

# Egalitarianism and Egalitarian Equivalence under Incomplete Information

Presenter: Geoffroy de Clippel (Brown University)

March 31, 2009

## Abstract

The theory of social choice has been applied extensively to determine collective actions. Nevertheless, the implication of informational constraints are not yet well understood. This is an important limitation. In many practical scenarios the participants already have some private information when they engage in the cooperative process. Extending the theory of social choice to characterize selection criteria that are applicable to the mechanism design problem is thus an important research agenda. As a step in that direction, we discuss in a first paper (joint with David Wettstein) possible extensions of the egalitarian solution to environments with asymmetric information. In a second paper (joint with David Perez-Castrillo and David Wettstein) we avoid interpersonal comparisons of interim utilities by studying egalitarian equivalence in an exchange economy under incomplete information. Both papers are work in progress at the time of submission to the conference, but I will submit a polished version of these papers before the conference, if accepted.

# Egalitarianism under Incomplete Information

Geoffroy de Clippel\*      David Wettstein†

**VERY PRELIMINARY - PLEASE DO NOT CIRCULATE**

**March 2009**

## 1. INTRODUCTION

The theory of social choice has been applied extensively to determine collective actions. Nevertheless, the implication of informational constraints are not yet well understood. This is an important limitation. In many practical scenarios the participants already have some private information when they engage in the cooperative process. Developing models of cooperation under incomplete information has long been considered and remains a significant open problem in economic theory, as pointed out, for instance, by Professor Aumann in his first presidential address to the Game Theory Society (reproduced in Aumann, 2003). To be more precise, an impressive amount of work, known as the theory of mechanism design, has already been devoted to understand which contracts are feasible under asymmetric information. Professors Hurwicz, Maskin, and Myerson were awarded the 2007 Sveriges riksbank prize in economic sciences in memory of Alfred Nobel for their path-breaking contributions on the topic. Yet, very little is known about what specific contract, among those that are feasible, should be chosen. Authors use the ex-ante utilitarian criterion, most often without justification, when they want to select a specific incentive compatible mechanism.<sup>1</sup> We feel that extending the theory of social choice to characterize selection criteria that will be applicable to the mechanism design problem is an important research agenda.

---

\*Department of Economics, Brown University. Email: declippel@brown.edu

†Department of Economics, Ben-Gurion University of the Negev. Monaster Center for Economic Research, Beer-Sheva 84105, Israel. E-mail: wettstn@bgu.ac.il.

<sup>1</sup>The only axiomatic justification of the ex-ante utilitarian criterion that we are aware of can be found in Nehring (2004). It can be shown that his methodology that consists in finding a social welfare ordering that is consistent with an ex-post criterion and interim Pareto comparisons cannot be applied to social welfare orderings that are anonymous, satisfy the Pigou-Dalton transfer principle, and are different from the utilitarian criterion (see de Clippel, 2009). In other words, his approach cannot help us find social welfare orderings that have even the slightest concern for equity. In the present paper we propose an alternative (interim) methodology based on social choice functions. It allows us to characterize a possible extension of the egalitarian principle.

As a first step in that direction, we discuss in the present paper possible extensions of the egalitarian solution to environments with asymmetric information. More specifically, we will look for solutions that are anonymous, determine mechanisms that are interim incentive efficient, and that satisfy some form of monotonicity, a property that is known to be characteristic of the egalitarian principle under complete information (see Kalai, 1977, for instance). Our first observation is that requiring monotonicity on the whole domain of social choice problems leads to an impossibility result. Contrary to the complete information case, this incompatibility remains even if one restricts attention to large decision sets that allow for transfers and prevent satiation (think for instance of “information rents”). The reason is that incentive constraints may lead to feasible sets of interim utilities that are non-comprehensive and with the possibility of satiation, even in very well-behaved problems (see Example 1). This difficulty will be present throughout the paper, whereby axiomatic results are far more difficult to derive than under complete information, because of the restriction imposed by the informational constraints.

Efficiency is our prime objective, and hence we must weaken our equity criterion, captured mostly by the monotonicity property, in order to escape the impossibility discussed in the previous paragraph. We feel that it may be unreasonable to require the monotonicity property when starting with a mechanism whose associated interim utilities belong to the relative boundary of the interim incentive Pareto frontier. In such cases, we cannot exclude that there exist alternative mechanisms that are better from an equity point of view, but were not selected because they are not second-best efficient. On the other hand, we can be sure that efficiency-first is not a binding constraint when the interim utilities associated to the mechanisms in the solution of the original problem belong to the relative interior of the interim incentive Pareto frontier. Indeed, in such cases, any kind of infinitesimal compensations for some types of some agents can be realized at the expense of others through other interim incentive efficient mechanisms. The restricted monotonicity axiom requires the monotonicity property to apply only in those cases. We prove that this weaker property is compatible with the properties of interim efficiency and anonymity. Actually, we offer a partial axiomatic characterization of the lex-min solution applied to interim utilities after adding the axioms of “interim welfarism,” “exhaustivity,” and “merging identical types” (see Theorem 1)

Interpersonal comparisons of interim utilities comes as a consequence of the axioms. We react in three ways to this fact. First we apply our criterion to classical examples in the mechanism design literature (taxes and public good) under the assumption that utilities are quasi-linear, in which case interpersonal comparisons are easiest to accept. Luckily, most examples in mechanism design fall in that category, simply because characterizing incentive compatible mechanisms in the more general case can be very hard. Second, we pursue Harsanyi’s (1963) methodology (see also Shapley, 1969, and Yaari, 1981) of endogenizing interpersonal comparisons so as to reconcile the utilitarian and the egalitarian principles. Here we will try to combine the ex-ante utilitarian and our interim egalitarian criterion by rescaling the interim utilities. Interestingly, it turns out (see Theorem 2) that this is always feasible, even while requiring our interim egalitarian criterion to hold with equality (no need to resort to the lex-min), it leads to a unique solution,

and results in a characterization of Myerson's (1979) solution, maximizing Harsanyi and Selten's (1972) weighted Nash product over the set of interim utilities that are achievable through some incentive compatible mechanism (see Weidner, 1972, for a direct axiomatic characterization of that solution under the assumption of independent types). Third, in a companion paper (de Clippel et al., 2009), we apply similar ideas to extend the concept of egalitarian equivalence to economies under asymmetric information. Though we do not have an axiomatic characterization of that solution, it has the advantage of being ordinally invariant.

Examples show that some agents may feel that the outcome of our egalitarian solution is actually biased in favor of some other agent *given the information they have*. This motivates another possible extension of egalitarianism in quasi-linear collective choice problems that selects the interim incentive efficient mechanisms that maximize the minimum of the type-agents ratios between their expected utility gains and the total surplus they expect the mechanism to realize. Since we do not have an axiomatic characterization of that second criterion, we only mention it in the concluding section, and study its properties in view of the axioms that have been introduced earlier. We find it interesting that the presence of incomplete information does not only create difficulties in finding selection criteria that satisfy some normative properties because of the incentive constraints, but also leads to different normative criteria, distinguishing a notion of equity from the point of view of an impartial designer and a notion of equity as perceived by the participants themselves, (a distinction that is of course irrelevant under complete information).

## 2. GENERAL MODEL

A *social choice problem under incomplete information* is a quintuple

$$\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}),$$

where  $I$  is the finite set of *individuals*,  $D$  is the set of *collective decisions*,  $d^* \in D$  is the *status-quo*,  $T_i$  is the finite set of individual  $i$ 's possible *types*,  $\pi \in \Delta(T)$  is the *common prior* determining the individuals' *beliefs* ( $T = \times_{i \in I} T_i$ ), and  $u_i : D \times T \rightarrow \mathbb{R}$  is individual  $i$ 's *utility function*, that will be used to determine his interim preferences via the expected utility criterion. We will assume for notational convenience that  $u_i(d^*, t) = 0$ , for all  $t \in T$ . This is without loss of generality if utilities are understood as utility gains over the status-quo.

A (direct) *mechanism* for  $\mathcal{S}$  is a function  $\mu : T \rightarrow \Delta(D)$ . If a mechanism  $\mu$  is implemented truthfully, then individual  $i$ 's expected utility when of type  $t_i$  is given by:

$$U_i(\mu|t_i) = \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) u_i(\mu(t), t).$$

If all the other individuals report their true type, while individual  $i$  reports  $t'_i$  instead of his true type  $t_i$ , then his expected utility is denoted  $U_i(\mu, t'_i|t_i)$ :

$$U_i(\mu, t'_i|t_i) = \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) u_i(\mu(t'_i, t_{-i}), t).$$

The mechanism  $\mu$  is *incentive compatible* if

$$U_i(\mu|t_i) \geq U_i(\mu, t'_i|t_i)$$

for each  $t_i, t'_i$  in  $T_i$  and each  $i \in I$ . The revelation principle (Myerson, 1979) implies that any agreement that is achievable through some form of communication can also be achieved through an incentive compatible direct mechanism. Hence we may restrict attention to those mechanisms without loss of generality.

An incentive compatible mechanism  $\mu$  is *interim individually rational* if  $U_i(\mu|t_i) \geq 0$ , for all  $t_i \in T_i$  and all  $i \in I$ . A mechanism is *feasible* if it is both incentive compatible and interim individually rational. The set of feasible mechanisms will be denoted  $\mathcal{F}(\mathcal{S})$ . The set of interim utilities that are achievable through some feasible mechanism will be denoted by  $\mathcal{U}(\mathcal{S})$ :

$$\mathcal{U}(\mathcal{S}) = \{\mathbf{u}(\mu) | \mu \in \mathcal{F}(\mathcal{S})\},$$

where  $\mathbf{u}(\mu) = (U_i(\mu|t_i))_{t_i \in T_i, i \in I}$ . For notational simplicity, we will restrict attention to social choice problems with the property that  $\mathcal{U}(\mathcal{S})$  is compact and for which there exists  $\mathbf{u} \in \mathcal{U}(\mathcal{S})$  such that  $\mathbf{u} \gg 0$ .

A *social choice function* is a correspondence  $\Sigma$  that associates a nonempty set of feasible mechanisms to each social choice problem:  $\Sigma(\mathcal{S}) \subseteq \mathcal{F}(\mathcal{S})$ , for each  $\mathcal{S}$ . Even though we allow for correspondences, we assume that the image of a social choice function is essentially unique, in the sense that all the individuals must be indifferent (whatever the information they have) between any two mechanisms that belong to the solution of any problem  $\mathcal{S} = (I, D, d^*(T_i)_{i \in I}, (u_i)_{i \in I}, p)$ :

$$(\forall \mu, \mu' \in \Sigma(\mathcal{S}))(\forall i \in I)(\forall t_i \in T_i) : U_i(\mu|t_i) = U_i(\mu'|t_i). \quad (1)$$

### 3. PARTIAL AXIOMATIC RESULT

We now require the social choice function to always exhaust all the benefit of cooperation at the time of agreeing, thereby selecting mechanisms that are efficient (as first formally defined by Holmström and Myerson (1983)).

**Interim Efficiency (I-EFF)** *Let  $\mathcal{S}$  be a social choice problem, and let  $\mu \in \Sigma(\mathcal{S})$ . Then there does not exist a  $\hat{\mu} \in \mathcal{F}(\mathcal{S})$  such that  $U_i(\hat{\mu}|t_i) \geq U_i(\mu|t_i)$ , for each  $t_i \in T_i$ , and each  $i \in I$ , with one of the inequalities being strict.*

The next two properties require the social choice function to be covariant with respect to renaming the individuals and/or their types.

**Anonymity (AN)** *Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  and  $\mathcal{S}' = (I, D, d^*(T'_i)_{i \in I}, (u'_i)_{i \in I}, p')$  be two social choice problems. Suppose that there exist an isomorphism  $f : I \rightarrow I$ , an isomorphism  $g : D \rightarrow D$ , and isomorphisms  $h_i : T_i \rightarrow T'_{f(i)}$  (one for each  $i \in I$ ) such that*

1.  $(\forall t \in T) : p(t) = p'(h(t))$ , and

$$2. (\forall t \in T)(\forall i \in I)(\forall d \in D) : u_i(d, t) = u'_i(g(d), h(t)),$$

with the convention  $h(t) = (h_i(t_i))_{i \in I}$ . Then  $\mu' \in \Sigma(\mathcal{S}')$  if and only if  $\mu \in \Sigma(\mathcal{S})$ , where  $\mu$  is the mechanism for  $\mathcal{S}$  defined as follows: the probability of implementing  $d \in D$  when first individual reports  $t_1 \in T_1, \dots$ , and the  $I^{\text{th}}$  individual reports  $t_I \in T_I$  is equal to the probability of implementing  $g(d)$  under  $\mu'$  when individual  $f(1)$  reports  $h_1(t_1), \dots$ , and individual  $f(I)$  reports  $h_I(t_I)$ .

There are many social choice functions that satisfy the three axioms listed so far. It is helpful to keep in mind the complete information case to make a comparison. Any social choice function that is derived from the maximization of an anonymous social welfare ordering would satisfy these axioms. Many other social choice functions (e.g. relative egalitarianism) would also satisfy them. We must impose a more substantive equity property to narrow down the set of acceptable social choice functions. Monotonicity is a traditional and appealing such property, that is also known to be intimately related to the egalitarian solution under complete information. It means that no individual in the society should be worse off when having additional collective decisions at one's disposal. This property is straightforward to phrase in our framework with incomplete information as well.

**Monotonicity (MON)** *Let  $\mathcal{S}$  and  $\mathcal{S}'$  be two social choice problem. Suppose that  $\mathcal{S}'$  differs from  $\mathcal{S}$  only in that more collective decisions are available:  $I = I', D \subseteq D', T_i = T'_i$ , and  $u_i(d, t) = u'_i(d, t)$ , for each  $i \in I$ , each  $d \in D$ , and each  $t \in T$ . If  $\mu \in \Sigma(\mathcal{S})$  and  $\mu' \in \Sigma(\mathcal{S}')$ , then  $U_i(\mu'|t_i) \geq U_i(\mu|t_i)$ , for each  $t_i \in T_i$ , and each  $i \in I$ .*

**Proposition 1** *There is no social choice function that satisfies I-EFF, AN, and MON.*

This impossibility result already holds, and was first observed by Luce and Raiffa (1957), under complete information. Yet the reader may be puzzled, as these very same axioms also lead to a characterization of the egalitarian solution in Kalai's (1977) framework. Luce and Raiffa's impossibility result does not hold in Kalai's framework because social choice problems are restricted to be comprehensive in the space of utilities and without any "flat parts" (at least at those utility vectors that are individually rational). So one may be tempted to follow a similar route under incomplete information, assuming that the set of collective decisions is rich enough to rule out free disposal and satiation. It is important to realize though that this will not prevent the impossibility stated in Proposition 1. That is, the set of interim utilities that can be achieved through some feasible mechanism need not be comprehensive, even for such rich sets of collective decisions, *because of the presence of incentive constraints*. It is well-known, for instance, that one may be constrained to give some type  $t_i$  of an individual  $i$  a higher expected utility than to another of his types, even in the most regular quasi-linear environments (cf. the concept of "information rent"). In such cases, it is impossible to find another mechanism that would decrease player  $i$ 's expected utility only when of type  $t_i$ , while keeping constant the expected utility of all his other types, and all the types of the other individuals. We illustrate this important observation on a simple example phrased in the language of our model.

**Example 1** Consider a social choice problem with two individuals, 1 and 2, that can be of two types,  $L$  and  $H$ . Each individual knows only his own type, and believes that the two types of the other individual are equally likely. Each individual has up to 10 hours available to work, and his productivity per hour is 1 if his type is  $L$ , and 2 if his type is  $H$ . Allowing for any kind of transfers and free disposal, the set of decisions is thus

$$D = \{(\alpha_1, \alpha_2, m_1, m_2) \in [0, 10]^2 \times \mathbb{R}^2 \mid m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u_i((\alpha, m), t) = \pi_i(t_i)\alpha_i + m_i,$$

for each  $(\alpha, m) \in D$ , each  $i \in \{1, 2\}$  and  $t_i \in \{L, H\}$ , with the convention  $\pi_i(L) = 1$  and  $\pi_i(H) = 2$ , for each  $i \in \{1, 2\}$ . One may think of each individual having access to a similar field, their payoffs being the quantity produced on their own field, which depend on their productivity, modified by any kind of subsidy and taxes.

Let's consider now a feasible mechanism  $(\alpha, m)$  that determines a decision in  $D$  as a function of the individuals' reports. The incentive constraints faced by the first individual can be written as follows:

$$\bar{m}_1(H) - \bar{m}_1(L) \leq \bar{\alpha}_1(L) - \bar{\alpha}_1(H) \leq \frac{\bar{m}_1(H) - \bar{m}_1(L)}{2} \quad (2)$$

where  $\bar{\alpha}_1(L)$  (resp.  $\bar{\alpha}_1(H)$ ) is the average quantity of time the first individual thinks he will have to work given the mechanism when of type  $L$  (resp.  $H$ ), i.e.

$$\bar{\alpha}_1(L) = \frac{1}{2}(\alpha_1(L, L) + \alpha_1(L, H)) \text{ and } \bar{\alpha}_1(H) = \frac{1}{2}(\alpha_1(H, L) + \alpha_1(H, H)),$$

and  $\bar{m}_1(L)$  (resp.  $\bar{m}_1(H)$ ) is the average quantity of time the first individual thinks he will have to work given the mechanism when of type  $L$  (resp.  $H$ ), i.e.

$$\bar{m}_1(L) = \frac{1}{2}(m_1(L, L) + m_1(L, H)) \text{ and } \bar{m}_1(H) = \frac{1}{2}(m_1(H, L) + m_1(H, H)).$$

This implies that  $\bar{m}_1(H) \leq \bar{m}_1(L)$  and  $\bar{\alpha}_1(L) \leq \bar{\alpha}_1(H)$ . If the mechanism is interim incentive efficient, then it must be that  $\bar{\alpha}_1(H) = 10$ . Otherwise, one could construct another feasible mechanism that interim Pareto dominates  $(\alpha, m)$  by slightly increasing both  $\bar{\alpha}_1(L)$  and  $\bar{\alpha}_1(H)$  by a similar amount, while keeping  $\alpha_2$  and  $m$  unchanged. Notice also that the second inequality in (2) must be binding if  $(\alpha, m)$  is interim incentive efficient. Indeed, suppose on the contrary that the inequality is strict. Hence  $\bar{\alpha}_1(L) < 10$  (as otherwise  $\bar{\alpha}_1(L) = \bar{\alpha}_1(H)$ , and (2) implies that  $\bar{m}_1(L) = \bar{m}_1(H)$ , which contradicts the fact that the second inequality is strict). Now we can construct another feasible mechanism that interim Pareto dominates  $(\alpha, m)$  by increasing a bit  $\bar{\alpha}_1(L)$ , while keeping  $\bar{\alpha}_1(H)$ ,  $\alpha_2$  and  $m$  unchanged. A similar reasoning applies to agent 2, and we have:

$$U_1((\alpha, m)|L) = 10 + \frac{\bar{m}_1(L) + \bar{m}_1(H)}{2} \text{ and } U_2((\alpha, m)|L) = 10 + \frac{\bar{m}_2(L) + \bar{m}_2(H)}{2}.$$

If the mechanism  $(\alpha, m)$  belongs to a solution that satisfies AN, then  $U_1((\alpha, m)|L) = U_2((\alpha, m)|L)$  (apply AN with  $f(1) = 2$ ,  $f(2) = 1$ , and  $g_1 = g_2 = id$ , together with (1)). The fact that  $m_1(t) + m_2(t) \leq 0$ , for each  $t \in \{L, H\}^2$  implies that  $\bar{m}_1(L) + \bar{m}_1(H) + \bar{m}_2(L) + \bar{m}_2(H) \leq 0$ , and hence we must conclude that  $U_1((\alpha, m)|L) \leq 10$ ,  $U_2((\alpha, m)|L) \leq 10$ ,  $\bar{m}_1(L) + \bar{m}_1(H) \leq 0$ , and  $\bar{m}_2(L) + \bar{m}_2(H) \leq 0$ . Hence it must also be that  $\bar{m}_1(H) \leq 0$  and  $\bar{m}_2(H) \leq 0$  (remember indeed that  $\bar{m}_1(H) \leq \bar{m}_1(L)$  and  $\bar{m}_2(H) \leq \bar{m}_2(L)$ ), and both  $U_1((\alpha, m)|H)$  and  $U_2((\alpha, m)|H)$  are no larger than 20. We conclude that any solution that satisfies AN and EFF will prescribe in our problem the set of mechanisms  $(\alpha, m)$ , where  $\bar{m}_i(L) = \bar{m}_i(H) = 0$ , and  $\alpha_i(t) = 10$ , for each  $i \in \{1, 2\}$  and each  $t \in \{L, H\}^2$ . There exist multiple such mechanisms, but of course they all lead to the same interim utilities of 10 for the low types and 20 for the high types.

Consider now a similar problem, but where the two individuals can work on a third field, in which case they need to work together and the total productivity is 3 per joint hour of work. For simplicity it will be assumed that the output is always shared equally. Formally, the set of collective decisions is

$$D' = \{(\alpha'_1, \alpha'_2, m_1, m_2) \in [0, 10]^2 \times \mathbb{R}^2 \mid m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u'_i((\alpha', m), t) = 1.5 \min\{\alpha'_1, \alpha'_2\} + m_i,$$

for each  $(\alpha', m) \in D'$ , each  $i \in \{1, 2\}$  and  $t_i \in \{L, H\}$ . AN, I-EFF, and (1) imply that  $(\alpha', m) \in \Sigma(\mathcal{S}')$  if and only if  $(\alpha', m) \in \mathcal{F}(\mathcal{S})$  and  $U_i((\alpha', m)|t_i) = 15$ , for all  $t_i \in \{L, H\}$  and all  $i \in \{1, 2\}$ . An example of such mechanism is given by  $(\alpha'_i(t), m(t)) = (10, 0)$ , for all  $i \in \{1, 2\}$  and all  $t \in \{L, H\}^2$ .

Finally, suppose that the impartial third party can choose to allocate the individuals' time between the three fields:

$$D'' = \{(\alpha_1, \alpha_2, \alpha'_1, \alpha'_2, m_1, m_2) \in [0, 10]^4 \times \mathbb{R}^2 \mid \alpha_1 + \alpha'_1 \leq 10, \alpha_2 + \alpha'_2 \leq 10, m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u''_i((\alpha, \alpha', m), t) = \pi_i(t_i)\alpha_i + 1.5 \min\{\alpha'_1, \alpha'_2\} + m_i,$$

for each  $(\alpha, \alpha', m) \in D''$ , each  $i \in \{1, 2\}$  and  $t_i \in \{L, H\}$ , with the convention  $\pi_i(L) = 1$  and  $\pi_i(H) = 2$ , for each  $i \in \{1, 2\}$ . Notice that

$$\sum_{i=1}^2 \sum_{t_i \in \{L, H\}} U_i((\alpha, \alpha', m)|t_i) = \sum_{t \in \{L, H\}^2} \frac{1}{2} \sum_{i=1}^2 u_i((\alpha(t), \alpha'(t), m(t)), t) \leq 65,$$

for each  $(\alpha, \alpha', m) \in \mathcal{F}(\mathcal{S}'')$ , the last equality following from the fact that the maximal total surplus is 40 when both individuals' type is H and is 30 otherwise. Hence there is no way to find a feasible mechanism that gives an interim utility of at least 15 to the low-type individuals and 20 to the high-type individuals, which contradicts MON, since  $D \subseteq D''$ ,  $D' \subseteq D''$ , and both  $u'_i(d, t) = u_i(d, t)$  and  $u''_i(d, t) = u_i(d, t)$ , for each  $i \in I$ , each  $d \in D$ , and each  $t \in \{L, H\}^2$ .



Free disposal and unlimited transfers are allowed in the three social choice problems from Example (1). Even so, the set of interim utilities that are achievable through feasible mechanisms need not be comprehensive, and improving the welfare of some types of some individuals at the expense of others may be impossible. In the problem  $\mathcal{S}$ , for instance, an individual with an L-type cannot get an interim utility higher than 10 at any mechanism that is interim incentive efficient. This may result in the incompatibility of the efficiency and equity criteria. Though this kind of tension has regularly been discussed in the literature (see Moulin, 1988, Chapter 1), it has never been seen as a consequence of informational constraints. Efficiency is our prime objective, and hence we must weaken our equity criterion, captured mostly by MON, in order to escape the impossibility stated in Proposition 1. We contend that it may be unreasonable to impose MON when comparing  $\mathcal{S}'$  to  $\mathcal{S}$  in Example 1, because the interim utilities associated to the mechanisms in the solution of  $\mathcal{S}$  belong to the relative boundary of  $\mathcal{U}^*(\mathcal{S})$ . In such cases, we cannot exclude the possibility there exist alternative mechanisms that are better from an equity point of view, but are not selected because they are not interim incentive efficient. On the other hand, we can be sure that “efficiency-first” is not a binding constraint when the interim utilities associated to the mechanisms in the solution belong to the relative interior of the interim incentive Pareto frontier. Indeed, in such cases, any kind of infinitesimal compensations for some types of some agents can be realized at the expense of others through other interim incentive efficient mechanisms. It is thus reasonable to impose MON in such cases. This is precisely the content of the next axiom. First we need to make precise what we mean by relative interior.

Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  be a social choice problem. The vector  $\mathbf{u} \in \mathcal{U}^*(\mathcal{S})$  belongs to the *relative interior* of  $\mathcal{U}^*(\mathcal{S})$  if for all set  $C \subseteq T$  there exists  $\mathbf{u}' \in \mathcal{U}^*(\mathcal{S})$  such that  $\mathbf{u}'_c > \mathbf{u}_c$ , for all  $c \in C$ .

**Restricted Monotonicity (R-MON)** *Let  $\mathcal{S}$  and  $\mathcal{S}'$  be two social choice problem. Suppose that  $\mathcal{S}'$  differs from  $\mathcal{S}$  only in that more collective decisions are available:  $I = I'$ ,  $D \subseteq D'$ ,  $T_i = T'_i$ , and  $u_i(d, t) = u'_i(d, t)$ , for each  $i \in I$ , each  $d \in D$ , and each  $t \in T$ . Let  $\mu \in \Sigma(\mathcal{S})$  be such that  $\mathbf{u}(\mu)$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , and let  $\mu' \in \Sigma(\mathcal{S}')$ . Then  $U_i(\mu'|t_i) \geq U_i(\mu|t_i)$ , for each  $t_i \in T_i$ , and each  $i \in I$ .*

R-MON remains silent when it comes to mechanisms  $\mu \in \Sigma(\mathcal{S})$  such that  $\mathbf{u}(\mu)$  that does not belong to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , contrary to MON. The monotonicity property may still hold in some of those cases, but we cannot be sure because the efficiency-first principle may be binding, as already explained.

R-MON is now compatible with both AN and I-EFF. Yet we will need three additional axioms to prove a partial characterization of a unique solution. Most of the theory of social choice under complete information is phrased under the welfarist assumption that only feasible utilities matter, not the underlying decisions that make them feasible.<sup>2</sup> Though it is certainly worthwhile to find more primitive properties to justify that

---

<sup>2</sup>This welfarist assumption that remains implicit in the way classical models are phrased was first emphasized by Roemer (1986).

assumption,<sup>3</sup> or to study context-dependent social choice functions that violate it,<sup>4</sup> we feel that it is important to start by trying to extend the most standard approach to frameworks under incomplete information before finding interesting ways of departing from the benchmark (see one such possible departure in the next section). Understanding what is the right notion of welfarism under incomplete information is not that obvious in itself. A first idea that may come to mind is to require that only the sets of utility vectors that are feasible *ex-post* (i.e. one set for each possible realization of the types), should be sufficient information to determine the solution. This approach is necessarily wrong, as it does not allow to keep track of incentive constraints. Indeed, the fact that a utility vector is feasible at some type profile does not allow to infer what would be the utility that a player would get by reporting a different type. Also, only interim preferences matter when taking an individual decision under incomplete information, and one may take the position that therefore only interim utilities should matter when taking a collective decision. This leads us to formulate our first axiom.

**Interim Welfarism (I-WELF)** *Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  and  $\mathcal{S}' = (I', D', (T'_i)_{i \in I}, (u'_i)_{i \in I}, p')$  be two social choice problems. If  $T_i = T'_i$ , for each  $i \in I$ , and  $\mathcal{U}(\mathcal{S}) = \mathcal{U}(\mathcal{S}')$ , then  $\mathcal{U}(\Sigma(\mathcal{S})) = \mathcal{U}(\Sigma(\mathcal{S}'))$ .*

Of course, this definition boils down to the usual notion of welfarism under complete information, i.e. when each type set is a singleton.

For the next axiom, say that two types  $t_i$  and  $t'_i$  are *identical* in the problem  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  if individual  $i$ 's interim beliefs are identical when of type  $t_i$  and  $t'_i$ , his utility function in any state  $t$  is identical to his utility function in the corresponding state  $(t'_i, t_{-i})$ , and all the other individuals see  $t_i$  and  $t'_i$  as equally likely:  $p(t_{-j}|t_j) = p(t'_{-j}|t_j)$  for all  $t_{-j} \in T_{-j}$ . As a mild consistency property, we require that identical types can be merged without making any essential change to the solution.

**Merging Identical Types (MIT)** *Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  be a social choice problem such that  $t_i$  and  $t'_i$  are identical. Let  $\mathcal{S}'$  be the problem derived from  $\mathcal{S}$  by “merging”  $t_i$  and  $t'_i$  into a single type  $\hat{t}_i$ . Then  $\mu \in \Sigma(\mathcal{S})$  if and only if  $\mu' \in \Sigma(\mathcal{S}')$ , where  $\mu'(t) = \mu(t)$ , for all  $t \in T$  such that  $t_i \neq \hat{t}_i$ , and  $\mu'(\hat{t}_i) = \mu(t_i, t_{-i})$  (also equal to  $\mu(t'_i, t_{-i})$ , by AN), for all  $t_{-i} \in T_{-i}$ .*

The last axiom requires that if a feasible mechanism generates the same interim utilities as another mechanism in the solution of a problem, then it also belongs to the solution of that problem. This property is always implicit in any welfarist model, but we must make it explicit in our model.

**Exhaustivity (EX)** *Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  be a social choice problem. If  $\mu \in \Sigma(\mathcal{S})$  and  $\mu'$  is a feasible mechanism such that  $U_i(\mu'|t_i) = U_i(\mu|t_i)$ , for all  $t_i \in T_i$  and all  $i \in I$ , then  $\mu' \in \Sigma(\mathcal{S})$ .*

We now define the lex-min solution that is partially characterized in the next theorem, and that boils down to the usual egalitarian criterion under complete information. A

---

<sup>3</sup>See e.g. ...

<sup>4</sup>See e.g. ...

mechanism  $\mu \in \mathcal{F}(\mathcal{S})$  belongs to the *lex-min solution* of the social choice problem  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$ ,  $\mu \in \Sigma^*(\mathcal{S})$  if and only if  $\theta(\mathbf{u}(\mu))$  maximizes  $\theta(\mathbf{u})$  according to the lexicographic order over all  $\mathbf{u} \in \mathcal{U}(\mathcal{S})$ , where  $\theta : \times_{i \in I} \mathbb{R}_+^{T_i} \rightarrow \times_{i \in I} \mathbb{R}_+^{T_i}$  is the function that rearrange the components of a vector increasingly.

**Theorem 1** *The lex-min solution  $\Sigma^*$  satisfies I-EFF, AN, R-MON, I-WELF, MIT. In addition, if  $\Sigma$  is a solution that satisfies the five axioms, then*

a) *If there exists  $\mu \in \Sigma^*(\mathcal{S})$  such that the vector  $(U_i(\mu|t_i))_{t_i \in T_i, i \in I}$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , then  $\Sigma(\mathcal{S}) = \Sigma^*(\mathcal{S})$ ;*

b) *If there exists  $\mu \in \Sigma(\mathcal{S})$  is such that the vector  $(U_i(\mu|t_i))_{t_i \in T_i, i \in I}$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , then  $\Sigma(\mathcal{S}) = \Sigma^*(\mathcal{S})$ .*

Sketch of Proof: It is not difficult to check that the lex-min solution satisfies the axioms. So we will start by proving a), while assuming that  $\Sigma$  satisfies the five axioms.

We may assume without loss of generality that all the type sets have equal cardinality. If not, then one can follow a similar reasoning after duplicating some types, and then conclude the proof by applying MIT. We can also assume without loss of generality that  $p$  is the uniform probability distribution. Indeed, otherwise one apply a similar reasoning to a modified problem with the same sets of types, the same set of collective decisions, a uniform common prior, and utility functions  $\hat{u}_i(d, t) := |T_{-i}|p(t_{-i}|t_i)u_i(d, t)$ . Notice that this modification does not change the individuals' interim evaluations of any mechanism since the products of the conditional probabilities with the state-contingent utilities remain constant (see Myerson, 1984, Section 3). Hence the set of interim utilities that are achievable through some feasible mechanism is the same in both problems, and one will be able to conclude the proof by applying I-WELF.

To summarize, we are looking at a problem  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$ , where  $|T_i| = |T_j|$ , for all  $i, j \in I$ , and  $p$  is uniform. Let  $\mu \in \Sigma^*(\mathcal{S})$  be such that the vector  $\mathbf{u}(\mu)$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ . In that case, it must be that  $U_i(\mu|t_i) = U_j(\mu|t_j)$ , for all  $t_i \in T_i$ , all  $t_j \in T_j$ , and all  $i, j \in I$ . Indeed, otherwise, let  $C$  be the set of couples  $(i, t_i)$  such that  $U_i(\mu|t_i) \leq U_j(\mu|t_j)$ , for all  $t_j \in T_j$  and all  $j \in I$ . Hence  $U_i(\mu|t_i) < U_j(\mu|t_j)$ , for all  $(i, t_i) \in C$  and all  $(j, t_j)$  that does not belong to  $C$ . Since  $\mathbf{u}(\mu)$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , there exists a mechanism  $\mu' \in \mathcal{F}(\mathcal{S})$  such that  $U_i(\mu'|t_i) > U_i(\mu|t_i)$ , for all  $(i, t_i) \in C$ . For each  $\epsilon \in ]0, 1[$ ,  $\mu^\epsilon := \epsilon\mu' + (1 - \epsilon)\mu \in \mathcal{F}(\mathcal{S})$ . For  $\epsilon$  small enough, the smallest component of  $\mathbf{u}(\mu^\epsilon)$  is strictly larger than the smallest component of  $\mathbf{u}(\mu)$ . Hence one reaches a contradiction with  $\mu \in \Sigma^*(\mathcal{S})$ . So it must be indeed that all the components of  $\mathbf{u}(\mu)$  are identical.

Since  $\mathbf{u}(\mu)$  belongs to the relative interior of  $\mathcal{U}^*(\mathcal{S})$ , there must exist  $\lambda \in \times_{i \in I} \mathbb{R}_{++}^{T_i}$  such that  $\mu$  maximizes  $\sum_{i \in I} \sum_{t_i \in T_i} \lambda_i(t_i) U_i(\nu|t_i)$ , over all  $\nu \in \mathcal{F}(\mathcal{S})$ . Let  $W^\lambda$  be the value of that weighted sum evaluated at  $\mu$ . Following Myerson's virtual utility construction (see also the Appendix of Weidner, 1992), it is possible to construct an auxiliary problem  $\mathcal{S}' = (I, D', (T_i)_{i \in I}, (u'_i)_{i \in I}, p)$ , where  $D' = D \cup \{d_{i,t_i} | t_i \in T_i, i \in I\}$ ,  $u'_i(d, t) = u_i(d, t)$ , for all  $d \in D$ ,  $t \in T$ , and  $i \in I$ , and such that  $\mathcal{U}(\mathcal{S}')$  is the convex hull of the vectors 0 and  $\mathbf{u}^{i,t_i}$ , for each  $t_i \in T_i$  and each  $i \in I$ , where  $\mathbf{u}_j^{i,t_i}(t_j) = 0$ , for all  $(j, t_j) \neq (i, t_i)$ , and  $\mathbf{u}_i^{i,t_i}(t_i) = W^\lambda / \lambda_{i,t_i}$ .

For each combination  $(i, t_i)$ , it is possible to find a collective decision  $\hat{d}_{i,t_i}$  and utility functions  $u''$  such that

1.  $U_i''(\hat{d}_{i,t_i}|t'_i) = 0$  if  $t'_i \neq t_i$ ;
2.  $U_i''(\hat{d}_{i,t_i}|t_i) = W^\lambda/\lambda_i(t_i)$ ;
3.  $U_j''(\hat{d}_{i,t_i}|t_j) = 0$ , for all  $j \in N \setminus \{i\}$  and all  $t_j \in T_j$ ;
4.  $\sum_{j \in I} u_j''(\hat{d}_{i,t_i}, t') \leq \sum_{j \in I} \frac{\lambda_j(t_j)}{p(t_j)} U_j''(\mu|t_j)$ , for all  $t' \in T$

(obtained by solving a system of linear equations - to be completed (details available upon request from the authors)). Let  $D'' = \{d^*\} \cup \{\hat{d}_{i,t_i} | i \in I, t_i \in T_i\}$ . Clearly,  $\mathcal{U}(D'') = \mathcal{U}(S')$ .

We conclude the proof with  $|I| = 2$  for simplicity (should be easy to extend to any number of individuals - remains to be done). Take  $\epsilon > 0$  sufficiently small such that such that the convex hull  $\mathcal{C}$  of the vectors  $(0, 0)$ ,  $(U_1(\mu|t_1), U_2(\mu|t_2))$ ,  $(U_1(\mu|t_1) + \epsilon, 0)$ , and  $(0, U_2(\mu|t_2) + \epsilon)$  is included in the half-space  $\{x \in \mathbb{R}^I | \sum_{i \in I} \lambda_i(t_i)x_i \leq \sum_{i \in I} \frac{\lambda_i(t_i)}{p(t_i)} U_i(\mu|t_i)\}$ , for all  $t \in T$ . Take now the problem  $S'''$  with the same information structure as all the other problems, but with a set of collective decisions  $D''' = \mathcal{C}^T$  and utility functions  $u'''$  defined as follows:  $u_i'''(d''', t) = d_i'''(t)$ , for all  $t \in T$ . The way  $\epsilon$  was chosen guarantees that  $\mathcal{U}(S''') = \mathcal{U}(S''')$ , where  $D'''' = D'' \cup D'''$ . Observe that the problem  $S'''$  is symmetric, and hence AN and (1) imply that the interim utility of any mechanism in the solution of that problem must give equal interim utility to all the players and whatever their private information. I-EFF and EX imply that the constant mechanism that gives  $(U_1(\mu|t_1), U_1(\mu|t_1))$ , for all  $t \in T$ , belongs to  $\Sigma(S''')$ . It is easy to check that the interim utility of that constant mechanism belongs to the relative interior of  $\mathcal{U}(S''')$ . R-MON implies that it also belong to  $\Sigma(S''')$ . EX implies that any feasible mechanism that gives the same vector of interim utilities also belongs to  $\Sigma(S''')$ . Let's choose one that can be expressed via lotteries on  $D''$ . R-MON implies that it must remain a solution to  $S''$ . I-WELF implies that any mechanism in the solution of  $S'$  must have the same interim utilities, and  $\mu$  must thus belong to  $\Sigma(S')$ , by EX. Finally, R-MON implies  $\mu \in \Sigma(S)$ , as desired.

As for Condition b), notice that R-MON implies that  $\mu$  must still be a solution in the extended "linearized" problem defined via Myerson's virtual utility (similar to the construction of  $S'$  earlier in the proof. Notice that part a) applies to that extended problem, since there exists an interim incentive efficient mechanism that gives the same interim utility to all the individuals, independently of their information. Hence the solution of that extended problem coincides with its lex-min solution, and  $\mu \in \Sigma^*(S)$ , since  $\Sigma^*$  satisfies R-MON. ■

The theorem can be used as follows. One can compute the lex-min solution to any given problem. If it falls in the relative interior of the interim incentive Pareto frontier, then we know, by a), that this is the outcome of any solution satisfying the axioms. Otherwise, we know, by b), that the outcome of any solution that satisfies the axioms must fall on the relative boundary of the interim incentive Pareto frontier, as does the

lex-min solution. Some slight variations of the axioms we have studied lead actually to an axiomatic characterization of the lex-min solution (see Imai, 1983, and Chun and Peters, 1989), but we have not been able to extend these results to social choice problems under incomplete information, nor find other solutions that satisfy those axioms. This remains an open question. As illustrated by Example 1, and the intricacy of the proof of Theorem 1, the difficulty to resolve this question lies in the fact that the presence of asymmetric information and, particularly, incentive constraints restricts the sets of interim utilities one may consider.

#### 4. APPLICATION IN QUASI-LINEAR ENVIRONMENTS

#### 5. EGALITARIANISM AND UTILITARIANISM RECONCILED

**Theorem 2** *Let  $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, (u_i)_{i \in I}, p)$  be a social choice problem, and let  $\mu^* \in \mathcal{F}(\mathcal{S})$ . Then*

$$\mu^* \in \arg \max_{\mu \in \mathcal{F}(\mathcal{S})} \prod_{i \in I} \prod_{t_i \in T_i} [U_i(\mu|t_i)]^{p(t_i)} \quad (3)$$

*if and only if  $\mu^*$  satisfies the two following conditions for some  $\lambda \in \times_{i \in I} \mathbb{R}_{++}^{T_i}$ :*

1.  $\mu^* \in \arg \max_{\mu \in \mathcal{F}(\mathcal{S})} \sum_{i \in I} \sum_{t_i \in T_i} p(t_i) U_i^\lambda(\mu|t_i)$ ,
2.  $(\forall t \in T)(\forall i \in I) : U_i^\lambda(\mu^*|t_i) = U_j^\lambda(\mu^*|t_j)$ ,

*where  $U_i^\lambda(\mu|t_i) := \lambda_i(t_i) U_i(\mu|t_i)$ , for each  $t_i \in T_i$  and each  $i \in I$ .*

Proof: Let  $W^* = \prod_{i \in I} \prod_{t_i \in T_i} [U_i(\mu|t_i)]^{p(t_i)}$ . Since we assume that there exists at least one element of  $\mathcal{U}(\mathcal{S})$  with only strictly positive components, it must be that  $W^* > 0$  and  $\{(U_i(\mu^*|t_i))_{t_i \in T_i, i \in I}\} \gg 0$  under (3). The sets  $\mathcal{U}(\mathcal{S})$  and  $\{\mathbf{u} \in \times_{i \in I} \mathbb{R}_+^{T_i} \mid \prod_{i \in I} \prod_{t_i \in T_i} \mathbf{u}_i(t_i)^{p(t_i)} \geq W^*\}$  are both closed and convex. Under (3), their intersection is the singleton  $\{(U_i(\mu^*|t_i))_{t_i \in T_i, i \in I}\}$ . Hence the separating hyperplane theorem implies that (3) is equivalent to the existence of a vector  $l \in \times_{i \in I} \mathbb{R}^{T_i}$  for which the two following conditions hold:

1.  $\mu^* \in \arg \max_{\mu \in \mathcal{F}(\mathcal{S})} \sum_{i \in I} \sum_{t_i \in T_i} l_i(t_i) U_i(\mu|t_i)$ ,
2.  $l$  is proportional to the gradient of the curve  $\{\mathbf{u} \in \times_{i \in I} \mathbb{R}_+^{T_i} \mid \prod_{i \in I} \prod_{t_i \in T_i} \mathbf{u}_i(t_i)^{p(t_i)} \geq W^*\}$  at  $(U_i(\mu^*|t_i))_{t_i \in T_i, i \in I}$ .

The second condition itself is equivalent to the existence of a strictly positive number  $\alpha$  such that

$$l_i(t_i) = \frac{\alpha p(t_i)}{U_i(\mu^*|t_i)},$$

for all  $t_i \in T_i$  and all  $i \in I$ . The result thus follows, by taking  $\lambda_i(t_i) = l_i(t_i)/p(t_i)$ , for all  $t_i \in T_i$  and all  $i \in I$ . ■

#### 6. CONCLUDING COMMENTS

## References

- Aumann, R. J., 2003. Presidential Address. *Games and Economic Behavior* **45**, 2-14.
- Chun, Y., and H. Peters, 1989. Lexicographic Monotone Path Solutions. *OR Spektrum* **11**, 43-47.
- de Clippel, G., 2009. A Comment on “The Veil of Public Ignorance.” *In preparation*.
- de Clippel, G. and E. Minelli, 2004. Two-Person Bargaining with Verifiable Information. *Journal of Mathematical Economics* **40**, 799-813.
- de Clippel, G., D. Perez-Castrillo, and D. Wettstein, 2009. Egalitarian Equivalence under Asymmetric Information. *Mimeo*.
- Harsanyi, J. C., 1963. A Simplified Bargaining Model for the n-Person Cooperative Game. *International Economic Review* **4**, 194-220.
- Harsanyi, J. C. and R. Selten, 1972. A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information. *Management Science* **18**, 80-106.
- Holmström, B. and R. B. Myerson, 1983. Efficient and Durable Decision Rules with Incomplete Information. *Econometrica* **51**, 1799-1819.
- Imai, H., 1983. Individual Monotonicity and Lexicographic Maxmin Solution. *Econometrica* **51**, 389-402.
- Kalai, E., 1977. Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons. *Econometrica* **45**, 1623-1630.
- Moulin, H., 1988. Axioms of cooperative decision making. Cambridge University Press.
- Myerson, R. B., 1979. Incentive Compatibility and the Bargaining Problem. *Econometrica* **47**, 62-73.
- Myerson, R. B., 1984. Two-Person Bargaining Problems with Incomplete Information. *Econometrica* **52**, 461-487.
- Nehring, K., 2004. The Veil of Public Ignorance. *Journal of Economic Theory* **119**, 247-270.
- Shapley, L. S., 1969. Utility Comparisons and the Theory of Games. In: Guilbaud, G. Th. (Ed.), *La Decision*, CNRS, 251-263.
- Yaari, M. E., 1981. Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-Examined. *Journal of Economic Theory* **24**, 1-39.

# Egalitarian Equivalence under Asymmetric Information

Geoffroy de Clippel\*  
Brown University

David Pérez-Castrillo†  
Universitat Autònoma de Barcelona

David Wettstein‡  
Ben-Gurion University of the Negev and Universitat Autònoma de Barcelona

April 2008

## Abstract

*JEL classification:* D62, C71.

*Keywords:* Pareto Efficient, Egalitarian Equivalent, Asymmetric Information, Implementation

## 1 Introduction

Fairness and efficiency are two criteria often adhered to by policy makers, arbitrators settling disputes, managers deciding on compensation packages and feature constantly in economic and social debates. These properties were first discussed and analyzed in a complete information setting. In economic environments, this discussion led to the notion of fair outcomes which are efficient and envy free allocations (Varian, 1974) and Egalitarian Equivalent allocations (Pazner and Schmeidler, 1978).<sup>1</sup>

In environments where agents hold private information, incentives must be taken into account. The role of information and incentives is one of the central

---

\*Department of Economics, Brown University. Email: Geoffroy\_DeClippel@brown.edu

†Department of Economics & CODE, Universitat Autònoma de Barcelona. E-mail: david.perez@uab.es.

‡Department of Economics, Ben-Gurion University of the Negev. Monaster Center for Economic Research, Beer-Sheva 84105, Israel. E-mail: wettstn@bgu.ac.il.

<sup>1</sup>These issues were also of great concern in the theory of cooperative games. For the special case of transferable utility environments the leading solution concept embodying fairness and efficiency is the Shapley value (Shapley, 1953). For more general classes of games there is the Nash bargaining solution (Nash, 1950) associating with each bargaining problem a reasonable compromise, the Kalai (1977) egalitarian approach and the various generalizations of the Shapley value to non-transferable utility environments (see, for instance, Harsanyi, 1959).

topics in economics since the end of the sixties. However, most effort has been devoted to understanding what is achievable in the presence of informational constraints<sup>2</sup> and analyzing non-cooperative games.<sup>3</sup> Only few papers discuss criteria to select a socially appealing incentive compatible mechanism (Harsanyi, 1959, Myerson, 1984, Widener, 1992, and Nehring, 2004). Hence, extending the theory of social choice to address problems under incomplete information is a seldom explored research agenda.

We look at economic environments with asymmetric information, and focus on the “interim” stage when agents are privately informed. Efficiency at this moment requires to take into account both the gains from insurance and the agents’ incentives to possibly misrepresent their information. The notion of interim incentive efficient mechanisms (Myerson, 1984) provides a suitable idea of efficiency.

In this paper, we extend the egalitarian approach underlying the concept of egalitarian equivalence to such environments. An allocation is egalitarian equivalent in an economy with complete information (see Pazner and Schmeidler, 1978) with respect to a reference bundle if all agents are indifferent between the proposed allocation and a bundle proportional to the reference bundle. That is, measuring the agents’ surplus in terms of the reference bundle, all obtain the same surplus. In a similar spirit, we say that a mechanism is interim egalitarian equivalent if, in any interim situation, all the agents are indifferent, in expected terms, between the proposed mechanism and receiving a fixed proportion of the reference bundle in each possible type profile. Therefore, in an interim egalitarian equivalent mechanism, at any interim stage, all the agents enjoy the same surplus (in terms of the reference bundle).

Under complete information, egalitarian equivalent allocations that are also (ex-post) Pareto efficient always exist. Our main result states that mechanisms that are both interim egalitarian equivalent and interim incentive also exist, at least in economies with non-exclusive information. That is, efficiency and egalitarianism are compatible concepts in economies with asymmetric but non-exclusive information. We also show that interim efficiency and egalitarian equivalence may not be compatible when the non-exclusive information assumption is violated. This incompatibility is reminiscent of the non-compatibility between equity and efficiency under complete information in general economic environments different than classical pure exchange economies (see Pazner and Schmeidler, 1974, and Maniquet, 1999)

We also address the natural question whether agents can reach egalitarian equivalent mechanisms through non-cooperative behavior. We show that simple adaptations of the constructions proposed by Crawford (1979) and Demange (1984) to implement egalitarian outcomes in complete information economies yield in any economy, as Bayesian equilibria, the set of mechanisms that are both interim egalitarian equivalent and interim incentive efficient.

In the next section, we present the framework and the classical definitions

---

<sup>2</sup>The revelation principles (Gibbard, 1973, Green and Laffont, 1977, and Myerson, 1979) have been a powerful tool in this task.

<sup>3</sup>See, for instance, the development of the principal-agent literature.



while, in Section 3, we introduce the notion of interim egalitarian equivalence. In Section 4 we prove our main existence result and, in Section 5 we present an extended example highlighting the incompatibility between egalitarianism and efficiency when non-exclusive information is violated. Section 6 is devoted to provide an implementation of our solution. Finally, in Section 7 we highlight additional properties of the solution, suggest a weaker concept for those economies where interim egalitarian equivalent and interim incentive efficient mechanisms do not exist, and further discuss our approach.

## 2 The Framework and Classical Definitions

An *economy* is a 6-tuple

$$(N, L, (T_i)_{i \in N}, \pi, e, (u_i)_{i \in N}),$$

where  $N$  is the set of agents,  $L$  is the set of goods,  $T_i$  is agent  $i$ 's set of possible types,  $\pi \in \Delta(T)$  ( $T = \times_{i \in N} T_i$ ) is the common prior describing the relative probability of the types,  $e \in \mathbb{R}_+^L \setminus \{0\}$  is the aggregate endowment of the economy in each possible state  $t$ , and  $u_i : \mathbb{R}^L \times T \rightarrow \mathbb{R}$  is a concave, continuous and strongly increasing utility function that represents the preferences of agent  $i$  (lotteries are evaluated according to the expected utility criterion). For easy of notation, we also denote by  $N$ ,  $L$  and  $T$  the number of elements in the corresponding sets. We assume without loss of generality that each type of each agent comes with a strictly positive probability, i.e. for all  $t_i \in T_i$  and all  $i \in N$  there exists  $t_{-i}$  such that  $\pi(t_i, t_{-i}) > 0$ .

Since types are private information, it may be profitable for the agents to communicate before agreeing on an allocation. Formalizing this idea, a *mechanism* is a function  $\mu : \times_{i \in N} M_i \rightarrow \mathbb{R}_+^{L \times N}$ , where  $M_i$  is any finite set of "messages." Agents are assumed to play a Bayes-Nash equilibrium in the game induced by the mechanism. The revelation principle (Myerson, 1979) allows us, without loss of generality, to restrict attention to direct mechanisms (i.e.  $M_i = T_i$ , for each  $i \in N$ ) for which truth-telling forms a Bayes-Nash equilibrium, that is, the mechanism is incentive compatible. To formally define this property, note that if all the other agents report their types truthfully, then agent  $i$ 's expected utility when reporting  $t'_i$  in the direct mechanism  $\mu$ , while being of type  $t_i$ , is

$$U_i(\mu, t'_i | t_i) = \sum_{t_{-i} \in T_{-i}} \pi(t_{-i} | t_i) u_i(\mu(t'_i, t_{-i}), t),$$

where  $\pi(t_{-i} | t_i)$  denotes the conditional probability of  $t_{-i}$  given  $t_i$ . For simplicity, we will write  $U_i(\mu | t_i)$  instead of  $U_i(\mu, t_i | t_i)$ . The mechanism  $\mu$  is *incentive compatible* if

$$U_i(\mu | t_i) \geq U_i(\mu, t'_i | t_i)$$

for each  $t_i, t'_i$  in  $T_i$  and each  $i \in N$ . A mechanism  $\mu$  is *incentive feasible* if it is incentive compatible and *feasible*, that is  $\sum_{i \in N} \mu_i(t) \leq e$ , for all  $t \in T$ .

Efficiency is a prerequisite for any cooperative solution. Its content was first formalized under incomplete information by Holmström and Myerson (1983). An incentive compatible mechanism  $\mu'$  *interim Pareto dominates* an incentive compatible mechanism  $\mu$  if  $U_i(\mu'|t_i) \geq U_i(\mu|t_i)$  for all  $t_i \in T_i$  and all  $i \in N$ , with at least one of the inequalities being strict. A mechanism is *interim incentive efficient* if it is incentive feasible, and it is not interim Pareto dominated by any other incentive feasible mechanism.

### 3 Egalitarian Equivalence

Efficiency is a necessary condition for a cooperative solution to be appealing, but it is not sufficient, as it remains silent regarding the distribution of the gains derived from cooperation. Pazner and Schmeidler (1978) made an interesting proposal to select a subset of Pareto efficient allocations under complete information, i.e. when the type sets are singletons. In order to obtain a solution that depends only on the ordinal information encoded in the agents' preferences, they proposed to measure cooperative gains in the space of goods following the direction given by a reference bundle  $d \in \mathbb{R}_+^L \setminus \{0\}$ . For each allocation  $a \in \mathbb{R}_+^{L \times N}$  and each agent  $i$ , let  $\lambda_i^a$  be the real number defined by the following equation:<sup>4</sup>

$$u_i(a_i) = u_i(\lambda_i^a d).$$

The allocation  $a$  is *egalitarian equivalent* (along  $d$ ) if  $\lambda_i^a = \lambda_j^a$  for all  $i, j \in N$ . Pazner and Schmeidler proposed to restrict attention to those allocations that are Pareto efficient and egalitarian equivalent, and prove existence and uniqueness under mild assumptions.

The purpose of our paper is to extend Pazner and Schmeidler's solution to environments with incomplete information (for any finite set  $T_i$ ,  $i = 1, \dots, n$ ), and study its properties. One may be tempted to simply look for the mechanism that associates to each  $t$  a Pareto efficient egalitarian equivalent allocation in that ex-post economy. This way to proceed is wrong for at least two reasons. First, that mechanism need not be incentive compatible, and thereby impossible to put into practice. Second, it does not exploit the possibility of mutually beneficial insurance. In other words, it would be incompatible with interim incentive efficiency in most economies. Agents know only their own type when choosing the mechanism. The solution concept should thus be based on their preferences at that point in time (*interim*, and not *ex-post*). Let  $d \in \mathbb{R}_+^L \setminus \{0\}$  be the reference vector.<sup>5</sup> For each incentive compatible mechanism  $\mu$  and each type  $t_i$  of each agent  $i$ , let  $\lambda_i^\mu(t_i)$  be the real number defined by the following equation:

$$U_i(\mu|t_i) = U_i(\lambda_i^\mu(t_i)d|t_i).$$

<sup>4</sup>We omit the vector  $t$  of types in the equation, since it is assumed to be common knowledge in this paragraph.

<sup>5</sup>Most of the analysis extends to the case where  $d$  varies with  $t$ , at the cost of heavier notations.

This means that agent  $i$  of type  $t_i$  is indifferent between participating to the mechanism  $\mu$  and receiving the fixed proportion  $\lambda_i^\mu(t_i)$  of the bundle  $d$  in each possible type profile (for the other agents). We propose a criterion according to which an incentive feasible mechanism  $\mu$  is “equitable” if, at any interim event, all the agents obtain the same (interim) gains (as measured by the vector  $d$ ).

**Definition 1** *An incentive compatible mechanism  $\mu$  is interim egalitarian equivalent if for all  $t \in T$  with  $\pi(t) > 0$ ,  $\lambda_i^\mu(t_i) = \lambda_j^\mu(t_j)$  for all  $i, j \in N$ .*

An incentive compatible mechanism is interim egalitarian if, given any interim situation where agents privately know their types  $(t_i)_{i \in N}$  and for any two agents  $i$  and  $j$  (whose types are  $t_i$  and  $t_j$ ), the gains of agent  $i$  coincide with the gains of agent  $j$  when both gains are measured in terms of the vector  $d$ .

The next section is devoted to the study of *interim equitable* mechanisms, defined as follows:

**Definition 2** *A mechanism is called interim equitable if it is both interim egalitarian equivalent and interim incentive efficient.*

## 4 On the Compatibility of Interim Egalitarian Equivalence and Interim Incentive Efficiency

We establish the existence of interim equitable mechanisms in environments characterized by *non-exclusive information* (NEI). An environment satisfies NEI if for any agent  $i$  and any  $t_{-i} \in \prod_{j \neq i} T_j$  there exist a unique  $t_i^* \in T_i$  such that  $\pi(t_i^*, t_{-i}) > 0$ . This property, introduced by Postlewaite and Schmeidler (1986), means the pooled information of any  $n-1$  agents uniquely determines the profile of types.

We also recall that, for a given  $\pi$ , a non-empty subset  $B \subset T$ , is said to be common knowledge if  $\pi(\hat{t}_{-i}, t_i) = 0$  for all  $i \in N$ ,  $t \in B$  and  $(\hat{t}_{-i}, t_i) \notin B$ . The following two lemmata will be used in the existence proof.

**Lemma 3** *For any  $t^* \in T$  with  $\pi(t^*) > 0$ , define the set  $B(t^*) = \{t \in T \mid \exists \text{ a finite sequence } t^s \in T \text{ with } \pi(t^s) > 0, s = 1, \dots, n_s \text{ and } t^1 = t^*, t^{n_s} = t \text{ such that for all } s \text{ there exists a } j \in N \text{ for which } t_j^s = t_j^{s+1}\}$ . The event  $B(t^*)$  is common knowledge.*

**Proof.** *The set is non-empty and for any  $(\hat{t}_{-i}, t_i) \notin B(t^*)$  it must be the case that  $\pi(\hat{t}_{-i}, t_i) = 0$ , otherwise it would have been in the set  $B(t^*)$ . ■*

**Lemma 4** *Let the environment satisfy NEI and let  $\hat{T}$  be the support of  $\pi$ . For any feasible allocation  $a : \hat{T} \rightarrow \mathbb{R}_+^{L \times N}$ , there exists an incentive feasible mechanism  $\mu$  over  $T$  such that  $U_i(\mu|t_i) = \sum \pi(t_{-i}|t_i)u_i(a(t), t)$ .*

**Proof.** *We define a mechanism  $\mu$  as follows:  $\mu(t) = a(t)$  for all  $t \in \hat{T}$  and  $\mu(t) = 0$  for all  $t \notin \hat{T}$ . If all agents report truthfully no agent can gain by deviating since by NEI any deviation will yield a  $t' \notin \hat{T}$  and result in receiving zero rather than a non-negative bundle, not increasing the deviating agent’s payoff. Hence, the mechanism  $\mu$  is incentive feasible. ■*

This lemma guarantees that, in environments with asymmetric information satisfying NEI, any feasible allocation yielding non-negative bundles to the agents on those situations that happen with positive probability, can be implemented as an interim feasible mechanism over  $T$ .

The next proposition shows that in environments satisfying NEI the set of interim equitable mechanisms is not empty and usually consists of a singleton.

**Proposition 5** *Let the environment satisfy NEI, then (a) An interim equitable mechanism exists.*

*(b) For any minimal common knowledge event  $B$ ,  $\lambda_i^\mu(t_i) = \lambda_j^\mu(t_j)$  for all  $t \in B$ ,  $i, j \in N$ , and interim equitable  $\mu$ .*

*(c) Agents are indifferent between any two interim equitable mechanisms  $\mu$  and  $\eta$ , i.e.,  $\lambda_i^\mu(t_i) = \lambda_i^\eta(t_i)$  for any  $i \in N, t_i \in T_i$ .*

*(d) If agents' preferences are strictly convex, then there exists a unique interim equitable mechanism.*

**Proof.** Denote by  $\mathcal{B}$  the set of minimal common knowledge events for  $\pi$ . It is easy to see that  $\mathcal{B}$  constitutes a partition of  $\widehat{T}$ . For any  $t \in \widehat{T}$ , agent  $i \in N$ , and feasible allocation  $a$ , agent  $i$  (of type  $t_i$ )' expected utility  $U_i(a|t_i)$  only depends on the set  $\{a(t^\circ)\}_{t^\circ \in B(t)}$ , where  $B(t)$  is the (minimal) common knowledge event defined in Lemma 3. For each  $B \in \mathcal{B}$ , define:

$$\lambda_B = \max \{ \lambda | \exists \text{ feasible allocation } a \text{ over } B \text{ s.t. } U_i(a|t_i) = U_i(\lambda d|t_i) \text{ for all } i \in N, t \in B \}.$$

The previous  $\lambda_B$  exists: the set of possible  $\lambda$ s is not empty given that  $a = 0$  is feasible, hence  $\lambda = 0$  belongs to the set; the set is bounded from above as it is not feasible to sustain unbounded utility levels for every agent; finally, by continuity of the utility functions, this set of  $\lambda$ s is closed and contains a maximum.

Denote by  $a_B$  a feasible allocation over  $B$  such that  $u_i(a_B|t_i) = u_i(\lambda_B d|t_i)$  for all  $i \in N, t \in B$ . Also define the feasible allocation  $a^*$  over  $\widehat{T}$  as  $a^*(t) = a_B(t)$  if  $t \in B$ . Finally, according to Lemma 4, we can expand  $a^*$  to an incentive feasible mechanism  $\mu^*$  over  $T$ . We claim that  $\mu^*$  is an interim equitable mechanism.

Suppose, by way of contradiction, that  $\mu^*$  is not an interim equitable mechanism. Given that, by construction,  $\lambda_i^{\mu^*}(t_i) = \lambda_j^{\mu^*}(t_j)$  for all  $i, j \in N$  and all  $t \in \widehat{T}$ , it must necessarily be the case that  $\mu^*$  is not interim incentive efficient. Let  $\mu^\circ$  be an incentive feasible mechanism that interim Pareto dominates  $\mu^*$ . Hence, there must exist  $B^\circ \in \mathcal{B}$ ,  $t^\circ \in B$  and  $i \in N$  for which  $u_i(\mu^\circ|t_i^\circ) > u_i(\mu^*|t_i^\circ)$  and  $u_j(\mu^\circ|t_j) \geq u_j(\mu^*|t_j)$  for all  $j \in N$  for all  $t \in T$ .

We now construct another allocation  $\mu^A$  that interim improves over  $\widehat{\mu}$  as follows:  $\mu_i^A(t^\circ) = \mu_i^\circ(t^\circ) - (n-1)\varepsilon d$ ,  $\mu_j^A(t^\circ) = \mu_j^\circ(t^\circ) + \varepsilon d$  for all  $j \neq i$ , and  $\mu_k^A(t) = \mu_k^\circ(t)$  for all  $k \in N$  if  $t \neq t^\circ$ . If  $\varepsilon > 0$  and small enough,  $u_i(\mu^A|t_i^\circ) > u_i(\mu^*|t_i^\circ)$  and  $u_j(\mu^A|t_j^\circ) > u_j(\mu^*|t_j^\circ)$  for all  $j \neq i$ . We now use any agent who is better off under the new allocation to improve still other types of other agents by further modifying the mechanism in the same manner. By Lemma 3 this process reaches any type vector in  $B^\circ$ , leading to a final mechanism that is incentive feasible and corresponds to a larger  $\lambda$  than  $\lambda_B$  in contradiction to the maximality of  $\lambda_B$ .

(b) By the above construction we see that the  $\lambda_i^\mu(t_i) = \lambda_j^\mu(t_j)$  for all  $t \in B$  and  $i, j \in N$  whenever  $B$  is a minimal common knowledge event.

(c) Since any interim equitable mechanisms  $\mu$  and  $\eta$  are interim Pareto efficient, it must be the case that  $\lambda_B^\mu = \lambda_B^\eta$  for any minimal common knowledge event  $B$ .

(d) Let  $\mu$  and  $\eta$  be two distinct interim equitable mechanisms. Consider  $\psi = \frac{\mu + \eta}{2}$ . Allocation  $\psi$  is feasible and, given the strict concavity of the utility functions, it interim Pareto dominates  $\mu$  and  $\eta$ . Indeed,  $U_i(\psi|t_i) \geq U_i(\mu|t_i)$  since

$$\begin{aligned} \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) u_i(\psi_i(t), t) &= \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) u_i \left( \left( \frac{\mu + \eta}{2} \right)_i(t), t \right) \geq \\ \frac{1}{2} \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) [u_i(\mu_i(t), t) + u_i(\eta_i(t), t)] &= \sum_{t_{-i} \in T_{-i}} \pi(t_{-i}|t_i) u_i(\mu_i(t), t). \end{aligned}$$

Moreover, since  $\mu$  and  $\eta$  are distinct, there is at least one strict inequality. ■

Proposition 5 shows that efficiency and egalitarianism are compatible in environments with asymmetrically informed agents when information is non-exclusive.

We note that the complete information setting is a special case of NEI, where  $\widehat{T}$  is “diagonal”, each profile of types in  $\widehat{T}$  is uniquely determined by any of its components. In this case any type profile  $t$  in  $\widehat{T}$  is a common knowledge event. For any such  $t$ ,  $\lambda_i^\mu(t_i) = \lambda_j^\mu(t_j)$  over all agents  $i$  and  $j$  when a mechanism  $\mu$  is interim equitable. Note however that the  $\lambda$ 's may well differ over distinct  $t$ 's.

More generally, when the partition of  $\widehat{T}$  into minimal common knowledge events is not trivial, the  $\lambda$ 's associated to all agents and all types must coincide over any minimal common knowledge event  $B$ , that is,  $\lambda_i^\mu(t_i) = \lambda_j^\mu(t_j)$  for all  $t \in B$ ,  $i, j \in N$ , and interim equitable  $\mu$ . However, the gains that agents obtain can vary over different common knowledge events. That is, it can perfectly be the case that  $\lambda_i^\mu(t_i) \neq \lambda_j^\mu(t_j)$  for an interim equitable mechanism  $\mu$  if the profile  $(t_i, t_j, t_{-ij})$  does not belong to any common knowledge event for any  $t_{-ij}$ .

The NEI condition enabled us to ignore to a large degree the incentive constraints. Another approach taken in the literature to circumvent the incentive constraints is to assume information is ex-post verifiable. That is the contracts are agreed upon in the interim stage, but when they are executed (in the ex-post stage) information is verifiable, agents' types are known. Proposition 5 with the assumption of verifiable information replacing the NEI assumption holds as well and furthermore is valid for the case of two agents with asymmetric information (recall that an NEI environment with two agents must be a complete information environment).

While the existence result might not be surprising given previous findings in complete information environments, it contrasts with recent findings regarding other fairness notions in economies with asymmetric information such as envy-freeness.

The set of fair (envy-free and Pareto efficient) allocations is always non-empty for allocation problems in classical exchange economies under complete

information, as it contains for instance the competitive equilibrium associated with an equal split of the total endowment. de Clippel (2008) proposes a natural extension of the concept of envy-freeness to exchange economies under asymmetric information. However, he shows that interim envy-freeness may be incompatible with interim Pareto efficiency.<sup>6</sup> The non-existence example in de Clippel (2008) involved variable aggregate endowments and verifiable information, however it can be modified to show that in our set-up with constant aggregate endowments and NEI, there are also situations where no envy-free and interim efficient allocations exist

**Example 6** *There are four agents and only one good (money). Each agent  $i = 1, 2$  has two possible types  $t_1^i, t_2^i$ , agent 3 has one type  $t^3$  and agent 4 has four types  $t_k^4, k = 1, \dots, 4$ , and  $\widehat{T} = \{(t_1^1, t_1^2, t^3, t_1^4), (t_1^1, t_2^2, t^3, t_2^4), (t_2^1, t_1^2, t^3, t_3^4), (t_2^1, t_2^2, t^3, t_4^4)\}$  with  $\pi(t) = 0.25$  for all  $t \in \widehat{T}$ . This environment satisfies NEI and agent 4 has complete information. The aggregate initial endowment and preferences at each state are described in the following table.*

State	$e(\cdot)$	$u_1(x, \cdot)$	$u_2(x, \cdot)$	$u_3(x)$	$u_4(x, \cdot)$
$(t_1^1, t_1^2, t_1^3, t_1^4)$	600	$x^{0.5}$	$x$	$x$	$x$
$(t_1^1, t_2^2, t_1^3, t_2^4)$	600	$2x^{0.5}$	$x$	$x$	$x$
$(t_2^1, t_1^2, t_1^3, t_3^4)$	600	$x^{0.5}$	$x$	$x$	$x$
$(t_2^1, t_2^2, t_1^3, t_4^4)$	600	$2x^{0.5}$	$x$	$x$	$x$

Similar to de Clippel (2008) it can be shown that interim envy-freeness entails that each agent at each state should receive 150. This allocation is interim dominated by the allocation where at state  $(t_1^1, t_1^2, t^3, t_1^4)$  agent 1 receives 140 and agent 3 receives 160 and at state  $(t_1^1, t_2^2, t^3, t_2^4)$  agent 1 receives 160 and agent 3 receives 140. Finally, notice that the interim equitable mechanism for this example (with one minimal common knowledge event) yield  $\lambda = \frac{2000}{13}$ . One of the interim equitable mechanisms has agents 2 and 4 receiving  $\frac{2000}{13}$  in all states, agent 1 receiving  $\frac{720}{13}$  in the first and third states and  $\frac{2880}{13}$  in the second and fourth states; finally, agent 3 receives the remaining aggregate endowment:  $\frac{3080}{13}$  in odd states and  $\frac{920}{13}$  in even states.

Existence of interim equitable mechanisms is not longer guaranteed in environments without the NEI assumption. Next sections provides an extended example in an economic environment where, for certain parameter values, an interim equitable mechanism does not exist.

<sup>6</sup>Production on the other hand has a similar effect under complete information. P&S (1974) show indeed that efficiency and envy-freeness may be incompatible in classical economies with production. This motivated P&S (1978) to propose egalitarian equivalence as a new ordinal notion of equity that would be compatible with efficiency in classical economic environments (with or without production). Our natural extension of egalitarian equivalence in exchange economies under asymmetric information achieves the same objective.

## 5 An economic example

This section provides a economic example, similar to the one provided by Myerson (1985), where first we characterize the interim equitable mechanisms under the NEI assumption and then show that the set of interim equitable mechanisms may be empty.

There are two commodities,  $q$  and  $m$ , where  $m$  can be viewed as money and  $q$  as consumption, the aggregate amounts of these commodities are given by  $Q$  and  $M$ . There are three agents, 1, 2 and 3; agent 1 has one type,  $t^1$  and agents 2 and 3 have two types  $(t_L^2, t_H^2)$  and  $(t_L^3, t_H^3)$ ,  $\pi(t^1, t_L^2, t_L^3) = p$ ,  $\pi(t^1, t_H^2, t_H^3) = 1-p$ . That is, agents 2 and 3 are completely informed while agent 1 is uncertain about the state of the world. This environment satisfies the NEI assumption. Let  $q_i$  denote the amount of consumption and  $m_i$  the amount of money assigned to agent  $i$ ,  $i = 1, 2, 3$ . The preferences of the agents are given by:

State	$e(\cdot)$	$u_1(\cdot, \cdot)$	$u_2(\cdot, \cdot)$	$u_3(\cdot, \cdot)$
$(t^1, t_L^2, t_L^3)$	$(Q, M)$	$m_1 + 2\sqrt{q_1}$	$m_2 + v_L q_2$	$m_3$
$(t^1, t_H^2, t_H^3)$	$(Q, M)$	$m_1 + 2\sqrt{q_1}$	$m_2 + v_H q_2$	$m_3$

where  $0 < v_L < v_H$ . A feasible allocation is a vector  $(q_{it}, m_{it})_{i=1,2,3;t=H,L}$  that satisfies  $q_{1t} + q_{2t} + q_{3t} \leq Q$  and  $m_{1t} + m_{2t} + m_{3t} \leq M$  for  $t = H, L$ .

Ex-post Pareto efficiency requires

$$q_{1t} = \frac{1}{v_t^2}, q_{2t} = Q - \frac{1}{v_t^2} \text{ and } q_{3t} = 0 \text{ for } t = H, L. \quad (1)$$

And, in this simple example, (1) are also the necessary and sufficient requirements for both ex-post and interim Pareto efficiency. Any sharing of the money  $(m_{it})_{i=1,2,3;t=H,L}$  that satisfies  $m_{1t} + m_{2t} + m_{3t} = M$ , for  $t = H, L$  is interim efficient.

Therefore, using (1), interim equitable mechanisms are characterized by  $m_{1t} + m_{2t} + m_{3t} = M$ , for  $t = H, L$ , and the existence of  $\lambda$  such that the following five equations hold:

$$u_1 = p \left[ m_{1L} + \frac{2}{v_L} \right] + (1-p) \left[ m_{1H} + \frac{2}{v_H} \right] = \lambda M + 2\sqrt{\lambda Q}, \quad (2)$$

$$u_{2L} = m_{2L} + v_L Q - \frac{1}{v_L} = \lambda M + v_L \lambda Q, \quad (3)$$

$$u_{2H} = m_{2H} + v_H Q - \frac{1}{v_H} = \lambda M + v_H \lambda Q. \quad (4)$$

$$u_{3L} = m_{3L} = \lambda M, \quad (5)$$

$$u_{3H} = m_{3H} = \lambda M. \quad (6)$$

The sum of (2),  $p$ -times (3) and (5) and  $(1-p)$ -times (4) and (6) gives (given  $m_{1t} + m_{2t} + m_{3t} = M$ , for  $t = H, L$ ):

$$M + p \left[ \frac{1}{v_L} + v_L Q \right] + (1-p) \left[ \frac{1}{v_H} + v_H Q \right] = 2\sqrt{Q}\sqrt{\lambda} + (3M + [pv_L + (1-p)v_H] Q) \lambda, \quad (7)$$

which is a second-degree equation in  $\sqrt{\lambda}$ . There is a unique positive value for  $\lambda$  that solves this equation. Then, in this example, there is also a unique interim equitable mechanism where the allocation of money is determined from equations (3) to (5) using the  $\lambda$  that solves (7).

In the environment specified so far, we were able, thanks to NEI, to ignore the incentive constraints. This is not longer possible if NEI is not satisfied. We now modify the environment by assuming that agent 3 is also uninformed. Hence, agent 2 has “real” private information and may decide not to truthfully report it, if it is in his interest to do so.

So the environment can now be described as follows:

State	$e(\cdot)$	$u_1(\cdot, \cdot)$	$u_2(\cdot, \cdot)$	$u_3(\cdot, \cdot)$
$(t^1, t_L^2, t^3)$	$(Q, M)$	$m_1 + 2\sqrt{q_1}$	$m_2 + v_L q_2$	$m_3$
$(t_1^1, t_H^2, t^3)$	$(Q, M)$	$m_1 + 2\sqrt{q_1}$	$m_2 + v_H q_2$	$m_3$

We start by determining the set of interim incentive efficient mechanisms. The two incentive constraints are:

$$m_{2H} + v_H q_{2H} \geq m_{2L} + v_H q_{2L}, \quad (8)$$

$$m_{2L} + v_L q_{2L} \geq m_{2H} + v_L q_{2H}. \quad (9)$$

In Appendix 1 we show that the set of interim incentive efficient mechanism is the union of the three following regions:

Region 1:

$$q_{1H} = \frac{1}{v_H^2}, q_{1L} = \frac{1}{v_L^2}, \text{ and}$$

$$\text{any allocation of money satisfying } m_{2L} - m_{2H} \in \left[ \frac{1}{v_L} - \frac{v_L}{v_H^2}, \frac{v_H}{v_L^2} - \frac{1}{v_H} \right].$$

Region 2:

$$q_{1H} = \frac{1}{v_H^2}, \quad \text{any } q_{1L} \geq \frac{1}{v_L^2} \text{ if } v_L \leq (1-p)v_H$$

$$\text{any } q_{1L} \in \left[ \frac{1}{v_L^2}, \frac{p^2}{(v_L - (1-p)v_H)^2} \right] \text{ if } v_L > (1-p)v_H, \quad \text{and}$$

$$\text{any allocation of money satisfying } m_{2L} - m_{2H} = v_H \left( q_{1L} - \frac{1}{v_H^2} \right).$$

Region 3:

$$\text{any } q_{1H} \in \left[ \frac{(1-p)^2}{(v_H - pv_L)^2}, \frac{1}{v_H^2} \right], q_{1L} = \frac{1}{v_L^2}, \text{ and}$$



any allocation of money satisfying  $m_{2L} - m_{2H} = v_L \left( \frac{1}{v_L^2} - q_{1H} \right)$ .

We can now show that for the economy where  $v_L = 1$ ,  $v_H = 2$ ,  $p = 0.75$ ,  $Q = 12$  and  $M = 20$ , there does not exist an interim equitable mechanism, taking  $d$  as the vector of endowments:  $d = (M, Q)$ . Egalitarian equivalence requires  $m_{1t} + m_{2t} + m_{3t} = 20$  for  $t = H, L$  and:

$$0.25(m_{1H} + 2\sqrt{q_{1H}}) + 0.75(m_{1L} + 2\sqrt{q_{1L}}) = 20\lambda + 2\sqrt{12\lambda} \quad (10)$$

$$m_{2H} + 2(12 - q_{1H}) = 44\lambda \quad (11)$$

$$m_{2L} + (12 - q_{1L}) = 32\lambda \quad (12)$$

$$0.25m_{3H} + 0.75m_{3L} = 20\lambda \quad (13)$$

We first note that the sum of (10), 0.25 times (11), 0.75 times (12) and (13) gives:

$$0.5\sqrt{q_{1H}} + 1.5\sqrt{q_{1L}} - 0.5q_{1H} - 0.75q_{1L} + 34 = 75\lambda + 4\sqrt{3}\sqrt{\lambda}. \quad (14)$$

We now proceed by examining the possible regions. In Region 1,  $q_{1H} = 0.25$  and  $q_{1L} = 1$ . Then, equation (14) yields  $\lambda = 0.406$  and subtracting (11) from (12) gives  $m_{2L} - m_{2H} = 7.878$ , which violates the upper-bound  $\frac{v_H}{v_L} - \frac{1}{v_H} = 1.5$ . In Region 2,  $q_{1H} = 0.25$  and  $m_{2L} - m_{2H} = 2q_{1L} - 0.5$ . Subtracting (11) from (12) gives  $q_{1L} = 12 - 12\lambda$ . Then, substituting this into (14) yields  $\lambda = .386$  which implies  $q_{1L} = 7.363$ , violating the upper-bound  $\frac{p^2}{(v_L - (1-p)v_H)^2} = 2.25$ . Finally, in Region 3  $q_{1L} = 1$  and  $m_{2L} - m_{2H} = 1 - q_{1H}$ . From (11) and (12) we obtain  $q_{1H} = 12 - 12\lambda$  and, substituting into (14) yields  $\lambda = .375$ , hence  $q_{1H} = 7.4998$ , which again violates the upper-bound  $\frac{1}{v_H} = 0.5$ .

There are of course several other instances where interim equitable mechanisms exist. In Appendix 2 we show that, if  $v_L \leq (1-p)v_H$  and  $Q$  large enough, there always exists an interim equitable mechanism; it lies in Region 2.

Hence moving from fairness to egalitarian equivalence while consistent with efficiency in NEI environments may clash with interim efficiency once incentive constraints are relevant. Equity considerations have faced the same problem in economic environments: equality may be incompatible with efficiency (see Pazner and Schmeidler, 1974, and Maniquet, 1999).

## 6 Implementation of the solution

In a complete information environment, Crawford (1979) and Demange (1984) have provided a non-cooperative foundation for the Pareto efficient egalitarian equivalent allocation rule by means of simple games that implement the set of

such allocations. We show here that a simple adaptation of these games to environments with asymmetric information generates interim equitable mechanisms as perfect Bayesian equilibria, that is, the set of interim equitable mechanisms can be weakly implemented in perfect Bayesian equilibrium. For any interim equitable mechanism, we construct a perfect Bayesian equilibrium of the game yielding it. We consider the following game, played at the interim stage:

Game  $\mathcal{G}$ :

At the first stage each agent  $i \in N$  announces a number  $\lambda^i$  in  $R_+$ . We let  $\lambda^* = \text{Max}_{i \in N}(\lambda^i)$  and call the agent announcing it  $\alpha$  (in case of a tie,  $\alpha$  is randomly chosen among the set of agents announcing the largest number).

At the second stage, the identity of  $\alpha$  and the winning bid  $\lambda^*$  are revealed and  $\alpha$  offers an incentive feasible mechanism  $\mu \in R^{L \times N \times T}$ .

The other agents reply sequentially with a YES or a NO. If they all answer YES the game ends and  $\mu$  is carried out.

The first NO that is encountered ends the game. The proposer gives each agent  $\lambda^*d$  and thus the final outcome is  $\lambda^*d$  for all agents other than  $\alpha$  and  $e - (N - 1)\lambda^*d$  for agent  $\alpha$ .

**Proposition 7** *Game  $\mathcal{G}$  weakly implements the set of interim equitable mechanisms.*

**Proof.** Let an interim equitable mechanism  $\mu$  be given. We now proceed to construct a perfect Bayesian equilibrium of  $\mathcal{G}$  yielding it.

Consider the following  $N$  – tuple of strategies and beliefs:

Stage 1: For any  $t \in \hat{T}$ , each agent  $i \in N$  announces  $\lambda_i^\mu(t_i)$ .

Stage 2: Agents' beliefs do not change, independently on the information revealed after Stage 1. Any type of the agent chosen as the proposer announces an incentive feasible mechanism for which, given their initial beliefs, all agents other than the proposer are indifferent between accepting it or rejecting it, that is, receiving  $d$  multiplied by the bid the proposer made. If it so happens that Stage 1 ended with  $\lambda^* = \lambda_B^\mu$ , where  $B$  is the minimal common knowledge event including  $t_\alpha$ , any proposer  $\alpha$  announces  $\mu$ .

Stage 3: Agents' beliefs do not change, independently on the information revealed after Stage 2. Any type of any agent announces YES for any proposal made by  $\alpha$  that is weakly preferred to the rejection outcome given his beliefs and announces NO for any other proposal.

Using this tuple of strategies clearly generates the interim equitable mechanism  $\mu$ . To see they form a perfect Bayesian equilibrium we start from Stage 3 where the strategies constitute a best response by construction. At Stage 2, the fact that  $\mu$  was interim equitable means it is a best response along the equilibrium path and all choices off the equilibrium path enjoy the best response property by construction. Also, having the offer accepted is certainly weakly preferred to having it rejected. At Stage 1, since  $\lambda_B^\mu$  supports an interim equitable mechanism when a vector of types in  $B$  has been realized, each agent is indifferent between winning or not. Hence, no agent can gain by lowering his bid since this will not change the proposed mechanism ( $\mu$ ) at Stage 2. Also, it

does not pay any agent to announce a larger bid to become the proposer with certainty, since his expected utility will decrease.

As regards beliefs, we stress that we take agents' beliefs to be constant on and off the equilibrium path. Having constant beliefs along the equilibrium path is consistent with Bayesian updating, since no further information is revealed along it. Outside the equilibrium path there is no constraint on the beliefs the agents might hold and we, in particular, assume they are constant. ■

We note that the implementation result does not require the NEI assumption. In any environment where interim equitable mechanisms exist, the game reaches the mechanisms as perfect Bayesian equilibria.

## 7 Discussion and further proposals

We start by discussing some additional properties of the interim equitable solution that associates to each economy the set of mechanisms that are interim incentive efficient and interim egalitarian equivalent. First, the proposal is invariant to affine transformations of the interim utilities, i.e. changing the utility function  $u_i$  of any agent  $i \in N$  by multiplying it at every  $t$  by a strictly positive coefficient, and/or adding a real number at every  $t$ , does not affect the solution.<sup>7</sup> Second, the interim equitable solution satisfies Myerson's (1984) probability invariance axiom, since it depends on the probabilities only through the computation of interim utilities. One could even have considered a more general framework with the agents' ordinal interim preferences as exogenous variables, instead of deriving those from the expected utility criterion applied to ex-post utilities. Indeed, both interim incentive efficiency and interim egalitarian equivalence depend only on those interim preferences, and the interim equitable solution is then ordinally invariant in this more general framework.

Third, the interim equitable solution is anonymous, meaning that renaming the agents, or even their types, will not change their payoffs. Fourth, the solution is also monotonic, meaning that increasing the total endowment  $e$  cannot make any agent of any type worse off. Fifth, we can also offer a weak comparison of the level of interim satisfaction achieved at mechanisms in the interim equitable solution and the level of satisfaction achieved for egalitarian equivalent allocations in the ex-post economies. Let  $\lambda^*(t)$  be the level reached at any Pareto efficient and egalitarian equivalent allocation, in the ex-post economy obtained should  $t$  realize. Let  $\lambda^*(E)$  be the level reached by any mechanism in the solution on the minimal common knowledge event  $E$ . If  $\pi$  satisfies the NEI condition, then  $\lambda^*(E) \geq \min_{t \in \hat{T} \cap E} \lambda^*(t)$  for each minimal common knowledge event  $E$ . Notice that the inequality is most often strict, because of the possibility of mutually beneficial insurance. On the other hand, the inequality does not extend to economies that do not satisfy NEI, because the incentive constraints can be so severe that it is impossible to guarantee even the minimum of the ex-post levels.

---

<sup>7</sup>In fact, our proposal is invariant to more general affine transformations: one can allow an affine transformation for each type  $t_i$  of agent  $i$ .

As we have shown through an economic example, interim equitable solutions may not exist in environments with asymmetric information when the NEI hypothesis does not hold. When there is tension between efficiency and fairness, a common remedy is to look for allocations (or mechanisms) that minimize the largest deviation from equal expected gains. To avoid multiplicity, a natural lexicographic refinement is often applied. We can follow the same path and, for any mechanism  $\mu$ , define the vector  $\Delta\gamma^\mu$  as a vector whose components are  $|\gamma_j^\mu(t_i) - \gamma_i^\mu(t_i)|$  for all  $i, j \in N$ ,  $i \neq j$ , and all  $t_i \in T_i$ . Let  $\alpha$  be the function that associates to each vector of real numbers the vector obtained by ordering its components decreasingly. Then, an interim incentive efficient mechanism  $\mu$  is said to be weakly interim equitable if it minimizes  $\alpha(\Delta\gamma^\mu)$  according to the lexicographic ordering over the set of interim incentive efficient mechanisms. The set of weakly interim equitable mechanisms is always non-empty.

## References

## References

- [1] Crawford, V., 1979. A Procedure for Generating Pareto-Efficient Egalitarian-Equivalent Allocations, *Econometrica* 47, 49-60.
- [2] de Clippel, G., 2008. Equity, Envy and Efficiency under Asymmetric Information. *Economics Letters* 99, 265-267.
- [3] Demange, G., 1984. Implementing Efficient Egalitarian Equivalent Allocations. *Econometrica* 52, 1167-1177.
- [4] Gibbard, A., 1973. Manipulation for Voting Schemes. *Econometrica* 41, 587-601.
- [5] Green, J. L., and J.-J. Laffont, 1979. Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods. *Econometrica* 45, 427-438.
- [6] Harsanyi, J.C., 1959. A Bargaining Model for the Cooperative n-Person Game. In *Contributions to the Theory of Games IV*, ed. by A.W. Tucker and R.D. Luce. Princeton: Princeton University Press, 325-355.
- [7] Holmström, B., and R. B. Myerson, 1983. Efficient and Durable Decision Rules with Incomplete Information. *Econometrica* 51, 1799-1819.
- [8] Kalai, E., 1977. Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons. *Econometrica* 45, 1623-1630.
- [9] Maniquet, F., 1999. A Strong Incompatibility between Efficiency and Equity in non-convex Economies. *Journal of Mathematical Economics* 32, 467-474.

- [10] Myerson, R. B., 1979. Incentive Compatibility and the Bargaining Problem. *Econometrica* 47, 61-73.
- [11] Myerson, R. B., 1984. Two Bargaining Problems with Incomplete Information. *Econometrica* 52, 461-88.
- [12] Nash, J., 1950. The Bargaining Problem. *Econometrica* 18 (2), 155–162.
- [13] Nehring, K., 2004. The Veil of Public Ignorance. *Journal of Economic Theory* 119, 247-270.
- [14] Pazner, E., and D. Schmeidler, 1974. A Difficulty in the Concept of Fairness. *Review of Economics Studies* 41, 441-443.
- [15] Pazner, E., and D. Schmeidler, 1978. Egalitarian Equivalent Allocations: A New Concept of Economic Equity. *Quarterly Journal of Economics* 92, 671-686.
- [16] Postlewaite, A., and D. Schmeidler, 1986. Implementation in Differential Information Economies. *Journal of Economic Theory* 39, 14-33.
- [17] Shapley, L.S., 1953. A Value for  $n$ -Person Games. In *Contributions to the Theory of Games II* (Annals of Mathematics Studies), H.W. Kuhn and A.W. Tucker (eds.), Princeton University Press, 307-317.
- [18] Varian, H., 1974. Equity, Envy and Efficiency. *Journal of Economic Theory* 9, 63-91.
- [19] Widener, F., 1992. The Generalized Nash bargaining Solution and Incentive Compatible Mechanisms. *International Journal of Game Theory* 21, 109-129.

## 8 Appendix 1

We characterize the set of interim incentive efficient (IIE) mechanism of the Example through some claims:

**Claim 8** (a) Any IIE mechanism involves  $q_{1H} \leq \frac{1}{v_H^2}$ . (b) Moreover,  $q_{1H} = \frac{1}{v_H^2}$  if (9) is not binding.

Indeed, from any mechanism, change  $q_{2H}$  by a small  $\delta$  and simultaneously change  $m_{2H}$  by an amount  $-\delta v_H$ . The utility obtained by both types of agent 2 and constraint (8) do not change. Constraint (9) is relaxed if  $\delta > 0$ . Finally, agent 1's utility level increases with the change when  $\delta > 0$  and  $q_{1H} > \frac{1}{v_H^2}$  or when  $\delta < 0$  and  $q_{1H} < \frac{1}{v_H^2}$ . Therefore,  $q_{1H} > \frac{1}{v_H^2}$  cannot be part of an IIE mechanism. Also,  $q_{1H} < \frac{1}{v_H^2}$  cannot be part of an IIE mechanism if (8) is not binding.

**Claim 9** (a) Any IIE mechanism involves  $q_{1L} \geq \frac{1}{v_L^2}$ . (b) Moreover,  $q_{1L} = \frac{1}{v_L^2}$  if (8) is not binding.

Similarly as before, from any allocation, change  $q_{2L}$  by a small  $\delta$  and simultaneously change  $m_{2H}$  by an amount  $-\delta v_L$ . Agent 2's utility and constraint (8) do not change. Constraint (8) is relaxed if  $\delta < 0$ . Agent 1's utility increases when  $\delta < 0$  and  $q_{1L} < \frac{1}{v_L^2}$  or when  $\delta > 0$  and  $q_{1L} > \frac{1}{v_L^2}$ .

From Claims 8 and 9,  $q_{1L} \geq \frac{1}{v_L^2} > \frac{1}{v_H^2} \geq q_{1H}$ . On the other hand, if (8) and (9) would both hold with equality, then  $q_{1L} = q_{1H}$ . Therefore:

**Claim 10** Both incentive constraints (8) and (9) can not bind simultaneously.

**Claim 11** Any IIE mechanism involves  $q_{3L} = q_{3H} = 0$ .

The proof is immediate given that agent 3 derives no utility from  $q$ .

We now analyze the three possible regions where IIE allocations can lie: no incentive constraint binding, or one of them is binding. To the equations identifying the IIE allocations below, we always have to add the ex-post efficient requirement  $q_{1t} + q_{2t} = Q$  and  $m_{1t} + m_{2t} + m_{3t} = M$  for  $t = H, L$ .

**Claim 12** In Region 1, where no incentive constraint is binding, the IIE allocations are characterized by:

$$q_{1H} = \frac{1}{v_H^2}, q_{1L} = \frac{1}{v_L^2}, \text{ and}$$

$$\text{any allocation of money satisfying } m_{2L} - m_{2H} \in \left[ \frac{1}{v_L} - \frac{v_L}{v_H^2}, \frac{v_H}{v_L^2} - \frac{1}{v_H} \right].$$

The allocations must be ex-post Pareto efficient if incentive constraints are not relevant, while the condition on  $m_{1H} - m_{1L}$  rewrites the constraints (8) and (9) for those values of  $q_{1H}$  and  $q_{1L}$ .

**Claim 13** In Region 2, where constraint (8) is binding, the IIE allocations are characterized by:

(a) If  $v_L \leq (1-p)v_H$

$$q_{1H} = \frac{1}{v_H^2}, \text{ any } q_{1L} \geq \frac{1}{v_L^2},$$

$$\text{any allocation of money that involves } m_{2L} - m_{2H} = v_H \left( q_{1L} - \frac{1}{v_H^2} \right).$$

(b) If  $v_L > (1-p)v_H$ , same conditions as in (a) except that  $q_{1L} \in \left[ \frac{1}{v_L^2}, \frac{p^2}{(v_L - (1-p)v_H)^2} \right]$ .

To show that these conditions characterize the IIE allocations in this region, we note that  $q_{1H} = \frac{1}{v_H^2}$  by Claim 8 and  $m_{1H} = m_{1L} + v_H \left( q_{1L} - \frac{1}{v_H^2} \right)$  given

that (8) is binding. Following some calculations, agents' utility as a function of  $q_{1L}$  and  $m_{1L}$  is:

$$u_1 = m_{1L} + 2p\sqrt{q_{1L}} + (1-p)v_H q_{1L} + (1-p)\frac{1}{v_H}. \quad (15)$$

$$u_{2H} = M - m_{1H} - m_{3H} + v_H(Q - q_{1L}) \quad (16)$$

$$u_{2L} = M - m_{1L} - m_{3L} + v_L(Q - q_{1L}) \quad (17)$$

$$u_3 = pm_{3L} + (1-p)m_{3H} \quad (18)$$

Any interim Pareto improvement of any allocation by increasing  $q_{1L}$  by a small  $\delta > 0$  requires a reduction of  $m_{1L}$  of, at least,  $\delta v_H$  to compensate agent 2 of type  $H$ . Such a change would improve agent 1's utility if:

$$-\delta v_H + \frac{p}{\sqrt{q_{1L}}}\delta + (1-p)v_H\delta = p\left(\frac{1}{\sqrt{q_{1L}}} - v_H\right) > 0,$$

which is never the case given that  $q_{1L}$  must be larger or equal than  $\frac{1}{v_L^2}$ . Similarly, a reduction of  $\delta$  in  $q_{1L}$  can go together with an increase of  $\delta v_L$  so that agent 2 of type  $L$  does not lose utility. Agent 1's utility improves if:

$$\delta v_L - \frac{p}{\sqrt{q_{1L}}}\delta - (1-p)v_H\delta > 0, \text{ i.e., } \sqrt{q_{1L}}\left(\frac{v_L - (1-p)v_H}{p}\right) > 1.$$

Therefore, if  $v_L \leq (1-p)v_H$  there is no incentive compatible allocation that dominates any  $(q_{1L}, m_{1L})$  that satisfies the two initial constraints. If  $v_L > (1-p)v_H$ , any allocation with  $q_{1L} > \frac{p^2}{(v_L - (1-p)v_H)^2}$  can be interim-incentive improved.

**Claim 14** *In Region 3, where constraint (9) is binding, the IIE allocations are characterized by:*

$$\text{any } q_{1H} \in \left[ \frac{(1-p)^2}{(v_H - pv_L)^2}, \frac{1}{v_H^2} \right], \quad q_{1L} = \frac{1}{v_L^2}, \quad \text{and}$$

$$\text{any allocation of money that involves } m_{2L} - m_{2H} = v_L \left( \frac{1}{v_L^2} - q_{1H} \right).$$

In this region, we note that  $q_{1L} = \frac{1}{v_L^2}$  by Claim 8 while  $m_{1H} = m_{1L} + v_L \left( \frac{1}{v_L^2} - q_{1H} \right)$ . Agents' utilities, as a function of  $q_{1H}$  and  $m_{1H}$  are:

$$u_1 = m_{1H} + 2(1-p)\sqrt{q_{1H}} + pv_L q_{1H} + p\frac{1}{v_L}.$$

$$u_{2H} = M - m_{1H} - m_{3H} + v_H(Q - q_{1H})$$

$$u_{2L} = M - m_{1L} - m_{1L} + v_L(Q - q_{1H})$$

$$u_3 = pm_{3L} + (1-p)m_{3H}$$

Increasing  $q_{1H}$  by a small  $\delta > 0$  requires a reduction of  $m_{1H}$  of, at least,  $\delta v_H$ . This would improve agent 1's utility if:

$$-\delta v_H + \frac{(1-p)}{\sqrt{q_{1H}}}\delta + pv_L\delta > 0, \text{ i.e., } \sqrt{q_{1H}} \left( \frac{v_H - pv_L}{1-p} \right) < 1.$$

Therefore, any allocation with  $q_{1H} < \frac{(1-p)^2}{(v_H - pv_L)^2}$  can be interim-incentive improved. On the other hand, a reduction of  $\delta$  in  $q_{1H}$  together with an increase of  $\delta v_L$  improves agent 1's utility if:

$$\delta v_L - \frac{(1-p)}{\sqrt{q_{1H}}}\delta - pv_L\delta = (1-p) \left( v_L - \frac{1}{\sqrt{q_{1H}}} \right) > 0$$

which is never the case since  $q_{1H} \leq \frac{1}{v_H^2}$ .

## 9 Appendix 2

We prove that, if  $v_L \leq (1-p)v_H$  and  $Q$  is large enough, there always exists an interim equitable mechanism; it lies in Region 2. Indeed, taking into account the characteristics of the IIE allocations, a mechanism is interim equitable if there exists  $\lambda \in (0, 1)$  such that:

$$pm_{1L} + 2p\sqrt{q_{1L}} + (1-p)m_{1H} + 2(1-p)\sqrt{q_{1H}} = \lambda M + 2\sqrt{\lambda Q}. \quad (19)$$

$$m_{2H} + v_H(Q - q_{1H}) = \lambda(M + v_H Q) \quad (20)$$

$$m_{2L} + v_L(Q - q_{1L}) = \lambda(M + v_L Q) \quad (21)$$

$$pm_{3L} + (1-p)m_{3H} = \lambda M \quad (22)$$

Subtracting (20) from (21), and recalling that in Region 2,  $q_{1H} = \frac{1}{v_H^2}$  and  $m_{2L} - m_{2H} = v_H \left( q_{1L} - \frac{1}{v_H^2} \right)$ , we obtain that  $\lambda = 1 - \frac{q_{1L}}{Q}$ . Also, adding up equations (19),  $(1-p)$  times (20),  $p$  times (21), and (22), we get:

$$M + 2p\sqrt{q_{1L}} + (1-p)\frac{1}{v_H} + pv_L(Q - q_{1L}) + (1-p)v_H Q = (3M + pv_L Q + (1-p)v_H Q)\lambda + 2\sqrt{\lambda Q}$$

Substituting  $\lambda$  in the previous expression, we obtain:

$$2p\sqrt{q_{1L}} + (1-p)\frac{1}{v_H} = 2M - q_{1L} \left( \frac{3M}{Q} + (1-p)v_H \right) + 2\sqrt{(Q - q_{1L})}$$

i.e.,

$$f(q_{1L}) \equiv 2M - q_{1L} \left( \frac{3M}{Q} + (1-p)v_H \right) + 2\sqrt{(Q - q_{1L})} - 2p\sqrt{q_{1L}} - (1-p)\frac{1}{v_H} = 0.$$



We are considering large  $M$ s, hence  $f(q_{1L} = 0) > 0$ . Also,

$$f(q_{1L} = \frac{1}{v_L^2}) \equiv 2M - \frac{1}{v_L^2} \left( \frac{3M}{Q} + (1-p)v_H \right) + 2\sqrt{\left(Q - \frac{1}{v_L^2}\right)} - 2p\sqrt{\frac{1}{v_L^2}} - (1-p)\frac{1}{v_H} > 0$$

if  $Q$  is large; finally

$$f(q_{1L} = Q) = -M - (1-p)v_H Q - 2p\sqrt{Q} - (1-p)\frac{1}{v_H} < 0.$$

Therefore, there exists  $q_{1L} \in \left(\frac{1}{v_L^2}, Q\right)$  for which  $f(q_{1L}) = 0$ . Together with  $m_{1L} = \frac{q_{1L}}{Q}M$ , that  $q_{1L}$  is part of an interim equitable mechanism with the corresponding  $\lambda = 1 - \frac{q_{1L}}{Q}$ .