Motives and implementation with rights structures

Foivos Savva*

Friday 4th June, 2021

Abstract

We study implementation with rights structures as in Koray and Yildiz (2018), under two different behavioural assumptions, partial honesty and social responsibility. Specifically, we show that unanimity is sufficient for implementation with partially honest agents and, for the case of social responsibility, we provide a full characterization of the implementable rules.

Keywords: Implementation, rights structure, partial honesty, social responsibility, motive.
JEL Classifications: C71; D47; D71.

*School of Economics and Finance, University of St Andrews. *email*: ss436@st-andrews.ac.uk

1 Introduction

Motives and behavioural traits have become increasingly prevalent in mechanism design. After the seminal contributions of Matsushima (2008) and Dutta and Sen (2012) who introduced a minimal honesty motive to the implementation problem,[1] the theory has explored various equilibrium notions and specifications.[2] However, the focus is clearly on noncooperative implementation notions, which implies that the results rely heavily on the specific equilibrium concept.

Instead, we follow the contribution of Koray and Yildiz (2018) who introduce the notion of a *rights structure*. In this setting, the social planner designs a state space, an outcome function that maps states to social outcomes, and an effectivity correspondence, by which she endows agents or coalitions with rights to change the status quo state. In equilibrium, no coalition that has the right to change the state has any incentive to do so. Implementation is achieved when the equilibrium outcomes coincide with the socially optimal outcomes, for any possible preference profile.

We examine two relevant motives in this framework: *partial honesty* and *social responsibility*. In the first case, a partially honest agent is one who prefers to tell the truth, when the welfare she derives from the outcome is not at stake. We assume the existence of at least one partially honest agent and show that *unanimity* is sufficient for implementation in this setting. In the second case, a socially responsible agent is one who breaks ties in favour of a socially optimal outcome. We provide a complete characterization of the implementable rules when all agents are socially responsible. Our conditions are weaker versions of *Maskin-monotonicity* and unanimity. Our results complement the cooperative implementation approach, which is becoming more central in the mechanism design literature.

2 Model

2.1 General setting

The society consists of a set of agents $N = \{1, ..., n\}$, where $n \geq 2$ and a (finite) set of social outcomes X. Each agent i is endowed with a weak preference relation R_i on X, where P_i and I_i is its strict and symmetric part respectively. The set of all possible preferences for each i is denoted by $\mathcal{R}_i$. An n-tuple $R = (R_1, ..., R_n) \in$

[1] For a survey on the frontier of behavioural implementation and some relevant open questions, see Dutta (2019), as well as other contributions in the same volume.

[2] Just to name a few, Lombardi and Yoshihara (2020), Korpela (2014), Savva (2018), Kimya (2017) or Mukherjee et al. (2017). For more recent contributions, see Altun et al. (2020) or Matsushima et al. (2020).

$\mathcal{R}_1 \times ... \times \mathcal{R}_n \equiv \mathcal{R}$, is called a preference profile. Finally, let $L_i(x, R) = \{y \in X | xR_iy\}$ and $I_i(x, R) = \{y \in X | xI_iy\}$.

A *social choice rule* ϕ is a correspondence $\phi : \mathcal{R} \rightrightarrows X$, such that for any R, $\phi(R) \subseteq X$ is nonempty. The image of ϕ is denoted by $\phi(\mathcal{R})$.

A *rights structure* is a triplet $\Gamma = (S, h, \gamma)$ where S is a state space, $h : S \to X$ is an outcome function that maps states to outcomes, and $\gamma : S \times S \rightrightarrows N$ is an effectivity correspondence such that, for any $(s, t) \in S \times S$, $\gamma(s, t)$ is the set of agents[3] who are effective to change the state from s to t. Finally, let $h(S) \equiv \{x \in X|$ there exists $s \in S$, $h(s) = x\}$ be the range of the rights structure.

2.2 Equilibrium notion

Given a profile R, for each $i \in N$, let i's *motive* $\succeq_i^R$ be a weak order on S and define a *motive profile* as $\succeq^R \equiv (\succeq_1^R, ..., \succeq_n^R)$. Given a rights structure $\Gamma = (S, h, \gamma)$, a preference profile R and a motive profile $\succeq^R$ on S, a state $s \in S$ is a *behavioural γ-equilibrium* at R, if for all t and $\{i\} \in \gamma(s, t)$, we have $s \succeq_i^R t$. Let $C(\Gamma, \succeq^R)$ be the set of behavioural γ-equilibrium states at R and $h(C(\Gamma, \succeq^R))$ be the set of outcomes that correspond to the behavioural γ-equilibrium states. Then, a rights structure Γ *behaviourally implements* ϕ, if for all R, $\phi(R) = h(C(\Gamma, \succeq^R))$.

2.3 Motives and implementation concepts

For motives, we consider two potential candidates that we present below:

2.3.1 Partial honesty

Let $\mathcal{G}^*$ be the set of all rights structures $\Gamma = (S, h, \gamma)$, such that $S = X \times \mathcal{R}$. Given $\Gamma \in \mathcal{G}^*$, a *truth correspondence* $T^\Gamma : \mathcal{R} \rightrightarrows S$ is such that, for any R, $T^\Gamma(R) = X \times \{R\}$. Now, given Γ and R, we define $\succsim_i^R$ as a weak order on S as follows:

An agent i is *partially honest* if for all $s, t \in S$ and $R \in \mathcal{R}$:

- If $h(s)I_ih(t)$, $s \in T^\Gamma(R)$ and $t \notin T^\Gamma(R)$, then $s \succ_i^R t$.
- Otherwise, $h(s)R_ih(t) \iff s \succsim_i^R t$.

A SCR ϕ is *behaviourally implementable with partially honest agents* if there exists at least on partially honest agent in N and there exists $\Gamma \in \mathcal{G}^*$, such that for all R, $h(C(\Gamma, \succsim^R)) = \phi(R)$.

[3]An effectivity correspondence can be defined more generally for coalitions. In our setting, we consider *individual-based* rights structures, which is without loss of generality for our weak-core solution concept, as shown by Korpela et al. (2018). For results on implementation in strong-core, see Lombardi et al. (2020).

2.3.2 Social responsibility

Another motive that we examine is social responsibility.[4] In this case, a socially responsible agent is one who breaks ties in favour of a state that corresponds to a socially optimal outcome, when facing indifference. Formally, given a SCR ϕ and $\Gamma = (S, h, \gamma)$, we define the weak order $\trianglerighteq_i^R$ on S as follows:

An agent i is socially responsible if for all R and $s, t \in S$:

- $h(s) I_i h(t)$, $h(s) \in \phi(R)$ and $h(t) \notin \phi(R)$ imply $s \triangleright_i^R t$.
- Otherwise, $h(s) R_i h(t) \iff s \trianglerighteq_i^R t$.

A SCR ϕ is *behaviourally implementable with socially responsible agents* if all agents in N are socially responsible and there exists Γ such that for all R, $h(C(\Gamma, \trianglerighteq^R)) = \phi(R)$.

3 Results

3.1 Previous results

Koray and Yildiz (2018) characterize the set of implementable rules without any behavioural assumptions,[5] in the domain of linear orderings. Korpela et al. (2018) extend the previous result to weak orderings. Their characterization consists of *Maskin-monotonicity*[6] and *unanimity*:

Definition 3.1. Let Y be a set of outcomes with $\phi(\mathcal{R}) \subseteq Y$. A SCR ϕ satisfies Maskin-monotonicity with respect to Y, if for all R, R' and $x \in \phi(R)$, whenever for all i we have $L_i(x, R) \cap Y \subseteq L_i(x, R')$, then $x \in \phi(R')$.

Definition 3.2. A SCR ϕ satisfies unanimity with respect to $Y \supseteq \phi(\mathcal{R})$, if for all R and $x \in Y$, whenever for all i we have $Y \subseteq L_i(x, R)$, then $x \in \phi(R)$.

3.2 Partial honesty

Our result on behavioural implementation with partially honest agents is as follows:

Theorem 1. A SCR ϕ is behaviourally implementable with partially honest agents if it satisfies unanimity with respect to Y.

[4]For a similar concept, see Hagiwara (2018), Lombardi and Yoshihara (2017), or Doğan (2017).

[5]That would be the case where for all i, R and $s, t \in S$, $s \trianglerighteq_i^R t \iff h(s) R_i h(t)$.

[6]Originally due to Maskin (1999).

Proof. Suppose that ϕ satisfies unanimity with respect to Y and consider the following rights structure:

- $S = \{(x, R) \in Y \times \mathcal{R}\}$.
- $h : S \to X$, such that for all $s = (x, R)$, $h(s) = x$.
- $\gamma : S \times S \rightrightarrows N$, such that, for all $s = (x, R)$:

  (i) If $x \in \phi(R)$, then for all i and t, $\{i\} \in \gamma(s, t) \iff xR_i h(t)$.

  (ii) Otherwise, for all i and t, $\{i\} \in \gamma(s, t)$.

We will show that Γ behaviourally implements ϕ. We break the proof into two parts:

Part 1: For all R, $\phi(R) \subseteq h(C(\Gamma, \succsim^R))$:

Consider $s = (x, R)$ with $x \in \phi(R)$. Then for all i and t, $\{i\} \in \gamma(s, t)$ if and only if $xR_i h(t)$, thus $s \in C(\Gamma, \succsim^R)$.

Part 2: For all R, $h(C(\Gamma, \succsim^R)) \subseteq \phi(R)$:

First, we show that there cannot exist $s \in C(\Gamma, \succsim^R)$, such that $s = (x, R')$, where $R' \neq R$. To see this, note that any partially honest agent i is effective to move from s to $t = (x, R)$, where $(x, R) \succ_i^R (x, R')$. So, for all $s \in C(\Gamma, \succsim^R)$, $s = (y, R)$.

Now, consider $s = (y, R) \in C(\Gamma, \succsim^R)$. Then, it must be that for all i, $Y \subseteq L_i(y, R)$. This fulfils the premises of unanimity, so $y \in \phi(R)$.

Thus, we established that $s \in C(\Gamma, \succsim^R) \iff s = (y, R)$ where $y \in \phi(R)$. This completes the proof.

□

3.3 Social responsibility

Below we present out two relevant conditions for behavioural implementation with socially responsible agents:

Definition 3.3. A SCR ϕ satisfies *SR-monotonicity* with respect to $Y \supseteq \phi(\mathcal{R})$, if for all R, R' and $x \in \phi(R)$, whenever for all i we have

$$L_i(x, R) \cap Y \subseteq SL_i(x, R') \cup [I_i(x, R') \setminus \phi(R')],$$

then $x \in \phi(R')$.

Definition 3.4. A SCR ϕ satisfies *SR-unanimity* with respect to $Y \supseteq \phi(\mathcal{R})$, if for all R and $x \in Y$, whenever for all i we have

$$Y \subseteq L_i(x, R) \text{ and } [I_i(x, R) \setminus \{x\}] \cap \phi(R) = \emptyset,$$

then $x \in \phi(R)$.

The intuition behind SR-monotonicity is the following: Suppose that x is socially optimal in R. Now, if for all agents, the outcomes that are ranked weakly below x in R are ranked weakly below x in R', and additionally, no other socially optimal outcome in R' that was weakly below x in R is tied with x in R', then x must be socially optimal in R' as well. SR-unanimity roughly states that, if all agents agree that x is the weakly best outcome in R, and it does not tie with any other socially optimal outcome, then it must be selected as socially optimal. We are now ready to state our second result:

Theorem 2. A SCR ϕ is behaviourally implementable with socially responsible agents if and only if it satisfies SR-monotonicity and SR-unanimity with respect to some Y.

Proof. First we prove the necessity part:

Consider a SCR ϕ that is implementable by $\Gamma = (S, h, \gamma)$ and let $Y = h(S) \supseteq \phi(\mathcal{R})$. Take $R, R', x \in \phi(R)$ such that for all i, $L_i(x, R) \cap Y \subseteq L_i(x, R') \setminus [I_i(x, R') \cap \phi(R')]$, but assume that $x \notin \phi(R')$.

By behavioural implementability, there exists $s \in C(\Gamma, \trianglerighteq^R)$, such that $h(s) = x$. Since $x \notin \phi(R')$, $s \notin C(\Gamma, \trianglerighteq^{R'})$, so there exists t and $\{i\} \in \gamma(s, t)$, such that $t \rhd_i^{R'} s$, where $h(t) \equiv y$. This implies that either $yP_i'x$, or $xI_i'y$ and $y \in \phi(R')$. In both cases, $y \notin L_i(x, R') \setminus [I_i(x, R') \cap \phi(R')]$ and by our assumption, either $y \notin Y$, which is rejected by the definition of Y, or $y \notin L_i(x, R)$. Since $\{i\} \in \gamma(s, t)$, this again contradicts that $s \in C(\Gamma, \trianglerighteq^R)$. So, ϕ satisfies SR-monotonicity.

Now consider R and $x \in Y$ such that for all i, $Y \subseteq L_i(x, R)$ and $[I_i(x, R) \setminus \{x\}] \cap \phi(R) = \emptyset$. Suppose $x \notin \phi(R)$. First, notice that $x \in Y$, so there exists s, $h(s) = x$. Since $x \notin \phi(R)$, there must exist t, and $\{i\} \in \gamma(s, t)$, such that $t \rhd_i^R s$, where $h(t) \equiv y$. So, either yP_ix, or yI_ix and $y \in \phi(R)$. Both cases contradict our premises, so ϕ satisfies SR-unanimity. We conclude with the sufficiency part.

Consider a SCR ϕ that satisfies both conditions with respect to a set Y and consider the following rights structure $\Gamma = (S, h, \gamma)$:

- $S = \{(x, R) \in Y \times \mathcal{R} | x \in \phi(R)\} \cup Y$.

- For all s, $h((x,R)) = h(x) = x$.
- For all i and s:
  - (i) If $s = (x,R)$, then for all t, $\{i\} \in \gamma(s,t) \iff xR_ih(t)$.
  - (ii) If $s = x$, then for all t, $\{i\} \in \gamma(s,t)$.

Part 1: For all R, $\phi(R) \subseteq h(C(\Gamma, \trianglerighteq^R))$:

Notice that for all i and t, $\{i\} \in \gamma(s,t) \iff h(t)R_ix$ and, since $x \in \phi(R)$, $x \in h(C(\Gamma, \trianglerighteq^R))$.

Part 2: For all R, $h(C(\Gamma, \trianglerighteq^R)) \subseteq \phi(R)$:

Let $s \in C(\Gamma, \trianglerighteq^R)$. We distinguish the following cases:

(i) $s = (y, R')$: Since $s \in C(\Gamma, \trianglerighteq^R)$, it must be that, for all $i \in N$ and $z \in L_i(y, R')$, either of the following is true:

- (a) $z \in SL_i(y,R)$, or
- (b) $z \in I_i(y,R)$ and either of the following is true:
  - (1) $z \in \phi(R)$, or
  - (2) $z \notin \phi(R)$.

Suppose that (b)(1) is true. Then, since $(y,R') \in C(\Gamma, \trianglerighteq^R)$, it must be that $y \in \phi(R)$ as well, and there is nothing else to prove. Finally, notice that, from (a) and (b)(2) we have that, for all $i \in N$, $L_i(y,R') \cap Y \subseteq SL_i(y,R) \cup [I_i(y,R) \setminus \phi(R)]$. So, the premises of SR-monotonicity are fulfilled, and we have that $y \in \phi(R)$.

(ii) $s = y$: Since $s \in C(\Gamma, \trianglerighteq^R)$, we have for all i, $Y \subseteq L_i(y,R)$. For the sake of contradiction, assume $y \notin \phi(R)$. Now suppose that there exists $z \in I_i(y,R) \cap \phi(R)$, with $z \neq x$. Then, since $y \in C(\Gamma, \trianglerighteq^R)$, it must be that $y \in \phi(R)$ as well, a contradiction. So, we have that for all $z \in Y \setminus \{x\}$, $I_i(y,R) \cap \phi(R) = \emptyset$. But now, the premises of SR-unanimity are fulfilled and we must have $y \in \phi(R)$, a contradiction. This concludes the proof.

□

It is not hard to show that SR-monotonicity and SR-unanimity are implied by Maskin-monotonicity and unanimity respectively, while they are clearly not equivalent. To conclude, we outline the significance of our Theorem 2 by providing some applications.[7] First, consider the following example:

R		R'	
R_1	R_2	R'_1	R'_2
w	x	w	x
x	y	xy	y
y	w	z	w
z	z		z

Table 1: Example, preferences

Table 1 above shows the preferences of agents 1 and 2 for the preference profiles R and R'. As usual, $\substack{a\\b}$ for agent i means that she strictly prefers a to b, while ab means that i is indifferent between a and b. Let $\phi(R) = \{x\}$, while $\phi(R') = \{y\}$. The reader can verify that ϕ violates Maskin-monotonicity and is thus not implementable without behavioural assumptions, while it satisfies SR-monotonicity. Given that it also satisfies unanimity (which implies SR-unanimity), it is implementable with socially-responsible agents.

To see another example, consider the strong Pareto rule $\phi^{SPO} : \mathcal{R} \rightrightarrows X$, such that for all R, $\phi^{SPO}(R) = SPO(X, R) \equiv \{x \in X|$ there is no $y \in X$ such that for all $i \in N$, yR_ix and for some $j \in N$, $yP_jx\}$. It is well-known that ϕ^{SPO} violates Maskin-monotonicity and is thus not implementable without any behavioural assumptions.[8] We show in Proposition 1 below that this is not the case with socially responsible agents.

Proposition 1. ϕ^{SPO} is behaviourally implementable with socially responsible agents.

Proof. We first show that ϕ^{SPO} satisfies SR-monotonicity with respect to some Y. Consider $x \in \phi^{SPO}(R)$ for some $R \in \mathcal{R}$ and take $Y = \phi^{SPO}(\mathcal{R})$. Now suppose that for all $i \in N$, $L_i(x, R) \cap Y \subseteq L_i(x, R') \setminus [I_i(x, R') \setminus \phi(R')]$ and assume $x \notin \phi^{SPO}(R')$. This implies that there exists $y \in \phi^{SPO}(R')$ such that, for all i, yR'_ix and for some j, yP'_jx. Then, for all i we have that $y \notin L_i(x, R') \setminus [I_i(x, R') \setminus \phi(R')]$. So, by our assumption, $y \notin L_i(x, R)$ or $y \notin Y$. The latter is rejected since $y \in \phi^{SPO}(R') \subseteq \phi^{SPO}(\mathcal{R}) = Y$. Assume the former. We then have for all i, $y \notin L_i(x, R)$, which contradicts that $x \in \phi^{SPO}(R)$.

[7]We are grateful to a referee for motivating us to pursue this.

[8]See Lombardi et al. (2020) for example.

It remains to show that ϕ^{SPO} satisfies SR-unanimity. This is straightforward to see, as ϕ^{SPO} obviously satisfies unanimity, which implies SR-unanimity. This completes the proof. □

Our examples outline the expansion of the set of implementable rules under social responsibility. The possibility of such expansion in a code of rights framework[9] though is still an open question.

[9] A code of rights is a rights structure where $S = X$ and h is the identity map. For results on implementation by codes of rights, see Korpela et al. (2020).

References

Altun, O. A., Barlo, M. and Dalkıran, N. A. (2020), 'Implementation with a sympathizer'.

Doğan, B. (2017), 'Eliciting the socially optimal allocation from responsible agents', *Journal of Mathematical Economics* **73**, 103 – 110.

Dutta, B. (2019), Complete information implementation and behavioral economics, *in* 'The Future of Economic Design', Springer, pp. 317–319.

Dutta, B. and Sen, A. (2012), 'Nash implementation with partially honest individuals', *Games and Economic Behavior* **74**(1), 154–169.

Hagiwara, M. (2018), 'A simple mechanism for double implementation with semi-socially-responsible agents', *Economics Letters* **171**, 51–53.

Kimya, M. (2017), 'Nash implementation and tie-breaking rules', *Games and Economic Behavior* **102**, 138–146.

Koray, S. and Yildiz, K. (2018), 'Implementation via rights structures', *Journal of Economic Theory* **176**, 479–502.

Korpela, V. (2014), 'Bayesian implementation with partially honest individuals', *Social Choice and Welfare* **43**(3), 647–658.

Korpela, V., Lombardi, M. and Vartiainen, H. (2018), 'Implementation by codes of rights', *Available at SSRN 3225865* .

Korpela, V., Lombardi, M. and Vartiainen, H. (2020), 'Do coalitions matter in designing institutions?', *Journal of Economic Theory* **185**, 104953.

Lombardi, M., Savva, F. and Zivanas, N. (2020), 'Implementation in strong core by codes of rights', *Available at SSRN 3563013* .

Lombardi, M. and Yoshihara, N. (2017), 'Natural implementation with semi-responsible agents in pure exchange economies', *International Journal of Game Theory* **46**(4), 1015–1036.

Lombardi, M. and Yoshihara, N. (2020), 'Partially-honest Nash implementation: a full characterization', *Economic Theory* **70**(3), 871–904.

Maskin, E. (1999), 'Nash equilibrium and welfare optimality', *The Review of Economic Studies* **66**(1), 23–38.

Matsushima, H. (2008), 'Role of honesty in full implementation', *Journal of Economic Theory* **139**(1), 353–359.

Matsushima, H. et al. (2020), Implementation, honesty, and common knowledge, Technical report.

Mukherjee, S., Muto, N. and Ramaekers, E. (2017), 'Implementation in undominated strategies with partially honest agents', *Games and Economic Behavior* **104**, 613 – 631.

Savva, F. (2018), 'Strong implementation with partially honest individuals', *Journal of Mathematical Economics* **78**, 27–34.