$1, 0$
θ_1''	$-4, 0$	$-4, 0$

b	θ_2	θ_2'
θ_1	$0, 0$	$0, 0$
θ_1'	$0, 2$	$0, 0$
θ_1''	$-4, 0$	$0, 0$

c	θ_2	θ_2'
θ_1	$0, 0$	$-4, 0$
θ_1'	$0, 0$	$0, 2$
θ_1''	$0, 0$	$0, 0$

d	θ_2	θ_2'
θ_1	$-4, 0$	$-4, 0$
θ_1'	$1, 0$	$-4, 0$
θ_1''	$0, 2$	$0, 2$

The planner wants the allocation $y_0 \in Y_0$ to maximize the sum of the agents' utilities at every type profile θ; thus he wants the allocation to depend to depend on type profile θ according to the function f_0 described in the following table:

f_0	θ_2	θ_2'
θ_1	a	a
θ_1'	b	c
θ_1''	d	d

(4)

In addition, balanced budget transfers are possible. Thus the planner must choose $(y_0, y_1, y_2) \in Y_0 \times \mathbb{R}^2$, with $y_1 + y_2 = 0$. Each agent has quasi-linear utility, so agent i's utility from (y_0, y_1, y_2) in payoff profile θ is $v_i(y_0, \theta) + y_i$. The planner maximizes the sum of utilities and so does not care about transfers; thus

$$F(\theta) = \left\{(y_0, y_1, y_2) \in Y_0 \times \mathbb{R}^2 : y_0 = f_0(\theta) \text{ and } y_2 = -y_1\right\}.$$

We first make a few observations regarding the ex post incentive constraints for truthtelling with *zero* transfers. Agent 1 always values the efficient alternatives at 0. The critical type for agent 1 is θ_1', where he values an inefficient alternative, either d or a (depending on the payoff type of agent 2 being θ_2 or θ_2'), at 1, and thus higher than the efficient alternative at that type profile. The remaining negative entries, -4, for agent 1 simply ensure that no other incentive constraints become relevant. Agent 2 always values the efficient allocation at 2, and every inefficient allocation at 0.

It is straightforward to establish that ex post implementation with balanced transfers is not feasible. Writing $f_i(\theta)$ for the transfer received by i at payoff type profile θ, we have the following ex post incentive constraints for agent 1:

$$v_1\left(f_0\left(\theta_1,\theta_2\right),\left(\theta_1,\theta_2\right)\right)+f_1\left(\theta_1,\theta_2\right)\geq v_1\left(f_0\left(\theta_1',\theta_2\right),\left(\theta_1,\theta_2\right)\right)+f_1\left(\theta_1',\theta_2\right),$$

$$v_1\left(f_0\left(\theta_1',\theta_2\right),\left(\theta_1',\theta_2\right)\right)+f_1\left(\theta_1',\theta_2\right)\geq v_1\left(f_0\left(\theta_1'',\theta_2\right),\left(\theta_1',\theta_2\right)\right)+f_1\left(\theta_1'',\theta_2\right),$$

and

$$v_1\left(f_0\left(\theta_1'',\theta_2'\right),\left(\theta_1'',\theta_2'\right)\right)+f_1\left(\theta_1'',\theta_2'\right)\geq v_1\left(f_0\left(\theta_1',\theta_2'\right),\left(\theta_1'',\theta_2'\right)\right)+f_1\left(\theta_1',\theta_2'\right),$$

$$v_1\left(f_0\left(\theta_1',\theta_2'\right),\left(\theta_1',\theta_2'\right)\right)+f_1\left(\theta_1',\theta_2'\right)\geq v_1\left(f_0\left(\theta_1,\theta_2'\right),\left(\theta_1',\theta_2'\right)\right)+f_1\left(\theta_1,\theta_2'\right),$$

Inserting the gross utilities $v_1(\cdot,\cdot)$, we can write the above set of inequalities as follows

$$f_1\left(\theta_1,\theta_2\right)\geq f_1\left(\theta_1',\theta_2\right)\geq f_1\left(\theta_1'',\theta_2\right)+1, \tag{5}$$

and

$$f_1\left(\theta_1'',\theta_2'\right)\geq f_1\left(\theta_1',\theta_2'\right)\geq f_1\left(\theta_1,\theta_2'\right)+1. \tag{6}$$

Next we consider the ex post incentive constraints for agent 2, at θ_1 and θ_1'', respectively. Here the social choice mapping prescribes allocations constant in the reported type profile of agent 2 and ex post incentive compatibility hence requires constant transfers as well, or $f_2\left(\theta_1,\theta_2\right)=f_2\left(\theta_1,\theta_2'\right)$ and $f_2\left(\theta_1'',\theta_2\right)=f_2\left(\theta_1'',\theta_2'\right)$. Using the balanced budget requirement by writing $f_2(\theta)=-f_1(\theta)$, we thus obtain

$$f_1\left(\theta_1,\theta_2\right)=f_1\left(\theta_1,\theta_2'\right),$$

and

$$f_1\left(\theta_1'',\theta_2\right)=f_1\left(\theta_1'',\theta_2'\right),$$

which leads to a contradiction with the inequalities (5) and (6).

Despite the failure of ex post implementation, we now show that we can satisfy the interim incentive compatibility conditions for every prior on the payoff type space. The sole determinant of the appropriate transfers is the belief of agent 1 with payoff type θ_1'. If type θ_1' assigns probability at least $\frac{1}{2}$ to agent 2 being of payoff type θ_2, then the following transfers to agent 1 (and corresponding balanced budget transfers for agent 2) are interim incentive compatible:

$$\begin{array}{cc} f_1\left(\theta_1,\theta_2\right)=0 & f_1\left(\theta_1,\theta_2'\right)=0 \\ f_1\left(\theta_1',\theta_2\right)=0 & f_1\left(\theta_1',\theta_2'\right)=-1 \\ f_1\left(\theta_1'',\theta_2\right)=-1 & f_1\left(\theta_1'',\theta_2'\right)=-1 \end{array} \tag{7}$$

Conversely, if type θ_1' assigns probability less than $\frac{1}{2}$ to the other agent being of payoff type θ_2, then the following transfers to agent 1 are interim incentive compatible:

$$\begin{array}{cc} f_1(\theta_1,\theta_2) = -1 & f_1(\theta_1,\theta_2') = -1 \\ f_1(\theta_1',\theta_2) = -1 & f_1(\theta_1',\theta_2') = 0 \\ f_1(\theta_1'',\theta_2) = 0 & f_1(\theta_1'',\theta_2') = 0 \end{array} \tag{8}$$

By symmetry of the payoffs, it will suffice to verify the incentive compatibility conditions for the first case. We first observe that all the ex post incentive constraints hold except for agent 1 at type profile $\theta_1'\theta_2'$, where he has a profitable deviation by misreporting himself to be of type θ_1. Suppose then that type θ_1' assigns probability p to the other agent being type θ_2. His expected payoff to truth-telling is

$$p(0+0) + (1-p)(0-1) = -(1-p)$$

while his expected payoff to mis-reporting type θ_1 is

$$p(-4+0) + (1-p)(1+0) = 1-5p,$$

and his expected payoff to mis-reporting type θ_1'' is given by

$$p(1-1) + (1-p)(-4-1) = -5(1-p).$$

Thus truth-telling is optimal as long as

$$-(1-p) \geq 1-5p \quad \Leftrightarrow \quad p \geq \frac{1}{3}. \tag{9}$$

The second set of transfers, described in (8), offer interim incentive compatibility for agent 1 provided that $p \leq \frac{2}{3}$. As either of the above transfer schemes satisfies the ex post incentive constraints of agent 2, it follows for every belief p by type θ_1', we can find interim incentive compatible transfers and hence F is interim implementable for all payoff type spaces.

However, on richer type spaces than the payoff type space, there may be many types with payoff type θ_1', some of whom are sure that the other agent is type θ_2 while others are sure that he is type θ_2'. That is the idea behind the following example of a "complete information" type space where F cannot be interim implemented. We consider the following type space:

	t_2^1	t_2^2	t_2^3	t_2^4	t_2^5	t_2^6	
t_1^1	$\frac{1}{6}$	0	0	0	0	0	θ_1
t_1^2	0	$\frac{1}{6}$	0	0	0	0	θ_1'
t_1^3	0	0	$\frac{1}{6}$	0	0	0	θ_1''
t_1^4	0	0	0	$\frac{1}{6}$	0	0	θ_1''
t_1^5	0	0	0	0	$\frac{1}{6}$	0	θ_1'
t_1^6	0	0	0	0	0	$\frac{1}{6}$	θ_1
	θ_2	θ_2	θ_2	θ_2'	θ_2'	θ_2'	

Thus there are six types for each agent, t_1^k and t_2^l. The entries in the cell describe the probabilities of the common prior, which puts all probability mass on the diagonal. The payoff type corresponding to each type appears at the end of the row/column corresponding to that type. Thus, for example, type t_1^3 of agent 1 has payoff type θ_1'' and believes that agent 2 has a payoff type θ_2 with probability one. It is in this sense, that we speak of complete information. We require that F is implemented even at "impossible" (zero probability) type profiles, but we could clearly adapt the example to have small probabilities off the diagonal.

Our impossibility argument will depend only on what happens at twelve critical type profiles: the diagonal profiles and the type profiles where agent 1 with type t_1^k claims to be one type higher, or t_1^{k+1}, and agent 2 with type t_2^l claims to be one type lower, or t_1^{l-1}. In the next table, we note which allocation must occur at these twelve profiles if F is to be implemented.

	t_2^1	t_2^2	t_2^3	t_2^4	t_2^5	t_2^6	
t_1^1	a					a	θ_1
t_1^2	b	b					θ_1'
t_1^3		d	d				θ_1''
t_1^4			d	d			θ_1''
t_1^5				c	c		θ_1'
t_1^6					a	a	θ_1
	θ_2	θ_2	θ_2	θ_2'	θ_2'	θ_2'	

We observe that the incentive constraints for agent 1 and agent 2 form jointly a cycle through the type space. We write y_{kl} for the transfer of agent 1 when the type profile is $t = \left(t_1^k, t_2^l\right)$. The incentive constraints corresponding to types t_1^k mis-reporting to be type

t_1^{k+1} (modulo 6) imply (for $k = 1, 2, .., 6$ respectively):

$$\begin{aligned} 0 + y_{11} &\geq 0 + y_{21} \\ 0 + y_{22} &\geq 1 + y_{32} \\ 0 + y_{33} &\geq 0 + y_{43} \\ 0 + y_{44} &\geq 0 + y_{54} \\ 0 + y_{55} &\geq 1 + y_{65} \\ 0 + y_{66} &\geq 0 + y_{16} \end{aligned} \tag{10}$$

The incentive constraints corresponding to types t_2^l mis-reporting to be type t_2^{l-1} imply, using the balanced budget to write the transfers to agent 2 as the negatives of agent 1, (for $l = 1, 2, .., 6$ respectively):

$$\begin{aligned} 2 - y_{11} &\geq 2 - y_{16} \\ 2 - y_{22} &\geq 2 - y_{21} \\ 2 - y_{33} &\geq 2 - y_{32} \\ 2 - y_{44} &\geq 2 - y_{43} \\ 2 - y_{55} &\geq 2 - y_{54} \\ 2 - y_{66} &\geq 2 - y_{65} \end{aligned} \tag{11}$$

The inequalities (10) and (11) have a very simply structure. With very few exceptions, the payoffs appearing on the lhs and rhs of the inequalities are identical and only the transfers differ. These inequalities are generated either by true and misreported types which induce only different transfer decisions but identical allocational decisions. The exceptions are the second and fifth inequality of agent 1, where a misreported type also leads to a different allocational decision. Re-arranging the inequalities, we obtain

$$\begin{array}{rclcrcl} 0 & \geq & y_{21} - y_{11}, & \quad & 0 & \geq & y_{11} - y_{16} \\ -1 & \geq & y_{32} - y_{22}, & & 0 & \geq & y_{22} - y_{21} \\ 0 & \geq & y_{43} - y_{33}, & & 0 & \geq & y_{33} - y_{32} \\ 0 & \geq & y_{54} - y_{44}, & & 0 & \geq & y_{44} - y_{43} \\ -1 & \geq & y_{65} - y_{55}, & & 0 & \geq & y_{55} - y_{54} \\ 0 & \geq & y_{16} - y_{66}, & & 0 & \geq & y_{66} - y_{65} \end{array}$$

When we sum these twelve constraints, the transfers on the right hand side of the inequalities cancel out and we are left with the desired contradiction for any arbitrary choice of probabilities, namely $-2 \geq 0$. The transfers cancelled out because the set of incentive constraints for agent 1 and agent 2 formed jointly a cycle through the type space.

4 Separable Environments

We now present general results about the relationship between ex post implementability and interim implementability on different type spaces. The first result is an immediate implication from the definition of ex post equilibrium.

Proposition 1 *If F is ex post implementable, then F is interim implementable on any type space.*

PROOF: If F is ex post implementable, then by hypothesis there exists $f^* : \Theta \to Y$ with $f^*(\theta) \in F(\theta)$ for all θ, such that for all i, all θ and all θ_i':

$$u_i\left(f^*(\theta), \theta\right) \geq u_i\left(f^*\left(\theta_i', \theta_{-i}\right), \theta\right)$$

Consider then an arbitrary type space $\mathcal{T}$ and the direct mechanism $f : T \to Y$ with $f(t) = f^*\left(\widehat{\theta}(t)\right)$. Incentive compatibility now requires

$$\begin{aligned} t_i &\in \arg\max_{t_i' \in T_i} \int_{t_{-i} \in T_{-i}} u_i\left(f\left(t_i', t_{-i}\right), \left(\widehat{\theta}_i(t_i), \widehat{\theta}_{-i}(t_{-i})\right)\right) d\widehat{\pi}_i(t_i) \\ &= \arg\max_{t_i' \in T_i} \int_{t_{-i} \in T_{-i}} u_i\left(f^*\left(\widehat{\theta}_i(t_i'), \widehat{\theta}_{-i}(t_{-i})\right), \left(\widehat{\theta}_i(t_i), \widehat{\theta}_{-i}(t_{-i})\right)\right) d\widehat{\pi}_i(t_i). \end{aligned}$$

This requires that

$$\begin{aligned} \widehat{\theta}_i(t_i) &= \arg\max_{\theta_i \in \Theta_i} \int_{t_{-i} \in T_{-i}} u_i\left(f^*\left(\theta_i, \widehat{\theta}_{-i}(t_{-i})\right), \left(\widehat{\theta}_i(t_i), \widehat{\theta}_{-i}(t_{-i})\right)\right) d\widehat{\pi}_i(t_i) \\ &= \arg\max_{\theta_i \in \Theta_i} \sum_{\theta_{-i} \in \Theta_{-i}} \left(\int_{\left\{t_{-i} : \widehat{\theta}_{-i}(t_{-i}) = \theta_{-i}\right\}} d\widehat{\pi}_i(t_i)\right) u_i\left(f^*(\theta_i, \theta_{-i}), \left(\widehat{\theta}_i(t_i), \theta_{-i}\right)\right). \end{aligned}$$

But by hypothesis of ex post implementability, truthtelling is a best response for every possible profile θ_{-i}, and thus it remains a best response for arbitrary expectations over Θ_{-i}. ■

The converse does not always hold, as shown by Examples 1 and 2 in the previous section. But we can identify important classes of problems for which the equivalence can be established.

4.1 Separable Environments

A social choice environment is *separable* if the outcome space has a common component and a private value component for each agent. Each agent cares only about the common component and his own private value component. The social choice correspondence picks a unique element from the common component, and has a product structure over all components.

Thus the environment and SCC can be represented in the following way:

$$Y = Y_0 \times Y_1 \times ... \times Y_I;$$

there exists $\widetilde{u}_i : Y_0 \times Y_i \times \Theta \to \mathbb{R}$ such that

$$u_i\left(\left(y_0, y_1, ..., y_I\right), \theta\right) = \widetilde{u}_i\left(y_0, y_i, \theta\right)$$

for all i, $y \in Y$ and $\theta \in \Theta$; and there exist a function $f_0 : \Theta \to Y_0$ and, for each agent i, a non-empty valued correspondence $F_i : \Theta \to 2^{Y_i}/\varnothing$ such that

$$F\left(\theta\right) = f_0\left(\theta\right) \times F_1\left(\theta\right) \times ... \times F_I\left(\theta\right).$$

We observe that the private component for agent i, determined by $F_i\left(\theta\right)$, is allowed to depend on the payoff type profile θ of all agents. The common component is determined by a function, whereas the private components are allowed to be correspondences. The strength of the separability condition, represented by the product structure, is that the set of permissible private components for agent i does not depend on the choice of the private component for the remaining agents.

There are two subsets of separable environments that we are particularly interested in.[13] First, there is the case of the single-valued private component where $Y_i = \{\overline{y}_i\}$ is a single allocation for all i. In this case, there exists a representation of the utility function $\widetilde{u}_i : Y_0 \times \Theta \to \mathbb{R}$ such that $\widetilde{u}_i$ depends only on the common component y_0 and the payoff type profile θ. Thus any social choice *function* is separable. Second, there is the case of the classic quasi-linear environment (described in Section 2). In this case, we set, for each agent i,

$$\begin{aligned} Y_i &= \mathbb{R}, \\ \widetilde{u}_i\left(y_0, y_i, \theta\right) &= v_i\left(y_0, \theta\right) + y_i. \\ \text{and } F_i\left(\theta\right) &= Y_i. \end{aligned}$$

In the quasi-linear environment, the common component $f_0\left(\theta\right)$ will often represent the problem of implementing an efficient allocation, so that

$$f_0\left(\theta\right) = \underset{y_0 \in Y_0}{\arg\max} v_i\left(y_0, \theta\right).$$

As the designer is only interested in maximizing the social surplus and the utilities are quasi-linear, there are no further restriction on the private components, here the monetary

[13] We would like to thank an anonymuous referee for suggesting that we incorporate these two special cases in the unified language of a separable environment.

transfers, offered to the agents. In contrast, in the next section, we shall investigate the quasi-linear environment *with* a balanced budget requirement as a canonical example of a non-separable environment. By requiring a balanced budget, the SCC contains an element of interdependence in the choice of the private components as the transfers have to add up to zero.

Proposition 2 *In separable environments, if F is interim implementable on every common prior payoff type space $\mathcal{T}$, then F is ex post implementable.*

PROOF: Suppose that F can be interim implemented on all type spaces. Then, in particular, it must be possible to interim implement F on the type space where agents other than i have type profile θ_{-i}. Thus for each i and $\theta_{-i} \in \Theta_{-i}$, there must exist $g^{i,\theta_{-i}} : \Theta_i \to Y$ such that i has an incentive to truthfully report his type:

$$u_i\left(g^{i,\theta_{-i}}(\theta_i), (\theta_i, \theta_{-i})\right) \geq u_i\left(g^{i,\theta_{-i}}(\theta_i'), (\theta_i, \theta_{-i})\right) \tag{12}$$

for all $\theta_i, \theta_i' \in \Theta_i$; and such that F is implemented, so that

$$g^{i,\theta_{-i}}(\theta_i) \in F(\theta). \tag{13}$$

If we have a separable environment, condition (13) can be re-written as

$$\begin{aligned} g_0^{i,\theta_{-i}}(\theta_i) &= f_0(\theta_i, \theta_{-i}) \\ \text{and } g_j^{i,\theta_{-i}}(\theta_i) &\in F_j(\theta_i, \theta_{-i}) \text{ for all } j = 1, .., I; \end{aligned}$$

condition (12) can be re-written as

$$u_i\left(f_0(\theta_i, \theta_{-i}), g_i^{i,\theta_{-i}}(\theta_i), (\theta_i, \theta_{-i})\right) \geq u_i\left(f_0(\theta_i', \theta_{-i}), g_i^{i,\theta_{-i}}(\theta_i'), (\theta_i, \theta_{-i})\right) \tag{14}$$

for all $\theta_i, \theta_i' \in \Theta_i$.

But these conditions ensure ex post implementation, by letting

$$f(\theta) = \left(f_0(\theta), g_1^{1,\theta_{-1}}(\theta_1), ..., g_i^{i,\theta_{-i}}(\theta_i), ..., g_I^{I,\theta_{-I}}(\theta_I)\right),$$

which completes the proof. ■

Proposition 2 immediately implies the following strong equivalence result for a separable environment.

Corollary 1 *In separable environments, the following are equivalent:*

1. *F is interim implementable on all type spaces;*

2. *F is interim implementable on all common prior type spaces;*

3. *F is interim implementable on all payoff type spaces;*

4. *F is interim implementable on all common prior payoff type spaces;*

5. *F is ex post implementable.*

PROOF. (1) $\Rightarrow$ (2), (3) and (4) follow by definition as we are asking for interim implementation on a smaller collection of type spaces. By Proposition 2, (4) $\Rightarrow$ (5). By Proposition 1, (5) $\Rightarrow$ (1). ■

Given Proposition 1, whenever we can show that interim implementability on a class of type spaces implies ex post implementability, it follows that there is equivalence between ex post implementation and interim implementation on any collection of type spaces including that class. In the remainder of the paper, we do not report these immediate corollaries.

Our two leading examples of separable environments are (1) when the social choice correspondence is single-valued; and (2) when the environment is quasi-linear. A recent literature has established positive and negative results concerning ex post implementation in quasi-linear environments (see footnote 9), motivating the ex post solution concept as reflecting the planner's ignorance about the true prior. Proposition 2 provides a foundation for the solution concept. In particular, it shows that the impossibility results in Jehiel and Moldovanu (2001) and Jehiel et al. (2004) for ex post implementation with multi-dimensional signals extend to interim implementation.

Proposition 2 and Corollary 1 would be true even without the restriction to separable environments if attention was restricted to truth-telling payoff type direct mechanisms, where outcomes depend only on the reported *payoff* types. This would just be the interdependent value analogue of the classic private values observation that direct implementation for all priors implies dominant strategy implementation (Ledyard (1978) and Dasgupta, Hammond and Maskin (1979)). If the social choice correspondence is single valued, then any implementing mechanism can only depend on payoff types, so the direct mechanism restriction is without loss of generality. But the assumption is not usually without loss of generality - as Examples 1 and 2 showed.

The proof of Proposition 2 used the fact that the class of all common prior payoff type spaces contains as a special case priors where there is only uncertainty about the payoff profile of agent i, but no uncertainty about the payoff profile, $\theta_{-i} \in \Theta_{-i}$, of the remaining agents. Thus a necessary condition of implementation on all type spaces is that, for every i and every $\theta_{-i} \in \Theta_{-i}$, it is possible to solve the agent i single agent implementation problem, when the payoff type profile of the remaining agents is known to be θ_{-i}. The separable condition is then enough to ensure that these necessary conditions spliced together replicate the ex post implementation problem for all agents (this is where the proof would break down in the cases of Examples 1 and 2). But, by construction, the priors used in this proof were not full support common priors. We will see in the next section the extent to which the equivalence result can be strengthened to full support common priors.

4.2 Full Support Conditions

One obvious supplementary condition to the separable environment is to introduce compactness. Thus we say that the environment is compact if each $\widetilde{u}_i(y_0, y_i, \theta)$ is continuous with respect to y_i and each $F_i(\theta)$ is a compact subset of Y_i. We observe that in the quasi-linear environment, the private component is given by $F_i(\theta) = \mathbb{R}$ for all $\theta \in \Theta$ and hence $F(\theta)$ is not compact. For this reason, we will separately prove the equivalence result for the compact environment and the quasi-linear environment.

Proposition 3 *In a compact separable environment, if F is interim implementable on every full support common prior payoff type space $\mathcal{T}$, then F is ex post implementable.*

PROOF. Suppose that F is interim implementable on every common prior full support payoff type space. Then, for every $p \in \Delta_{++}(\Theta)$, there exists for each i, $g_i^p : \Theta \to Y_i$ such that $g_i^p(\theta) \in F_i(\theta)$ for all θ and

$$\begin{aligned} &\sum_{\theta_{-i}} p(\theta_i, \theta_{-i}) \widetilde{u}_i(f_0(\theta_i, \theta_{-i}), g_i^p(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) \\ \geq &\sum_{\theta_{-i}} p(\theta_i, \theta_{-i}) \widetilde{u}_i\left(f_0\left(\theta_i', \theta_{-i}\right), g_i^p\left(\theta_i', \theta_{-i}\right), (\theta_i, \theta_{-i})\right) \end{aligned} \tag{15}$$

for all θ_i and θ_i'. Consider a sequence of priors with $p^n \to p^*$ where $p^*(\theta_{-i}) = 1$. By compactness of each $F_i(\cdot)$, we can choose a convergent subsequence of $g_i^{p^n}$. Writing $g_i^{\theta_{-i}}$ for the limit of that subsequence, we have

$$\begin{aligned} &\widetilde{u}_i\left(f_0(\theta_i, \theta_{-i}), g_i^{\theta_{-i}}(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})\right) \\ \geq\ &\widetilde{u}_i\left(f_0\left(\theta_i', \theta_{-i}\right), g_i^{\theta_{-i}}\left(\theta_i', \theta_{-i}\right), (\theta_i, \theta_{-i})\right), \end{aligned} \tag{16}$$

for all i, θ, θ_i', which ensures ex post incentive compatibility.■

Consider next the quasi-linear environment in which the social choice correspondence is unbounded in the private component. With quasi-linear utilities, it is useful to express the ex post incentive constraints as a set of linear constraints. The only data of the problem that interests us is the incentive of a payoff type θ_i to manipulate the choice of $y_0 \in Y_0$ by mis-reporting his payoff type. His ex post gain to reporting himself to be type θ_i' when he is type θ_i and he is sure that others have type profile θ_{-i} is:

$$\delta_i\left(\theta_i' | \theta_i, \theta_{-i}\right) \triangleq v_i\left(f_0\left(\theta_i', \theta_{-i}\right), \theta\right) - v_i\left(f_0(\theta_i, \theta_{-i}), \theta\right). \tag{17}$$

A set of transfer functions $f = (f_1, ..., f_I)$, each $f_i : \Theta \to \mathbb{R}$, then satisfy ex post incentive compatibility if

$$f_i(\theta_i, \theta_{-i}) - f_i(\theta_i', \theta_{-i}) \geq \delta_i(\theta_i' | \theta_i, \theta_{-i})$$

for all i, θ_i, θ_i' and θ_{-i}.

Proposition 4 *In a quasi-linear environment, if F is interim implementable on every full support common prior payoff type space $\mathcal{T}$, then F is ex post implementable.*

PROOF. We first show that a solution to the following maxmin problem exists for any fixed θ_{-i}:

$$\max_{f_i:\Theta_i \to \mathbb{R}} \left\{ \min_{(\theta_i, \theta_i') \in \Theta_i \times \Theta_i} \left\{ f_i(\theta_i) - f_i(\theta_i') - \delta_i(\theta_i' | \theta_i, \theta_{-i}) \right\} \right\}. \tag{18}$$

To show this, let M be the maximal gain or loss from misreporting of types,

$$M \triangleq \max_{(\theta_i, \theta_i') \in \Theta_i \times \Theta_i} \left| \delta_i(\theta_i' | \theta_i, \theta_{-i}) \right|,$$

let $\overline{F}_i$ be the set of transfer rules bounded by $[-2M, 2M]$,

$$\overline{F}_i = \{ f_i : \Theta_i \to [-2M, 2M] \},$$

and write $\Delta_i(f_i)$ for the lowest incentive to tell the truth under transfer rule f_i,

$$\Delta_i(f_i) \triangleq \min_{(\theta_i, \theta_i') \in \Theta_i \times \Theta_i} \left\{ f_i(\theta_i) - f_i(\theta_i') - \delta_i\left(\theta_i' \middle| \theta_i, \widehat{\theta}_{-i}\right) \right\}.$$

Now observe that for all $f_i \in F_i$, there exists $\overline{f}_i \in \overline{F}_i$ with $\Delta_i(f_i) \leq \Delta_i(\overline{f}_i)$. To see this, let $f_i^0(\theta_i) = 0$ for all θ_i; note that $f_i^0 \in \overline{F}_i$ and $\Delta_i(f_i^0) \geq -M$. If

$$\max_{(\theta_i, \theta_i') \in \Theta_i \times \Theta_i} \left| f_i(\theta_i) - f_i(\theta_i') \right| > 2M$$

then $\Delta_i(f_i) < -M \leq \Delta_i(f_i^0)$. If

$$\max_{(\theta_i, \theta_i') \in \Theta_i \times \Theta_i} \left| f_i(\theta_i) - f_i(\theta_i') \right| \leq 2M,$$

fix any $\overline{\theta}_i$ and let $\widetilde{f}_i(\theta_i) = f_i(\theta_i) - f_i(\overline{\theta}_i)$. Now $\widetilde{f}_i \in \overline{F}_i$ and $\Delta_i(f_i) \leq \Delta_i\left(\widetilde{f}_i\right)$. But now we have that the maximum in expression (18) is attained on a compact subset, so the maxmin exists.

Now suppose that ex post implementation is infeasible. Then there exists j and $\widehat{\theta}_{-j}$ such that, for every $f_j : \Theta_j \to \mathbb{R}$,

$$f_j\left(\theta_j, \widehat{\theta}_{-j}\right) - f_i(\theta_j', \theta_{-j}) < \delta_i\left(\theta_j' \middle| \theta_j, \widehat{\theta}_{-j}\right)$$

for some θ_j, θ_j'. Since we have shown that a solution to

$$\max_{f_j:\Theta_j\to\mathbb{R}}\left\{\min_{\left(\theta_j,\theta_j'\right)\in\Theta_j\times\Theta_j}\left\{f_j\left(\theta_j\right)-f_j\left(\theta_j'\right)-\delta_j\left(\theta_j'\middle|\theta_j,\widehat{\theta}_{-j}\right)\right\}\right\} \tag{19}$$

exists, there exists $\eta>0$ such that, for every $f_j:\Theta_j\to\mathbb{R}$,

$$\min_{\left(\theta_j,\theta_j'\right)\in\Theta_j\times\Theta_j}\left\{f_j\left(\theta_j\right)-f_j\left(\theta_j'\right)-\delta_j\left(\theta_j'\middle|\theta_j,\widehat{\theta}_{-j}\right)\right\}\leq-\eta. \tag{20}$$

Now suppose that F is interim equilibrium implementable on the payoff type space for all priors $p\in\Delta\left(\Theta\right)$. Consequently, for every p there must exist a set of transfers functions, $f_i^p:\Theta\to\mathbb{R}$, and associated interim payments:

$$f_i^p\left(\theta_i\right)\triangleq\sum_{\theta_{-i}\in\Theta_{-i}}f_i^p\left(\theta_i,\theta_{-i}\right)p\left(\theta_{-i}|\theta_i\right),$$

such that $\forall i,\forall\theta_i,\theta_i'$:

$$f_i^p\left(\theta_i\right)-f_i^p\left(\theta_i'\right)\geq\sum_{\theta_{-i}\in\Theta_{-i}}\delta_i\left(\theta_i'\left|\theta_i,\theta_{-i}\right.\right)p\left(\theta_{-i}|\theta_i\right). \tag{21}$$

Let

$$\xi\left(p\right)=\sup_{f_j:\Theta_j\to\mathbb{R}}\left\{\min_{\left(\theta_j,\theta_j'\right)\in\Theta_j\times\Theta_j}\left\{f_j^p\left(\theta_j\right)-f_j^p\left(\theta_j'\right)-\sum_{\theta_{-j}\in\Theta_{-j}}\delta_j\left(\theta_j'\left|\theta_j,\theta_{-j}\right.\right)p\left(\theta_{-j}|\theta_j\right)\right\}\right\}.$$

For all full support p, we have

$$\xi\left(p\right)\leq-\eta+p\left(\widehat{\theta}_{-j}|\theta_j\right)M$$

by (20) and

$$\xi\left(p\right)\geq0$$

by (21). This yields a contradiction if we choose p with $p\left(\widehat{\theta}_{-j}|\theta_j\right)$ sufficiently close to 1. ■

The argument is straightforward, but distinct from the argument in Proposition 3. It proceeds by contrapositive and relies on the linearity in monetary transfers f_i in two crucial steps. First, we can show that the problem of maximizing the minimal ex post benefits from truthtelling over all profiles and all agents is well-defined and admits a finite solution, even though the set of feasible transfers and utilities is unbounded. This allows us to conclude that if ex post implementation is infeasible, then the social choice function which maximizes the minimal benefits of ex post truthtelling (i.e. solves (18)) leads to a *strictly negative* solution. Second, we use the linearity to separate in the incentive constraints the contribution of the utility from the allocation $v_i(y_0,\theta)$ and the monetary transfer $f_i(\theta_i,\theta_{-i})$. The monetary transfer has the further property that the value of the transfer for agent i does neither depend on the allocation y_0 nor on his own true payoff profile. This allows us to evaluate the value of transfers in expectations, thereby eliminating the payoff types of the other agents, exclusively on the basis of the reported type, $\widehat{\theta}_i$, of agent i. But then we are back at the ex post incentive constraints, from which know from the first step, that they have a strict gap, and hence so do interim incentive constraints for distributions close by.

While a similar argument will apply under some weakenings of the quasi-linear assumption, there is not a lot of slack. Suppose each agent's utility takes the form $u_i(y_0,\theta) + v_i(y_i,\theta_i)$, where each v_i is supermodular in (y_i,θ_i), strictly increasing in θ_i and has range $\mathbb{R}_+$. Now each agent's benefit from his transfer is allowed to depend on his own type only. This seems like a minimal weakening of the quasi-linear assumption, yet we have constructed a simple example where interim implementation on all full support payoff type spaces is possible, even though ex post implementation is impossible. We report this example in the appendix, along with an elaborate set of sufficient conditions that do extend the quasi-linear result.

5 The Quasi-Linear Environment with Budget Balance

We now consider the quasi-linear environment with budget balance as a canonical example of a nonseparable environment. There are three reasons for studying this case.

First, we are able to establish some more limited ex post equivalence results in this case. We show that if either there are only two agents or, for an arbitrary number of agents, the payoff space of each agent is binary then the equivalence between ex post implementation and interim implementation on all type spaces holds.

Second, unlike in the case of separable environments of the previous section, we are able to identify an important class of economic environments when there is a gap between interim implementation on all type spaces and interim implementation on all payoff type spaces: in the two agent case, we show that ex post implementation is equivalent to the former but not to the latter. This confirms that our concern with the richness of the type space is not misplaced.

Finally, we know that our results are tight: once there are more than two agents and at least one agent has at least three types, we can show that there is no longer equivalence between ex post implementation and interim implementation on all type spaces. Thus within the budget balanced quasi-linear environments of this section, we are able to establish the limits to ex post equivalence.

Formally, the budget balance requirement is introduced in the quasi-linear environment by imposing budget balance on the private components. Thus we take the definition of a quasi-linear environment in Section 2.1 but let

$$Y = \left\{ (y_0, y_1, ..., y_I) \in Y_0 \times \mathbb{R}^I : \sum_{i=1}^{I} y_i = 0 \right\}.$$

Example 3 was an example of a quasi-linear environment with budget balance.

We exploit a dual characterization of when ex post implementation is possible. The dual approach builds on the classic work of d'Aspremont and Gerard-Varet (1979) and the more recent works of d'Aspremont, Cremer and Gerard-Varet (1995, 2002). In contrast to these works, we use the ex post rather than the interim dual. The dual variables of our characterization will be the multipliers of the budget balance constraints, ν, and the multipliers of the incentive constraints, λ.

Our first result concerns the two agent case. The critical type space in the argument will be the complete information type space. We used a subset of this type space earlier in Example 3 and describe it now more precisely. Let each $T_i = \Theta$ and hence a type of agent i will be written as $t_i = \theta^i \in \Theta$, where $\theta^i = \left(\theta^i_1, ..., \theta^i_I\right)$. We also write θ^i_{-i} for the vector θ^i excluding θ^i_i. We assume that $\widehat{\theta}_i\left(\theta^i\right) = \theta^i_i$ and $\widehat{\pi}_i$ satisfies

$$\widehat{\pi}_i\left(\theta^i\right)[t_{-i}] = \begin{cases} 1, \text{ if } t_j = \theta^i \text{ for all } j \neq i \\ 0, \text{ otherwise} \end{cases}.$$

Thus we require that for each θ, there is a type of agent i who has payoff type θ_i and assigns probability 1 to his opponents each having type θ. The complete information type space is $T = \times_{i=1}^{I} T_i = \left[\times_{i=1}^{I} \Theta_i\right]^I$.

Recall from (17) that we write $\delta_i\left(\theta'_i \,|\theta\right)$ for the ex post incentive of agent i to misreport himself to be type θ'_i when the true type profile is θ. With two agents, the ex-post incentive constraints are given by:

$$\begin{aligned} f_1(\theta) - f_1(r_1, \theta_2) &\geq \delta_1(r_1 \,|\theta), \ \forall r_1, \\ f_2(\theta) - f_2(\theta_1, r_2) &\geq \delta_2(r_2 \,|\theta), \ \forall r_2. \end{aligned} \tag{22}$$

We can use the budget balance condition, $f_1(\theta) + f_2(\theta) = 0$, or $f_1(\theta) = -f_2(\theta)$, to combine the ex post incentive constraints (22) and observe that ex post implementation with budget balance exists if and only if there exists $f_1(\cdot)$ such that:

$$f_1(\theta_1, r_2) - f_1(r_1, \theta_2) \geq \delta_1(r_1 \,|\theta) + \delta_2(r_2 \,|\theta), \forall \theta, \forall r. \tag{23}$$

Proposition 5 (Equivalence with Budget Balance: $I = 2$)
If $I = 2$ and F is interim implementable on all complete information type spaces, then F is ex post implementable.

PROOF: We argue by contrapositive and thus suppose that F is not ex post implementable. Then, by Farkas' Lemma, there exists a nonnegative vector $(\lambda(\theta, r))_{(\theta,r)\in\Theta^2}$ such that for every $\theta \in \Theta$:

$$\sum_r \lambda((\theta_1, r_2), (r_1, \theta_2)) = \sum_r \lambda((r_1, \theta_2), (\theta_1, r_2)), \tag{24}$$

and

$$\sum_{\theta,r} \lambda(\theta, r)[\delta_1(r_1|\theta) + \delta_2(r_2|\theta)] > 0. \tag{25}$$

Let $\nu(\theta)$ denote the common value of the lhs and rhs term in (24).

For $(\theta, r) \in \Theta^2$, we define $q(\theta, r)$ as follows:

$$q(\theta, r) \triangleq \frac{\sum_{r'} \lambda((\theta_1, \theta_2), (r_1, r'_2)) \lambda((r_1, r_2), (r'_1, \theta_2))}{\nu(r_1, \theta_2)}. \tag{26}$$

Therefore, by (24),

$$\sum_{r_2} q((\theta_1, \theta_2), (r_1, r_2)) = \sum_{r_2} \lambda((\theta_1, \theta_2), (r_1, r_2)), \tag{27}$$

and

$$\sum_{r_1} q((r_1, r_2), (\theta_1, \theta_2)) = \sum_{r_1} \lambda((\theta_1, \theta_2), (r_1, r_2)), \tag{28}$$

so that

$$\sum_r q(\theta, r) = \sum_r q(r, \theta). \tag{29}$$

We now show that F is not implementable under the complete information common prior. In contradiction, suppose that $\left(f_1(\theta, \theta'), f_2(\theta, \theta')\right)_{(\theta,\theta')\in\Theta^2}$ is a budget balanced vector of transfers in the complete information setting and interim implements the social choice problem, i.e. for all $\theta \in \Theta$:

$$f_1(\theta, \theta) - f_1(r, \theta) \geq \delta_1(r_1|\theta), \ \forall r \in \Theta, \tag{30}$$

and

$$f_2(\theta, \theta) - f_2(\theta, r) \geq \delta_2(r_2|\theta), \ \forall r \in \Theta. \tag{31}$$

It follows that with positive weights $q(\theta, r)$ and $q(r, \theta)$, as defined in (26), we can sum the inequalities (30) and (31), and obtain:

$$\sum_{\theta,r} q(\theta, r)[f_1(\theta, \theta) - f_1(r, \theta) - \delta_1(r_1|\theta)] +$$
$$\sum_{\theta,r} q(r, \theta)[f_2(\theta, \theta) - f_2(\theta, r) - \delta_2(r_2|\theta)] \geq 0.$$

Using the budget balance requirement, we can write the above inequality as

$$\sum_{\theta,r} q(\theta,r)[f_1(\theta,\theta) - f_1(r,\theta) - \delta_1(r_1|\theta)] + \tag{32}$$
$$\sum_{\theta,r} q(r,\theta)[f_1(\theta,r) - f_1(\theta,\theta) - \delta_2(r_2|\theta)] \geq 0.$$

Regarding the transfers, (29) implies that:

$$\sum_{r} q(\theta,r) f_1(\theta,\theta) = \sum_{r} q(r,\theta) f_1(\theta,\theta), \ \forall\theta,$$

and the remaining transfer terms cancel as well as:

$$-\sum_{\theta,r} q(\theta,r) f_1(r,\theta) + \sum_{\theta,r} q(r,\theta) f_1(\theta,r) = 0,$$

The remaining terms in the inequality (32) can be written as:

$$\sum_{\theta,r_1} \delta_1(r_1|\theta_1,\theta_2) \sum_{r_2} q((\theta_1,\theta_2),(r_1,r_2)) + \sum_{\theta,r_2} \delta_2(r_2|\theta_1,\theta_2) \sum_{r_1} q((r_1,r_2),(\theta_1,\theta_2)). \tag{33}$$

Using (27) and (28), we can rewrite (33) as:

$$\sum_{\theta,r_1} \delta_1(r_1|\theta_1,\theta_2) \sum_{r_2} \lambda((\theta_1,\theta_2),(r_1,r_2)) + \sum_{\theta,r_2} \delta_2(r_2|(\theta_1,\theta_2)) \sum_{r_1} \lambda((\theta_1,\theta_2),(r_1,r_2)). \tag{34}$$

Now (32) implies that expression (34) is less than or equal to zero, contradicting a property of the ex post dual solution (25). ■

Since the equivalence holds for all complete information type spaces, it must also hold for *all type spaces*. Example 3 considered a balanced budget problem with two agents. It already indicated the crucial role of the type space for the interim implementation result. The main feature of the example was that ex post implementation and interim implementation on all type spaces was impossible, yet interim implementation on all payoff type spaces was possible. This illustrates that the equivalence result for $I = 2$ does not hold if all *complete information type spaces* is replaced with all *payoff type spaces*.

For $I = 2$ we directly used the budget balance to combine the ex post incentive constraints for agent 1 and agent 2 at a true payoff type profile θ against reports r_1 and r_2, respectively, into a single constraint for the true state θ and pair of misreports $r = (r_1, r_2)$. The resulting dual variable $\lambda(\theta, r)$ of the ex post constraint has the same dimension as the interim incentive constraints of the complete information type space with true type profile θ and report r. We then directly used the existence of $\lambda(\theta, r)$ to prove that interim implementation on the complete information type space is impossible.[14]

[14] We would like to thank an anonymous referee for suggesting the direct argument presented here.

With more than two agents we have to consider the ex post incentive constraints of each agent separately and then link them through the additional budget balance constraints:

$$f_i\left(\theta_i',\theta_{-i}\right)-f_i\left(\theta_i,\theta_{-i}\right)+\delta_i\left(\theta_i'\left|\theta_i,\theta_{-i}\right.\right)\leq 0,\qquad \forall i,\forall\theta, \tag{35}$$

and the balanced budget constraint

$$\sum_{i=1}^{I}f_i\left(\theta\right)=0,\ \forall\theta. \tag{36}$$

The dual problem to (35) and (36) with the multipliers $\lambda_i:\Theta_i\times\Theta_i\times\Theta_{-i}\to\mathbb{R}_+$ and $\nu:\Theta\to\mathbb{R}$ is given the ex post flow condition (EF):

$$\nu\left(\theta\right)=\sum_{\theta_i'\in\Theta_i}\lambda_i\left(\theta_i',\theta_i,\theta_{-i}\right)-\sum_{\theta_i'\in\Theta_i}\lambda_i\left(\theta_i,\theta_i',\theta_{-i}\right) \tag{37}$$

for all $\theta\in\Theta$ and all i; and the ex post weighting condition (EW):

$$\sum_{i=1}^{I}\sum_{\theta\in\Theta}\sum_{\theta_i'\in\Theta_i}\lambda_i\left(\theta_i',\theta_i,\theta_{-i}\right)\delta_i\left(\theta_i'\left|\theta_i,\theta_{-i}\right.\right)>0. \tag{38}$$

Thus ex post implementation is impossible if and only if there exist (λ,ν) satisfying EF and EW. In the case where each agent has exactly two types, we can use this ex post dual characterization to show the impossibility of interim implementation on all payoff type spaces. In particular, if ex post implementation fails, we can construct a payoff type spaces where interim implementation fails: whenever

$$\sum_{\theta_{-i}}\lambda_i\left(\theta_i,\theta_i',\theta_{-i}\right)>0$$

for some $\theta_i\neq\theta_i'$, let type θ_i assign probability

$$\frac{\lambda_i\left(\theta_i,\theta_i',\theta_{-i}\right)}{\sum\limits_{\theta_{-i}}\lambda_i\left(\theta_i,\theta_i',\theta_{-i}'\right)} \tag{39}$$

to his opponents type profile θ_{-i} (this construction is well defined exactly because there is only one possible $\theta_i'\neq\theta_i$). Now summing interim incentive compatibility constraints will give a contradiction.

We will show the stronger result that ex post implementation is equivalent to interim implementation on all *common prior* payoff type spaces. For this, it is necessary to establish properties of the ex post multipliers; we will show that any solution to EF and EW takes a simple form. Given a dual solution to the ex post program, we refer to $\lambda_i\left(\theta_i',\theta_i,\theta_{-i}\right)>0$ as an *outflow* from (θ_i, θ_{-i}) and correspondingly as an inflow into (θ_i', θ_{-i}). Consistent with this language, we refer to the profile (θ_i, θ_{-i}) as a *source* if there are only outflows:

$$\sum_{\theta_i' \in \Theta_i} \lambda_i \left(\theta_i', \theta_i, \theta_{-i}\right) > 0 \text{ and } \sum_{\theta_i' \in \Theta_i} \lambda_i \left(\theta_i, \theta_i', \theta_{-i}\right) = 0,$$

and refer to (θ_i, θ_{-i}) as a *sink* if there are only inflows:

$$\sum_{\theta_i' \in \Theta_i} \lambda_i \left(\theta_i', \theta_i, \theta_{-i}\right) = 0 \text{ and } \sum_{\theta_i' \in \Theta_i} \lambda_i \left(\theta_i, \theta_i', \theta_{-i}\right) > 0.$$

In the simple solution, every payoff profile θ is either a sink or source, the ex post incentive multipliers, $\lambda_i \left(\theta_i', \theta_i, \theta_{-i}\right)$, are either 0 or 1, and the budget balance multipliers, $\nu(\theta)$, are either -1 or $+1$. In graph-theoretic terms, the multipliers (ν, λ) form the unique solution to the two-coloring problem, which we illustrate below for the case of $I = 3$.

Insert Figure 1 Here

Given this simple structure of the ex post dual, the flow equality ensures that the posteriors can be generated from a common prior. In fact the resulting common prior $p(\cdot)$ puts uniform probability on all sources and zero probability on all the sinks as illustrated in the above figure. The resulting type space is a common prior payoff type space with correlation. Finally, when we add up all the interim incentive constraints under these posteriors, due to the $0, 1$ property of the posteriors and the balanced budget postulate, all the transfers cancel out and we are exactly left with the sum appearing in the ex post weighting inequality. By the hypothesis of the ex post dual the sum is positive and hence the interim incentive constraints cannot be satisfied either.[15]

[15] The dual argument for the hypercube encompasses the cubical array lemma in Walker (1980) which establishes necessary and sufficient conditions for dominant strategy implementation with budget balance in a private value model. Walker considers dominant strategy implementation when the set of possible preferences is given by the class of all utility functions on a given set of allocations. This allows him to assert that the only dominant strategy incentive compatible transfer functions (without regard to budget balance) are the exact Groves schemes. As the Groves schemes represent the marginal contributions of each agent at each type profile, budget balance can be translated into an equality constraint on the sum of the differences of the social valuations at the true profiles on the hypercube. Generically, the social values will not satisfy the equality. In the current model, we are only considering a finite set of preferences for each agent, and hence the set of dominant strategy incentive compatible transfers (without regard to budget balance) is larger then the set of exact Groves schemes. For the hypercube, this implies that the sum of differences in Walker is always strictly larger than our weighting inequality, and thus if budget balance fails on the hypercube and the weighting inequality is positive, then the Groves schemes will necessarily fail as well.

Proposition 6 (Equivalence with Budget Balance: $I > 2$**)**
If $\#\Theta_i \leq 2$ *for all* i*, and* F *is interim implementable on all payoff type spaces, then* F *is ex post implementable.*

PROOF: We first note that if any agent has only one type, then a well known argument establishes that budget balance has no bite, since the single type can absorb the budget surpluses or deficit (see Mas-Collel et al. (1995), p.881). Thus suppose that $\#\Theta_i = 2$ for all i. The proof is by contradiction. Thus suppose F is not ex post implementable and hence there does not exist a solution to the ex post incentive constraints and budget balance constraints, (35) and (36). By Farkas' Lemma with equality constraints, it then follows that there must exist a solution to the dual problem (37) and (38) satisfying $\lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right) \geq 0$, $\forall i, \forall \theta$.

Next we show that if a solution $\left\{v\left(\theta\right), \lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right)\right\}$ exists, then there also exists a solution such that for all i, θ, and θ_i' :

$$v\left(\theta\right) \in \left\{-1, 1\right\}, \ \lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right) \in \left\{0, 1\right\}, \ \ \lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right) + \lambda_i\left(\theta_i, \theta_i', \theta_{-i}\right) = 1. \tag{40}$$

The binary payoff type space implies that for a given θ_i, payoff type $\theta_i' \neq \theta_i$, is uniquely determined. We first observe that a necessary condition for interim (and ex post) *incentive compatibility* on all payoff type spaces is that for all i and all θ_{-i} :

$$f_i\left(\theta_i', \theta_{-i}\right) - f_i\left(\theta_i, \theta_{-i}\right) + \delta_i\left(\theta_i' \left|\theta_i, \theta_{-i}\right.\right) \leq 0,$$

and

$$f_i\left(\theta_i, \theta_{-i}\right) - f_i\left(\theta_i', \theta_{-i}\right) + \delta_i\left(\theta_i \left|\theta_i', \theta_{-i}\right.\right) \leq 0.$$

By summing up the two inequalities we obtain that for all i and all θ_{-i}:

$$\delta_i\left(\theta_i' \left|\theta_i, \theta_{-i}\right.\right) + \delta_i\left(\theta_i \left|\theta_i', \theta_{-i}\right.\right) \leq 0. \tag{41}$$

Based on the given solution $\left\{v\left(\theta\right), \lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right)\right\}$, we then propose a new solution $\left\{v\left(\theta\right), \widehat{\lambda}_i\left(\theta_i', \theta_i, \theta_{-i}\right)\right\}$ defined by:

$$\widehat{\lambda}_i\left(\theta_i', \theta_i, \theta_{-i}\right) \triangleq \max\left\{\lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right) - \lambda_i\left(\theta_i, \theta_i', \theta_{-i}\right), 0\right\},$$

and correspondingly

$$\widehat{\lambda}_i\left(\theta_i, \theta_i', \theta_{-i}\right) \triangleq \max\left\{\lambda_i\left(\theta_i, \theta_i', \theta_{-i}\right) - \lambda_i\left(\theta_i', \theta_i, \theta_{-i}\right), 0\right\}.$$

By construction the new solution satisfies the equality constraints (37) under the original values $\nu\left(\theta\right)$ and by (41) weakly increases the right side of the inequality constraint (38). Accordingly, the equalities (37) simplify to either

$$v\left(\theta\right) = \widehat{\lambda}_i\left(\theta_i', \theta_i, \theta_{-i}\right), \forall i, \tag{42}$$

or

$$v(\theta) = -\widehat{\lambda}_i\left(\theta_i, \theta_i', \theta_{-i}\right), \ \forall i. \tag{43}$$

Due to the binary property of the type space Θ_i and the fact that the above equalities, (42) and (43) have to hold for all agents simultaneously, we obtain a bipartition of the type space Θ into subsets, Θ' and Θ'' (in graph-theoretic terms they form the unique solution to the two-coloring problem) such that for all $\theta \in \Theta'$, $v(\theta) > 0$ and for all $\theta \in \Theta''$, $v(\theta) < 0$. We can finally normalize $v(\theta)$ and $\widehat{\lambda}_i\left(\theta_i', \theta_i, \theta_{-i}\right)$ by dividing through $|v(\theta)|$ to obtain a solution, denoted by $\left\{v^*(\theta), \lambda_i^*\left(\theta_i', \theta_i, \theta_{-i}\right)\right\}$, with the desired properties described in (40). The inequality (38) now reads:

$$\sum_{i=1}^{I} \sum_{\theta \in \Theta'} \delta_i\left(\theta_i' \left| \theta_i, \theta_{-i}\right.\right) > 0. \tag{44}$$

We obtain a contradiction to (44) by considering the interim implementation for the payoff prior which puts uniform probability on all $\theta \in \Theta'$ and zero probability on all $\theta \in \Theta''$. By hypothesis of interim implementability, the interim incentive constraints for every i and every θ_i:

$$\sum_{(\theta_i, \theta_{-i}) \in \Theta'} \left[f_i\left(\theta_i', \theta_{-i}\right) - f_i\left(\theta_i, \theta_{-i}\right) + \delta_i\left(\theta_i' \left| \theta_i, \theta_{-i}\right.\right)\right] p\left(\theta_{-i} \left| \theta_i\right.\right) \leq 0,$$

can be satisfied with a balanced budget transfer scheme. By summing up the interim incentive constraints over all agents and omitting the constant (on Θ') probability, $p\left(\theta_{-i} \left| \theta_i\right.\right)$, we get:

$$\sum_{i=1}^{I} \sum_{(\theta_i, \theta_{-i}) \in \Theta'} \left[f_i\left(\theta_i', \theta_{-i}\right) - f_i\left(\theta_i, \theta_{-i}\right) + \delta_i\left(\theta_i' \left| \theta_i, \theta_{-i}\right.\right)\right] \leq 0,$$

and by the balanced budget stipulation, the transfers drop out, and we are left with:

$$\sum_{i=1}^{I} \sum_{(\theta_i, \theta_{-i}) \in \Theta'} \delta_i\left(\theta_i' \left| \theta_i, \theta_{-i}\right.\right) \leq 0,$$

which provides the desired contradiction to (44). ■

In an earlier version of this paper, Bergemann and Morris (2003), we demonstrated, by means of an example, the tightness of the ex post equivalence results obtained in Proposition 5 and 6. Example 4 consisted of three agents in which the first agent had three payoff types, and the remaining two agents had binary payoff type spaces. With this minimal relaxation of either of the above sufficient conditions, we have an example where ex post implementation is impossible while interim implementation (using a single mechanism) is possible on all type spaces. We conjecture that ex post equivalence results may again be obtained in a general environment with $I > 2$ and $\#\Theta_i > 2$ only after imposing suitable restrictions on the environment such as single crossing or supermodularity conditions.

6 Discussion and Conclusion

6.1 The Universal Type Space

We wanted to address the following question. Suppose that the planner knows that the payoff environment is common knowledge, but the planner knows nothing else about agents' types; in particular, he knows nothing about their beliefs and higher order beliefs about other types. Can the planner implement the social choice correspondence? We formalized this question by requiring the planner to implement the social choice correspondence on any type space (consistent with the payoff environment).

An alternative approach to formalizing this question is to work with a universal type space about the payoff environment, maintaining the assumption that agents know their own payoff types. Thus there is a (known own payoff type) universal type space

$$\mathcal{T}^* = \left(T_i^*, \widehat{\theta}_i^*, \widehat{\pi}_i^*\right)_{i=1}^{I},$$

with the properties that the range of $\left(\widehat{\theta}_i^*, \widehat{\pi}_i^*\right)$ is $\Theta_i \times \Delta\left(T_{-i}^*\right)$. The existence of such a universal type space can be shown by a slight adaption of the standard arguments of Mertens and Zamir (1985) and Brandenburger and Dekel (1993); in fact, the universal type space can be explicitly constructed by identifying T_i^* with all coherent sequences of agents' beliefs and higher order beliefs about payoff types.[16]

All the positive and negative results reported in this paper would be unchanged if we replaced "implementable on all type spaces" by "implementable on the (known own payoff type) universal type space". Since the universal type space is an example of a type space, implementability for all type spaces trivially implies implementability on the universal type space. On the other hand, when we show a failure of implementability for all type spaces, we do so by constructing a finite type space where implementability is impossible. But those finite type spaces are isomorphic to belief closed subsets of the universal type space. And if it is not possible to implement on a given type space, it is not possible to implement on any type space (such as the universal type space) which contains that given type space as a belief closed subset. Thus whenever implementability is impossible on those finite type spaces, it is also impossible on the universal type space.

We emphasize that this equivalence between "all type spaces" and the "universal type space" holds for the questions we ask in this paper and holds under our maintained assumption that it is common knowledge that agents know their own payoff types. For other questions and other environments, this equivalence may not hold. If we fix a set of payoff types, there is sometimes a difference between analyzing strategic behavior on the universal

[16] See Heifetz and Neeman (2004) for a more detailed discussion of such constructions.

type space constructed for those payoff types and analyzing strategic behavior on all type spaces.[17] The universal type space constructions rely on topological assumptions that might be problematic for other applications.[18]

6.2 A Classical Debate

An old debate in the Bayesian implementation literature went as follows. Some scholars pointed out that - as a practical matter - the planner was unlikely to know the true prior over the type space. Therefore, it would be desirable to have a mechanism which was going to work independent of the prior. For a private values environment, Dasgupta, Hammond and Maskin (1979), Ledyard (1978, 1979) and Groves and Ledyard (1987) observed that if a direct mechanism was going to implement a social choice correspondence for every prior on the type space, then there must be dominant strategies implementation. Other scholars pointed out that if the planner did not know the prior (and the agents do) then we should not restrict attention to direct mechanisms. Rather, we should allow the mechanism to elicit reports of the true prior from the agents (since this information is non-exclusive in the sense of Postlewaite and Schmeidler (1986), this elicitation will not lead to any incentive problems). A formal application of this folk argument appears in the recent work of Choi and Kim (1999). How do our results fit into this debate between the "practical designers" and the "implementation purists"?

Our results allow for interdependent values, but we believe they clarify this debate when restricted to private values (recall that ex post incentive compatibility implies dominant strategies incentive compatibility under private values).

- In some environments, even if the designer was allowed to elicit the true prior, implementation for every prior on the fixed type space implies dominant strategy implementation. In these environments, the practical designers' conclusion is immune to the purists' criticism. These environments include separable environments (Proposition 2) and quasi-linear environments with budget balance but at most two types for each player (Proposition 6).[19]

[17] For example, if each player has only one possible payoff type, then he has only one type in the constructed universal type space, and equilibrium on the type space implies Nash equilibrium play in the implied complete information game. But if we look at all possible type spaces consistent with those payoff types, any rationalizable action might be played.

[18] Heifetz and Samet (1998) provide a non-constructive proof of the existence of a universal type space without topological restictions.

[19] Mookerjee and Reichelstein (1992) examine the relation between Bayesian implementation and dominant strategy implementation in the private value environment. If dominant strategy implementation of an allocation rule is possible, then it is possible to do so in a way that generates any expected transfer levels achievable under Bayesian implementation of that allocation rule.

- In some environments, the purists' criticism binds. That is, dominant strategy implementation is impossible but Bayesian implementation is possible for every prior on a fixed type space. This was true of Examples 1 and 2 and we can also construct quasi-linear environments with budget balance where it is true.

- A second practical criticism of the classical Bayesian implementation literature is that not only may the planner not know the true prior over the payoff types, but the agents may not know the true prior either. We have formalized this criticism by requiring implementation on type spaces larger than the payoff type space. And we have shown that in some environments, implementation on all type spaces implies dominant strategies implementation even when interim implementation for all priors on the payoff type space does not (Example 3 and Proposition 5).[20]

6.3 Genericity

If we restrict attention to "generic" priors on the payoff type space (or any fixed finite type space), it is possible to obtain very permissive implementation results. Thus arguments in d'Aspremont, Cremer and Gerard-Varet (1979) and d'Aspremont, Cremer and Gerard-Varet (1995, 2002) establish that it is possible to implement any allocation rule in a quasi-linear environment with budget balance for a generic set of priors on all fixed type spaces. This contrasts with our results showing that implementation in some quasi-linear environments with budget balance for all priors on (the fixed) payoff type space is equivalent to ex post implementation, which is known to be impossible under quite general conditions.

As emphasized by Neeman (2001), "generic" priors entail some counterintuitive properties, e.g., that a planner can infer an agent's valuation of an object from that agent's beliefs about other agents' types. In any case, the justification for fixing a set of types, "generically" picking a prior and *then* assuming common knowledge of that prior is not clear. Some current work tries to identify more natural ways of thinking about genericity.[21] In this work, we have not discussed any results that rely on genericity notions.[22]

[20] Example 3 has interdependent values, but we could mechanically turn it into a private value example the same way we constructed Example 2 as a private value version of Example 1.

[21] Morris (2002) and Dekel, Fudenberg and Morris (2004) examine ways of defining "strategic topologies" on types in the universal type space that might suggest useful topological notions of genericity. Heifetz and Neeman (2004) argue that among common prior belief closed subspaces of the universal type space, the type spaces often described as generic are not "prevalent" in the sense of Christensen (1974), Hunt et al. (1992) and Anderson and Zame (2001).

[22] Genericity issues are discussed at greater length in the working paper version of this paper, Bergemann and Morris (2003).

6.4 Augmented Ex Post Equivalance

In non-separable environments, ex post implementability may be a strictly stronger requirement than interim implementability on all type spaces. Is there a natural weakening of ex post implementability that is necessary? Consider an augmented mechanism where each agent's report consists of his payoff type and a supplemental message. An agent's strategy is truthful if he always correctly reports his payoff type. A decision rule mapping message profiles into outcomes is augmented ex post incentive compatible if an agent who expects all other agents to report truthfully has a truthful best response. A social choice correspondence F is augmented ex post implementable if there exists an augmented ex post incentive compatible decision rule that (under truthful strategies) always achieves outcomes in F. We showed in the working paper version of this paper (Bergemann and Morris (2003)) that (up to some technical restrictions) augmented ex post incentive compatibility is equivalent to interim implementability on all type spaces. Now an interesting way to characterize implementation problems is how many supplemental messages are needed. For separable environments, no extra messages are needed. In the worst case, the supplemental message might consist of the agent's belief over T_{-i}^* in the universal type space and we would be looking at a direct mechanism on the universal type space. An interesting problem for future research is the characterization of how many supplemental messages are required for different classes of problems.

References

[1] Anderson, R. and W. Zame. 2001. "Genericity with Infinitely Many Parameters." *Advances in Theoretical Economics* 1.

[2] d'Aspremont, C. and L.-A. Gerard-Varet. 1979. "Incentives and Incomplete Information." *Journal of Public Economics* 11: 25-45.

[3] d'Aspremont, C., J. Cremer and L.-A. Gerard-Varet. 1995. "Correlation, Independence and Bayesian Incentives."

[4] d'Aspremont, C., J. Cremer and L.-A. Gerard-Varet. 2002. "Balanced Bayesian Mechanisms."

[5] Battigalli, P. and M. Siniscalchi. 2003. "Rationalization and Incomplete Information." *Advances in Theoretical Economics* 3.

[6] Bergemann, D. and S. Morris. 2003. "Robust Mechanism Design." Cowles Foundation Discussion Paper 1421, available at http://ssrn.com/abstract=412497.

[7] Bergemann, D. and S. Morris. 2004. “Robust Implementation: The Role of Rich Type Spaces.” http://www.econ.yale.edu/sm326/rmd-full.pdf.

[8] Bergemann, D., S. Morris and Ilya Segal. 2004. “Robust Optimal Auctions.” in progress.

[9] Bergemann, D. and J. Valimaki. 2002. “Information Acquisition and Mechanism Design.” *Econometrica* 70: 1007-1033.

[10] Choi, J. and T. Kim. 1999. “A Nonparametric, Efficient Public Good Decision Mechanism: Undominated Bayesian Implementation.” *Games and Economic Behavior* 27: 64-85.

[11] Christiansen, J. 1974. *Topology and Borel Structure*, North Holland Mathematical Studies Vol. 10, Amsterdam: North-Holland.

[12] Chung, K. and J. Ely. 2003. "Implementation with Near-Complete Information." *Econometrica* 71: 857-871.

[13] Cremer, J. and R. McLean. 1985. “Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist when Demands are Interdependent.” *Econometrica* 53: 345-361.

[14] Dasgupta, P., P. Hammond and E. Maskin. 1979. “The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility.” *Review of Economic Studies* 46: 185-216.

[15] Dasgupta, P. and E. Maskin. 2000. “Efficient Auctions.” *Quarterly Journal of Economics* 115: 341-388.

[16] Dekel, E., D. Fudenberg and S. Morris. 2004. “Dense Types." mimeo.

[17] Duggan, J. and J. Roberts. 1997. “Robust Implementation.” University of Rochester.

[18] Eliaz, K. 2002. “Fault Tolerant Implementation.” *Review of Economic Studies* 69, 589-610.

[19] Groves, T. and J. Ledyard. 1987. “Incentive Compatibility Since 1972,” in *Information, Incentives and Economic Mechanisms* (T. Groves, R. Radner, and S. Reiter, Eds.). University of Minnesota Press: Minneapolis.

[20] Hagerty, K. and W. Rogerson. 1987. “Robust Trading Mechanisms.” *Journal of Economic Theory* 42: 94-107 .

[21] Harsanyi, J. 1967/68. "Games with Incomplete Information Played by Bayesian agents." *Management Science* 14, 159-182, 320-334, 486-502.

[22] Heifetz, A. and D. Samet. 1988. "Topology-Free Typology of Beliefs." *Journal of Economic Theory* 82: 324-341.

[23] Heifetz, A. and Z. Neeman. 2004. "On the Generic Infeasibility of Full Surplus Extraction in Mechanism Design." mimeo

[24] Holmstrom, B. and R. Myerson. 1983. "Efficient and Durable Decision Rules with Incomplete Information." *Econometrica* 51: 1799-1820.

[25] Hunt, B., T. Sauer and J. Yorke. 1992. "Prevalence: A Translation-Invariant 'Almost Every' on Infinite-Dimensional Spaces," *Bulletin (New Series) of the American Mathematical Society* 27, 217-238.

[26] Hurwicz, L. 1972. "On Informationally Decentralized Systems," in *Decision and Organization* (R. Radner and C. McGuire, Eds.). North Holland: Amsterdam.

[27] Jehiel, P. and B. Moldovanu. 2001. "Efficient Design with Interdependent Valuations." *Econometrica* 65: 1237-1259.

[28] Jehiel, P., M. Meyer-ter-Vehn and B. Moldovanu. 2004. "The Limits of Ex-Post Implementation."

[29] Kalai, E. 2002. "Large Robust Games." Northwestern University.

[30] Ledyard, J. 1978. "Incentive Compatibility and Incomplete Information." *Journal of Economic Theory* 18: 171-189.

[31] Ledyard, J. 1979. "Dominant Strategy Mechanisms and Incomplete Information." In *Aggregation and Revelation of Preferences*, ed. Jean-Jacques Laffont. Amsterdam: North Holland Chapter 17, pp. 309-319.

[32] Lopomo, G. 1998. "The English Auction is Optimal among Simple Sequential Auctions." *Journal of Economic Theory* 82: 144-166.

[33] Lopomo, G. 2000. "Optimality and Robustness of the English Auction." *Games and Economic Behavior* 36: 219-240.

[34] Mas-Colell, A., M. Whinston and J. Green. 1995. *Microeconomic Theory.* Oxford: Oxford University Press.

[35] Maskin, E. 1992. "Auctions and Privatization." In: *Privatization: Symposium in Honor of Herbert Giersch*, ed. H. Siebert. Tuebingen: J.C.B.Mohr pp. 115-136.

[36] Mertens, J.-F. and S. Zamir. 1985. "Formulation of Bayesian Analysis for Games of Incomplete Information." *International Journal of Game Theory* 14: 1-29.

[37] Mookherjee, D. and S. Reichelstein. 1992. "Dominant Strategy Implementation of Bayesian Incentive Compatible Allocation Rules." *Journal of Economic Theory* 56: 378-399.

[38] Morris, S. 2002. "Typical Types." http://www.econ.yale.edu/~sm326/typical.pdf.

[39] Morris, S. and H. S. Shin. 2003. "Global Games: Theory and Applications" in *Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society)* (M. Dewatripont, L. Hansen and S. Turnovsky, Eds). Cambridge, England: Cambridge University Press.

[40] Myerson, R. 1991. *Game Theory.* Cambridge: Harvard University Press.

[41] Neeman, Z. 2001. "The Relevance of Private Information in Mechanism Design." *Journal of Economic Theory,* forthcoming.

[42] Perry, M. and P. Reny. 2002. "An Ex Post Efficient Auction." *Econometrica* 70: 1199-1212.

[43] Walker, M. 1980. "On the Nonexistence of a Dominant Strategy Mechanism for Making Optimal Public Decisions." *Econometrica* 48, 1521-1540.

[44] Wilson, R. 1985. "Incentive Efficiency in Double Auctions." *Econometrica* 53: 1101-1115.

[45] Wilson, R. 1987. "Game-Theoretic Analyses of Trading Processes." In Advances in Economic Theory: Fifth World Congress, ed. Truman Bewley. Cambridge: Cambridge University Press chapter 2, pp. 33-70.

7 Appendix

In this appendix we provide a set of sufficient conditions which encompass the quasilinear environment. With these conditions, we can generalize the equivalence results presented in Proposition 4 for common prior full support payoff type spaces. The first condition replaces the compactness condition, the second and third condition generalizes the important features of the quasilinear utility model. The appendix ends with an example which is meant to illustrate that the conditions, in particular, the condition on bounded allocation differences, is not easily dispensed with.

We substitute the compactness condition by essential compactness. To this end let

$$\Delta_i(\theta_{-i}, f_i) \triangleq \min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i} \begin{bmatrix} \widetilde{u}_i(f_0(\theta_i,\theta_{-i}), f_i(\theta_i,\theta_{-i}), (\theta_i,\theta_{-i})) \\ -\widetilde{u}_i(f_0(\theta_i',\theta_{-i}), f_i(\theta_i',\theta_{-i}), (\theta_i,\theta_{-i})) \end{bmatrix}.$$

Definition 7 *(Essential compactness).*
Each $\widetilde{u}_i(y_0, y_i, \theta)$ is continuous with respect to y_i. For each θ and i, there is a compact set $\overline{F}_i(\theta) \subseteq F_i(\theta)$, such that for each i and θ_{-i}, there exists $f_i^* : \Theta \to Y_i$ with $f_i^*(\theta) \in \overline{F}_i(\theta)$ such that $\Delta_i(\theta_{-i}, f_i^*) \geq \Delta_i(\theta_{-i}, f_i)$ for all $f_i : \Theta \to Y_i$ with each $f_i(\theta) \in F_i(\theta)$.

Definition 8 *(Transferable utility).*
For every $\psi \in \Delta(\Theta_{-i})$ and $f_i : \Theta \to Y_i$ with $f_i(\theta) \in F_i(\theta)$ for all θ, there exists $\overline{f}_i : \Theta_i \to Y_i$ with $f_i(\theta) \in F_i(\theta)$ for all θ such that

$$\begin{aligned} &\sum_{\theta_{-i}\in\Theta_{-i}} \psi_i(\theta_{-i})\, \widetilde{u}_i\left(f_0(\theta_i',\theta_{-i}), \overline{f}_i(\theta_i'), (\theta_i,\theta_{-i})\right) \\ = &\sum_{\theta_{-i}\in\Theta_{-i}} \psi_i(\theta_{-i})\, \widetilde{u}_i\left(f_0(\theta_i',\theta_{-i}), f_i(\theta_i',\theta_{-i}), (\theta_i,\theta_{-i})\right) \end{aligned}$$

for all θ_i and θ_i'.

Definition 9 *(Bounded Allocation Differences).*
There exists M such that

$$\left|\widetilde{u}_i(y_0, y_i, \theta) - \widetilde{u}_i(y_0', y_i, \theta')\right| \leq M \text{ for all } i,\ y_0, y_0' \in Y_0,\ y_i \in Y_i \text{ and } \theta,\ \theta' \in \Theta.$$

Essential compactness ensures that the problem of maximizing the minimal gains from thruthtelling in the ex post incentive constraints always has a well defined solution. Transferable utility ensures that every utility compensation for agent i can be achieved by assigning the private component $\overline{f}_i$ conditionally on the reported payoff type θ_i' of agent i only. A sufficient condition for transferable utility is that $\widetilde{u}_i(y_0, y_i, \theta)$ is continuous with respect to y_i and that the utility is positively and negatively unbounded in y_i for every y_0 and θ. Bounded allocation differences ensures that the differences in utilities due to the public component and the payoff type for every given private component is bounded.

Lemma 1 *If F is essentially compact, separable with transferable utility and satisfies bounded allocation differences and F is interim implementable on every full support common prior payoff type space $\mathcal{T}$, then F is ex post incentive implementable.*

PROOF. Suppose F is not ex post compatible. By essential compactness, for each i and θ_{-i}, there exists $\delta > 0$, such that for all $f_i : \Theta \to Y_i$,

$$\min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i}\left[\begin{array}{c} \widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}\right),f_i\left(\theta_i,\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) \\ -\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),f_i\left(\theta_i',\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) \end{array}\right] \leq -\delta. \tag{45}$$

Now suppose that F is interim implementable on every full support common prior payoff type space. Then, for every $p \in \Delta_{++}(\Theta)$, there exists for each i, $f_i^p : \Theta \to Y_i$ such that $f_i^p(\theta) \in F_i(\theta)$ for all θ and

$$\min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i}\left[\begin{array}{c} \sum\limits_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}\right),f_i^p\left(\theta_i,\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) \\ -\sum\limits_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),f_i^p\left(\theta_i',\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) \end{array}\right] \geq 0,$$

for all θ_i and θ_i'. By transferable utility, there exists $\overline{f}_i^p : \Theta_i \to Y_i$ such that $\overline{f}_i^p(\theta) \in F_i(\theta)$ for all θ and

$$\min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i}\left[\begin{array}{c} \sum\limits_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}\right),\overline{f}_i^p\left(\theta_i\right),\left(\theta_i,\theta_{-i}\right)\right) \\ -\sum\limits_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}\right)\right) \end{array}\right] \geq 0.$$

By bounded allocation differences, for any $\theta_{-i}^* \in \Theta_{-i}$,

$$\left|\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}\right)\right)-\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}^*\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}^*\right)\right)\right| \leq M.$$

We can express the expected utility under $p(\cdot)$ from f_0 and $\overline{f}_i^p$ as

$$\begin{aligned} &\sum_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}\right)\right) = \\ &\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}^*\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}^*\right)\right) + \\ &\sum_{\theta_{-i}\neq\theta_{-i}^*} p\left(\theta_i,\theta_{-i}\right)\left[\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}\right)\right)-\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}^*\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}^*\right)\right)\right] \end{aligned}$$

So

$$\left|\begin{array}{c} \sum\limits_{\theta_{-i}\in\Theta_{-i}} p\left(\theta_i,\theta_{-i}\right)\left(\widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}\right),\overline{f}_i^p\left(\theta_i\right),\left(\theta_i,\theta_{-i}\right)\right)-\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}\right)\right)\right) \\ -\left(\widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}^*\right),\overline{f}_i^p\left(\theta_i\right),\left(\theta_i,\theta_{-i}^*\right)\right)-\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}^*\right),\overline{f}_i^p\left(\theta_i'\right),\left(\theta_i,\theta_{-i}^*\right)\right)\right) \end{array}\right|$$

$$\leq \left(1 - p\left(\theta_i, \theta^*_{-i}\right)\right) 2M$$

But consider a sequence of priors with $p^n \to p^*$ and $p^*\left(\theta^*_{-i}\right) = 1$. Since

$$\min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i} \left[\begin{array}{c} \sum\limits_{\theta_{-i}\in\Theta_{-i}} p^n\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i,\theta_{-i}\right), \overline{f}_i^{p^n}\left(\theta_i\right), \left(\theta_i,\theta_{-i}\right)\right) \\ -\sum\limits_{\theta_{-i}\in\Theta_{-i}} p^n\left(\theta_i,\theta_{-i}\right)\widetilde{u}_i\left(f_0\left(\theta_i',\theta_{-i}\right), \overline{f}_i^{p^n}\left(\theta_i'\right), \left(\theta_i,\theta_{-i}\right)\right) \end{array} \right] \geq 0$$

for all n we have that:

$$\min_{(\theta_i,\theta_i')\in\Theta_i\times\Theta_i} \left[\begin{array}{c} \widetilde{u}_i\left(f_0\left(\theta_i,\theta^*_{-i}\right), \overline{f}_i^{p^n}\left(\theta_i\right), \left(\theta_i,\theta^*_{-i}\right)\right) \\ -\widetilde{u}_i\left(f_0\left(\theta_i',\theta^*_{-i}\right), \overline{f}_i^{p^n}\left(\theta_i'\right), \left(\theta_i,\theta^*_{-i}\right)\right) \end{array} \right]$$

tends to 0 as $n \to \infty$. But this contradicts (45).■

It is immediate that

Lemma 2 *Essential compactness, transferable utility and bounded allocation differences are satisfied in the quasilinear environment.*

PROOF. (1) Essential Compactness. Let

$$M = \max_{y_0,y_0'} \left|v_i\left(y_0,\theta\right) - v_i\left(y_0',\theta\right)\right|$$

and let $F_i = \{f_i : \Theta \to \mathbb{R}\}$ and $\overline{F}_i = \{f_i : \Theta \to [-2M, 2M]\}$. To show essential compactness, it is enough to show that for all $f_i \in F_i$ and θ_{-i}, there exists $\overline{f}_i \in \overline{F}_i$ with $\Delta_i\left(\theta_{-i}, f_i\right) \leq \Delta_i\left(\theta_{-i}, \overline{f}_i\right)$. To see this, let $f_i^0\left(\theta_i\right) = 0$ for all θ_i; note that $f_i^0 \in \overline{F}_i$ and $\Delta_i\left(\theta_{-i}, f_i^0\right) \geq -M$. If

$$\max_{\theta_i,\theta_i'} \left|f_i\left(\theta_i\right) - f_i\left(\theta_i'\right)\right| > 2M$$

then $\Delta_i\left(\theta_{-i}, f_i\right) < -M \leq \Delta_i\left(\theta_{-i}, f_i^0\right)$. If

$$\max_{\theta_i,\theta_i'} \left|f_i\left(\theta_i\right) - f_i\left(\theta_i'\right)\right| \leq 2M,$$

fix any $\overline{\theta}_i$ and let $\widetilde{f}_i\left(\theta_i\right) = f_i\left(\theta_i\right) - f_i\left(\overline{\theta}_i\right)$. Clearly, $\widetilde{f}_i \in \overline{F}_i$ and $\Delta_i\left(\theta_{-i}, f_i\right) \leq \Delta_i\left(\theta_{-i}, \widetilde{f}_i\right)$.

(2) Transferable utility. This is immediate: just set

$$\overline{f}_i\left(\theta_i\right) \triangleq \sum_{\theta_{-i}\in\Theta_{-i}} \psi_i\left(\theta_{-i}\right) f_i\left(\theta_i, \theta_{-i}\right).$$

(3) Bounded Allocation Differences. Also immediate: set

$$M = \max_{i,\theta,\theta',y_0,y_0'} \left|v_i\left(y_0,\theta\right) - v_i\left(y_0',\theta'\right)\right|,$$

which completes the proof.■

The following example satisfies essential compactness and transferable utility, but does not satisfy bounded allocation differences. Yet, it is arguably a very slight departure from the quasilinear model.

EXAMPLE 4. Consider the following example with two agents, $i = 1, 2$. The payoff type space of agent 1 is $\Theta_1 = \{1, 2\}$ and of agent 2 it is $\Theta_2 = \{1, 2, 3\}$. We consider an additive utility function

$$u_i(y_0, \theta) + v_i(y_i, \theta_i)$$

as a minimal extension of the quasilinear utility function. The allocation rule for the common component is:

$$\begin{array}{cccc} f_0 & \theta_2 = 1 & \theta_2 = 2 & \theta_2 = 3 \\ \theta_1 = 1 & a & c & d \\ \theta_1 = 2 & b & c & d \end{array} \tag{46}$$

and the utility from the common component $u_1(y_0, \theta)$ for agent 1 is given by:

$$\begin{array}{cccc} u_1(a,\cdot) & 1 & 2 & 3 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 \end{array} \quad \begin{array}{cccc} u_1(b,\cdot) & 1 & 2 & 3 \\ 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{array} \quad \begin{array}{cccc} u_1(c,\cdot) & 1 & 2 & 3 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 \end{array} \quad \begin{array}{cccc} u_1(d,\cdot) & 1 & 2 & 3 \\ 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \end{array} \tag{47}$$

The social choice correspondence for the private component is $F_i = \mathbb{R}_+$ and the utility from the private component is given by

$$v_i(y_i, \theta_i) = \sum_{k=1}^{\theta_i} y_i^k, \tag{48}$$

or

$$v_i(y_i, 1) = y_i, \quad v_i(y_i, 2) = y_i + y_i^2.$$

It is easy to verify that the utility function is supermodular in (y_i, θ_i) and thus well behaved with respect to implementation constraints. We exclusively focus attention on the incentive problem of agent 1 at $\theta_2 = 1$ and use the additional payoff type of agent 2, $\theta_2 = 2, 3$, as conditioning devices in the interim implementation.

By (46) and (47), the SCC F is not ex post incentive compatible for agent 1 at $\theta_2 = 1$. The ex post incentive constraints for agent 1 at $\theta_2 = 1$ are given by:

$$\begin{array}{l} u_1(f_0(1,1),(1,1)) + v_1(f_1(1,1),1) \geq u_1(f_0(2,1),(1,1)) + v_1(f_1(2,1),1) \\ u_1(f_0(2,1),(2,1)) + v_1(f_1(2,1),2) \geq u_1(f_0(1,1),(2,1)) + v_1(f_1(1,1),2) \end{array} \tag{49}$$

After inserting the payoffs from the common component, and rearranging the utility from the private component, we have

$$v_1(f_1(1,1),1) - v_1(f_1(2,1),1) \geq 1,$$
$$v_1(f_1(2,1),2) - v_1(f_1(1,1),2) \geq 1,$$

but both inequalities cannot be satisfied simultaneously as it follows from (48) that:

$$v_1(f_1(2,1),2) - v_1(f_1(1,1),2) > 0 \Leftrightarrow v_1(f_1(2,1),1) - v_1(f_1(1,1),1) > 0.$$

However we can interim implement the social choice correspondence F for every full support prior. It is easiest to demonstrate this with an independent prior:

$$p_1(\theta_2 = 1 \mid \cdot) = 1 - 2\varepsilon, \quad p_1(\theta_2 = 2 \mid \cdot) = p_1(\theta_2 = 3 \mid \cdot) = \varepsilon, \ \forall \theta_1 \in \Theta_1.$$

We offer different rewards for each payoff type θ_1 of agent 1 at different realizations of θ_2. We use the fact that $v_1(y_1, \theta_1)$ grows at different rates to obtain interim incentive compatibility. More precisely, the following rewards as a function of the announced type accomplish interim implementation for all ε satisfying $0 < \varepsilon < \frac{1}{2}$.

$$f_1(1,1) = 0, \ f_1(1,2) = \frac{3}{\varepsilon}, \ f_1(1,3) = \frac{5}{\varepsilon}, \tag{50}$$

and

$$f_1(2,1) = 0, \ f_1(2,2) = \frac{1}{\varepsilon}, \ f_1(2,3) = \frac{6}{\varepsilon}. \tag{51}$$

(With correlated rather than independent priors, we could use differential probabilities as well as differential rewards to guarantee interim incentive compatibility.) To verify interim incentive compatibility, it suffices to establish that for $\theta_1 = 1$:

$$\sum_{\theta_2 \in \Theta_2} p_1(\theta_2)\left(v_1(f_1(1,\theta_2),1) - v_1(f_1(2,\theta_2),1)\right) \geq 1, \tag{52}$$

and conversely for payoff type $\theta_1 = 2$:

$$\sum_{\theta_2 \in \Theta_2} p_1(\theta_2)\left(v_1(f_1(2,\theta_2),2) - v_1(f_1(1,\theta_2),2)\right) \geq 1. \tag{53}$$

Inserting $f_1(\theta_1, \theta_2)$ from (50) and (51) into (52) and (53), we find:

$$\varepsilon\left(\left(\frac{3}{\varepsilon} + \frac{5}{\varepsilon}\right) - \left(\frac{1}{\varepsilon} + \frac{6}{\varepsilon}\right)\right) \geq 1,$$

and

$$\varepsilon\left(\left(\frac{1}{\varepsilon} + \left(\frac{1}{\varepsilon}\right)^2 + \frac{6}{\varepsilon} + \left(\frac{6}{\varepsilon}\right)^2\right) - \left(\frac{3}{\varepsilon} + \left(\frac{3}{\varepsilon}\right)^2 + \frac{5}{\varepsilon} + \left(\frac{5}{\varepsilon}\right)^2\right)\right) \geq 1,$$

and it is easy to verify that both inequalities are satisfied.