From Equals to Despots: The Dynamics of Repeated Group Decision Taking with Private Information*

(Preliminary Draft)

Vinicius Carrasco† William Fuchs‡

Abstract

This paper considers the problem faced by n agents who repeatedly have to take a joint action, cannot resort to side payments, and each period are privately informed about their favorite actions. We study the properties of the optimal contract in this environment. We establish that first best values can be arbitrarily approximated (but not achieved) when the players are extremely patient. Also, we show that the provision of intertemporal incentives necessarily leads to a *dictatorial* mechanism: in the long run the optimal scheme converges to the adoption of one player's favorite action.

1 Introduction

There are many situations in which, repeatedly, a group of agents have to take a common action, cannot resort to side payments, and each period are privately informed about their favorite actions. Examples include many supranational organizations such as a Monetary Union or a Common Market. In the former, monetary policy must be jointly taken and, in the latter, a common tariff with the outside world must be adopted each period. At the national level, political coalitions which must jointly decide on policy issues are also a good example. In this paper we study the properties of the optimal contract for environments with such features.

We first show that efficiency can be arbitrarily approximated, but never attained, when players' are sufficiently patient. The intuition goes as follows. In a repeated setting, the promise of continuation (equilibrium) values can play a similar role to the one side payments play in static mechanism design problems. The difference between side payments and continuation values is that the latter can only imperfectly transfer utility across players. In particular, to transfer continuation utility from a player i to another player j in any period t, allocations (decisions) for periods $\tau > t$ must be altered. When players' are sufficiently patient (i.e., their discount rate (δ) is close to one), their current payoff, which is weighted by $(1-\delta)$, becomes insignificant relative to the promised continuation values. Hence, in order to guarantee truth-telling in the

*We have benefited from conversations with Simon Board, Sergio Firpo, Roger Myerson, Alessandro Pavan, Phil Reny, Leo Rezende, Yuliy Sannikov, Andy Skrzypacz, Nancy Stokey, and Balazs Szentes. Ezra Oberfield provided excellent research assistance. All errors are ours.

†Department of Economics, PUC-Rio.

‡Department of Economics, The University of Chicago.

current period, continuation values have to vary only minimally. Since, the (Incentive Compatible) Utility Possibilities Frontier is locally linear, this implies that the associated efficiency losses from the variation in continuation values are arbitrarily small in the limit. The attainment of full efficiency, however, would call for no variation whatsoever in continuation values. This is clearly at odds with the provision of incentives needed for an efficient action to be taken. Hence, full efficiency is not attainable.

Although the limiting efficiency result is of interest, our main focus is on understanding the dynamic properties of the optimal allocation rule for the case in which (although potentially large) the discount factor is strictly smaller than 1 ($\delta << 1$). In order to understand these properties it is useful to keep in mind, as a benchmark, what the first best allocation would entail. The first best would call for a constant weighted average of the players' types. The problem with this allocation when types are private is that the agents away from the center have an incentive to exaggerate their positions. If they expect the other types to be to the left (right) of them they would have an incentive to claim to be far right (far left) and in that way bring the chosen allocation closer to their preferred point. Compared to the case in which players do not care about the future ($\delta = 0$), the optimal allocation is more sensitive to extreme announcements of preferences. This is the case because, when the agents care about the future, they can trade decision power in the current allocation for decision power in the future. More extreme types are given more weight in the current decision but they pay for it by having less influence in future allocations.

As known from static mechanism design, once Incentive Compatibility constraints are taken into account, the agents' utilities have to be adjusted to incorporate the rents derived from their private information. Following Myerson (1981), the adjusted utility is referred to as virtual utility. In our repeated setting, virtual utilities also play a key role. In fact, we show that the dynamics of decision taking is fully determined by: (i) a decision rule that, at each period, maximizes the weighted sum of the agents' (instantaneous) *virtual* utilities and (ii) a process that governs the evolution of the weights given to the agents' virtual utilities on decisions. The dynamics of the decision taking leads to our second and most interesting result. Continuation values vary from period to period reflecting the agents' weights in the allocation rule.[1] Indeed, continuation values tend to increase (higher future decision power) for the agent that reports a less extreme preferred action, and to decrease (lower future decision power) for a player that reports to prefer an extreme action. Such dynamics eventually lead to one player becoming a dictator. Put differently, in the limit, only the preferences of one agent are taken into account to determine current and future allocations.

Our approximate efficiency result can be contrasted with the one obtained by Sonnenschein and Jackson (2007). They study a "budgeting mechanism" which allows the agents to report each possible type (they have a discrete type space) a fixed number of times.[2] The number of times they can report a given type is given by the frequency with which that type should be realized. They prove (Corollary 2 in their paper) that, for any $\epsilon > 0$, their "budgeting mechanism" is, for a finite (but large) number of periods of interaction, less than ϵ inefficient relative to the first best if players are patient. The sources of the inefficiency in their mechanism and in our scheme are quite different though. In their setting, when the last periods get close, agents may not be able to report truthfully, as they might have run out of their budgeted reports for a particular type. Instead, in our setting, the inefficiency arises from the fact that the weights each agent has

[1] Our approach to the analysis of the optimal allocation rule in the repeated game relies on the factorization results of Abreu, Pearce and Stacchetti (1990), which show that the agent's payoff can be split into a current value and a continuation value.

[2] Although not necessarily efficient for given δ, their mechanism has the nice feature that it is robust i.e. the planner does not need to know the exact functional form of the agents preferences.

on the choice of the allocation must vary over time. More remarkably, we slowly drift towards one of the agents becoming a dictator. Indeed, we show that the optimal way to link decisions over time necessarily leads to a dictatorship ex-post.

In a somewhat simpler environment in which there is a binary choice each period and agents can have either have weak or strong preferences for either option, Casella (2005) studies a mechanism in which agents are given a vote every period which they can use over time.[3] The possibility of shifting votes intertemporally allows agents to concentrate their votes when preferences are more intense. Therefore, if one agent has a long string of strong preferences and the other doesn't, the other agent will accumulate a lot of votes and will be able to outvote him in the future. In a voting setting with two players, two binary issues, a continuum of preference intensities, and where votes across issues are cast simultaneously, Hortala-Vallve (2007) shows that if players are allowed to freely distribute a given number of votes across the two issues, the ex-ante efficient decision can be attained. In our setting the continuation values play the same role as the number of remaining votes in Casella's mechanism. However, by considering the optimal mechanism instead of a particular scheme, we don't restrict the accrual of votes to one per period and we allow agents to borrow votes from the future. In our case one of the agents will eventually "run out of votes" and the other will dictate the allocation in the future. In contrast to Hortala-Vallve (2007), decisions are made sequentially in our model. Future decisions are used as an instrument to provide incentives for current decisions. This, in turn, leads to some inefficiency ex-ante and to a dictatorship ex-post.

Our paper also relates to a series of papers that show that continuation values are close substitutes to side payments in repeated settings. Among those, the closest to ours is Athey and Bagwell (2001), who analyze an infinitely repeated Bertrand duopoly, and establish that, for a finite discount factor, monopoly profits can be exactly attained if firms make use of asymmetric continuation values. The difference from our setting is that, for some states that occur with positive probability, firms in their paper can transfer profits perfectly. This allows them to reconcile the variation in promised continuation values – which, in our setting, leads to a dictatorial mechanism – with those values being provided at the region of the Frontier that is linear.

In dynamic insurance problems with one-sided private information such as Thomas and Worral (1990), the privately informed agent's marginal utility is driven to $-\infty$, so that his consumption goes to zero. This is known as the immiseration result.[4] This is related to our dictatorship result. Indeed, an agent who reports to have an extreme type in a given period is like an agent that reports to have a low income realization. The optimal mechanism will respond by giving that agent more weight in the current allocation decision (similarly a higher transfer today). Incentive compatibility then calls for the agent to "pay" for this by forgoing future weight in the allocation decision (future consumption). In their setting, a Principal designing an optimal insurance policy trades-off risk-sharing (which calls for a constant consumption stream) and the provision of incentives – through varying continuation values. No matter if there is one agent or a continuum, this leads to inmisseration in the limit.

In our model, in contrast, privacy of information would not pose a problem if there were just one agent, nor it is a problem in the case in which there is a continuum of agents. In the one agent case, there is no incentive problems for the agent because his report will not be weighted with any other reports. In the continuum of agents case, any report has no effect on the allocation and hence, there is no incentive to lie

[3] Skrzypacz and Hopenhayn (2004) use a similar chip mechanism to sustain collusion in a repeated auctions environment.

[4] Atkeson and Lucas (1990) establish a similar result for an economy with a continuum of agents. They show that the income distribution fans out and in the limit almost all agents are impoverished.

either. In our environment, incentives are harder to provide when there is a small number of agents.

Despite the similarities with the dynamic insurance literature, we find it remarkable that an optimal incentive scheme among ex-ante identical agents leads to the granting of all bargain power to a single player. Also, note that as opposed to immiseration, dictatorship is an absorbing state: once an agent is granted all the decision rights incentive constraints will not bind any longer and the continuation values will be constant rather than constantly drifting towards $-\infty$.

The paper is organized as follows. We introduce the model in section 2. The optimal mechanism and its properties are characterized in Section 3. All proofs are relegated to the Appendix.

2 The Model

There are $n < \infty$ ex-ante symmetric players, $i = 1, 2, ..., n$ who, at every period, must take a new joint action, a. At each period $t \in \{0, 1, ...\}$, they receive privately preference shocks $\theta_i \in [0, 1]$. The preference shocks are i.i.d. over time and across players, and are drawn from a distribution $F(.)$, with density $f(\theta_i) > 0$, which is symmetric around $\frac{1}{2}$.

Their instantaneous utility is a twice continuously differentiable function[5]

$$u(a, \theta_i),$$

with

$$u(\theta_i, \theta_i) \geq u(a, \theta_i) \;\; for\ all\ a,$$

and

$$u_{a,\theta_i}(a, \theta_i) > 0 > u_{aa}(a, \theta_i).$$

Put in words, their preferences are single peaked, with θ_i representing their favorite action.

We additionally assume that they are symmetric around $\frac{1}{2}$: for all $a, \theta_i \in [0, 1]$,

$$u(a, \theta_i) = u(1 - a, 1 - \theta_i).^{6}$$

Symmetry of preferences and the distribution of types around $\frac{1}{2}$ makes the problem itself symmetric around that point. Therefore, it is natural to measure extreme preferences in terms of how distant they are from $\frac{1}{2}$.

After the players observe their preference shocks, they make reports $\hat{\theta}_i, i = 1, ..., n$. Letting the initial history h^0 be the empty set, a public history at time $t > 0$, h^t, is a sequence of (i) past announcements of all players, and (ii) past realized actions:

[5] All the results extend to the case in which the individual players utility function have different forms. That is:

$$u_i(a, \theta_i).$$

This, in turn, implies that our dictatorship result extends to the case in which the welfare criterion is not utilitarian. In fact, for the general case in which player $i's$ Pareto weight on the welfare functional is λ_i, we can proceed as in the Utilitarian criterion case in the text with

$$\widetilde{u}_i(a, \theta_i) = \lambda_i u_i(a, \theta_i).$$

[6] Note that, in particular, this holds whenever an agent with type θ_i is indifferent between any two actions a and b that are equidistant from θ_i.

$$h^t = \{\varnothing, \left(\hat{\theta}_1^0, \hat{\theta}_2^0, ..., \hat{\theta}_n^0, a^0\right), ..., \left(\hat{\theta}_1^{t-1}, \hat{\theta}_2^{t-1}, ..., \hat{\theta}_n^{t-1}, a^{t-1}\right)\}.$$

Given the reports and the history of the game, a history dependent allocation is determined according to a contract, which is a sequence of functions of the form

$$\left\{a_t\left(\hat{\theta}_1^t, \hat{\theta}_2^t, ..., \hat{\theta}_n^t, h^t\right) : [0,1]^n \times [0,1]^{(n+1)t} \to [0,1]\right\}_{t=0}^{\infty}.$$

This contract is chosen a priori before the agents learn their preference shocks.

Let H^t be the set of all public histories h^t. A public strategy for player i is a sequence of functions $\{\hat{\theta}_i^t(.,.)\}_{t,}$ where

$$\hat{\theta}_i^t : H^t \times [0,1] \to [0,1].$$

Each profile of strategies $\hat{\theta} = \left(\left\{\hat{\theta}_1^t(.)\right\}_t, \left\{\hat{\theta}_2^t(.)\right\}_t, ..., \left\{\hat{\theta}_n^t(.)\right\}\right)$ defines a probability distribution over public histories. Letting $\delta \in [0,1)$ denote the discount factor, player i's discounted expected payoff is given by:

$$E\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t u\left(a_t(\hat{\theta}^t, h^t); \theta_i^t\right)\right].$$

We analyze this game using the recursive methods developed by Abreu, Pearce and Stacchetti (1990). More specifically, letting $W \subset \Re^n$ be the set of Public Pure Strategy Equilibria (PPSE) payoffs for the players, we can decompose the payoffs into a current utility $u(a, \theta_i)$ and a continuation value $v_i(\hat{\theta}) \in W$:

$$E_\theta[(1-\delta)u(a\left(\hat{\theta}\right), \theta_i) + \delta v_i(\hat{\theta}_i, \hat{\theta}_{-i})],$$

In other words, any PPSE can be summarized by the actions to be taken in the current period and equilibrium continuation values as a function of the announcements.

3 Properties of the Optimal Allocation Rules

We can use this decomposition to write the Bellman equation that characterizes the frontier of equilibrium values that can be attained in this environment. Let $v = (v_1, v_2, ...v_{n-1})$ denote the expected values for players $1, .., n-1$, denote $V(v)$ as the highest value to player n given that other players expected values are v. Let $\underline{v}$ be lowest value the designer can assign to an agent, and define $\overline{v}$ as the players' payoff when their preferred action is always taken, $\overline{v} = E_\theta[u(\theta_i, \theta_i)]$.[7]

Letting $\theta = (\theta_1, \theta_2, ..., \theta_n)$, we can write $V(v)$ as:

$$V(v) = \max_{a:[0,1]^n \to [0,1], \{w_i:[0,1]^n \to [\underline{v}, \overline{v}]\}_{i=1,...,n-1}} E_\theta[(1-\delta)u(a(\theta), \theta_n) + \delta V(w(\theta))]$$

s.t.

$$E_\theta[(1-\delta)u(a(\theta), \theta_i) + \delta w_i(\theta)] = v_i, i = 1, ..., n-1 \qquad \text{(Promise Keeping } i\text{)}$$

$$E_{\theta_{-i}}[(1-\delta)u(a(\theta), \theta_i) + \delta w_i(\theta)] \geq E_{\theta_{-i}}\left[(1-\delta)u\left(a\left(\theta_{-i}, \widehat{\theta}_i\right), \theta_i\right) + \delta w_i\left(\theta_{-i}, \widehat{\theta}_i\right)\right] \text{ for all } \theta_i, \hat{\theta}_i, i = 1, ..., n-1 \qquad \text{(ICi)}$$

[7] We assume $\underline{v} \leq E_\theta[u(\theta_j, \theta_i)]$ for $i \neq j$, that is the value when somebody else chooses the allocation to equal their type.

$$E_{\theta_{-n}}\left[(1-\delta)\,u\left(a\left(\theta\right),\theta_n\right)+\delta V\left(w\left(\theta\right)\right)\right]\geq E_{\theta_{-n}}\left[(1-\delta)\,u\left(a\left(\widehat{\theta}_n,\theta_{-n}\right),\theta_n\right)+\delta V\left(w\left(\widehat{\theta}_n,\theta_{-n}\right)\right)\right]\ for\ all\ \theta_n,\hat{\theta}_n \tag{ICn}$$

$$w_i\left(\theta\right)\in\left[\underline{v},\overline{v}\right],\ i=1,...n-1,\ for\ all\ \theta. \tag{Feasibility}$$

In this setting, the first best allocation would be a weighted average of the agents' types, with the weights being constant over time. The difficulty of implementing such an allocation is that, whenever his preference shock is different than $\frac{1}{2}$, the agent would have an incentive to exaggerate his report towards the extremes. In particular, for the case in which $\delta=0$, the only way that the principal has to prevent the agents from lying is by making the allocation more insensitive to the their reports. As a result, the allocation is biased towards the center.[8]

When $\delta>0$, continuation values can be used as an additional instrument to get agents to report truthfully. Now, an agent that reports an extreme type can be allowed to have a larger impact on the allocation in an incentive compatible way. The key is to present the agents with a trade-off between the benefit of a larger influence in the current allocation vs. the loss they will incur in future continuation values. This allows the mechanism to take into account the intensity of the agents' preferences, which, in turn, leads to efficiency gains when compared to a static decision taking problem.

Continuation values play a similar role to side payments in standard static incentive problems. The difference between side payments and continuation values is that the latter can only imperfectly transfer utility across players. In particular, to transfer continuation utility from player i to player j in any period t, *allocations* for periods $\tau>t$ must be altered. This, together with the lack of observability, implies that one cannot attain *exact* efficiency as an equilibrium outcome. Indeed, exact efficiency would call for an equilibrium in which for all histories future allocations would not respond to current announcements. Hence, truthtelling would have to be a static best response for the players, and this cannot be attained with an efficient allocation.

Although efficiency cannot be attained, one can arbitrarily approximate it as the players become patient. In fact, when $\delta\rightarrow 1$, the utility value of the current period, which is weighted by $(1-\delta)$, becomes insignificant relative to the continuation values. Hence, in order to guarantee truth-telling in the current period continuation values have to vary only minimally. Since $V(v)$ is locally linear, the associated losses from the variation in continuation values become negligible.

Indeed, letting

$$a^*\left(\theta\right)=\arg\max_a\sum_{i=1}^n u\left(a,\theta_i\right),$$

be the (ex-ante) symmetric Pareto efficient allocation, and $v^{FB}=E_\theta\left[\sum_{i=1}^n u\left(a^*\left(\theta\right),\theta_i\right)\right]$, we have

[8] See Carrasco and Fuchs (2008) for a complete analysis of the static problem with two agents.

Proposition 1 (Approximate Efficiency) *Given $\epsilon > 0$, there exists $\overline{\delta} < 1$ such that, for all $\delta > \overline{\delta}$ the sum of players PPE payoffs at an optimum are within ϵ of v^{FB}.*

This result can be contrasted with the one obtained by Jackson and Sonnenschein (2007). They study a "budgeting mechanism" which allows the agents to report each possible type (they have a discrete type space) a fixed number of times. The number of times they can report a given type is given by the frequency with which that type should statistically be realized. They prove (Corollary 2 in their paper) that, for any $\epsilon > 0$, their "budgeting mechanism" is less than ϵ inefficient relative to the first best if players are patient and face sufficiently many similar problems. Although it appears to operate very differently from the storable votes mechanism proposed by Casella (2005) or our own mechanism, in essence, the budgeting mechanism also presents the players with a trade-off between current allocation and continuation values. The way continuation values vary in Jackson and Sonnenschein (2007) with the current reports is not efficient, but, for δ close to 1, they are sufficient to sustain incentive compatibility and the inefficiency becomes negligible. The sources of the efficiency losses in their mechanism and in our scheme are quite different though. In their setting, when the last periods get close, agents may not be able to report truthfully, as they might have run out of their budgeted reports for a particular type. Instead, in our setting, the inefficiency arises from the fact that the weight each agent has on the choice of the allocation must vary over time.

3.1 The Dynamics of Decision Taking

By assigning multipliers $\{\lambda_i(\theta_i)\}_{\theta_i}$ to the first order condition counterparts of ICi, and γ_i to the Promise Keeping Constraints, we show in the appendix that the first order condition for the optimal current allocation is

$$\begin{aligned} &\left[\tfrac{\partial u(a(\theta),\theta_n)}{\partial a} f(\theta_n) - \tfrac{d\lambda_n(\theta_n)}{d\theta_n}\tfrac{\partial u(a(\theta),\theta_n)}{\partial a} - \lambda_n(\theta_n)\tfrac{\partial^2 u(a(\theta),\theta_n)}{\partial\theta_n\partial a}\right]\prod_{i\neq n} f(\theta_i) \\ &+\sum_{i=1}^{n-1}\left[\left[\gamma_i\tfrac{\partial u(a(\theta),\theta_i)}{\partial a} f(\theta_i) - \tfrac{d\lambda_i(\theta_i)}{d\theta_i}\tfrac{\partial u(a(\theta),\theta_i)}{\partial a} - \lambda_i(\theta_i)\tfrac{\partial^2 u(a(\theta),\theta_i)}{\partial\theta_i\partial a}\right]\prod_{j\neq i} f(\theta_j)\right] = 0. \end{aligned} \qquad \text{(FOC1)}$$

As suggested by Myerson (1984), it is convenient to think about the Lagrangian that yields this first order condition as representing the weighted sum of the agents' virtual utilities.[9] Indeed, defining new multipliers

$$\begin{aligned} \widetilde{\lambda}_i(\theta_i) &= \frac{\lambda_i(\theta_i)}{\gamma_i},\ i = 1, ..., n-1 \\ \widetilde{\lambda}_n(\theta_n) &= \lambda_n. \end{aligned}$$

and, letting agent i's instantaneous virtual utility be

$$\widetilde{u}(a(\theta),\theta_i) = u(a(\theta),\theta_i) - \frac{d\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)\,d\theta_i}u(a(\theta),\theta_i) - \frac{\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)}\frac{\partial u(a(\theta),\theta_i)}{\partial\theta_i},$$

it can be seen from the first order condition for $a(.)$, that the optimal mechanism maximizes the weighted sum of the agents' virtual instantaneous utilities, with the weight given to agent n being equal to one, and

[9] See also Myerson's notes on virtual utility at http://home.uchicago.edu/~rmyerson/research/virtual.pdf

the weight given to agent $i \neq 1$ being equal to γ_i.[10]

In solving for the optimal continuation value for player i, we obtain the following condition:

$$\frac{dV(v)}{dv_i} = E^{\mathcal{Q}^i}\left[\frac{dV(w(\theta))}{dw_i}\right]. \tag{Martingale}$$

By the Envelope Theorem,

$$-\frac{dV(v)}{dv_i} = \gamma_i.$$

so that the weight (γ_i) agent i's virtual utility is given when the action is taken is a martingale process with respect to a distribution $\mathcal{Q}^i$.

It follows that the dynamics of decision taking is fully determined by (i) a decision rule that, at each period, maximizes the weighted sum of the agents' instantaneous virtual utilities, and (ii) the process that governs the evolution of the weights the agents' virtual utilities are given on decisions.

The distribution $\mathcal{Q}^i$ differs from the true distribution of the players' types by an explicit account – through the multipliers and their derivatives – of the incentive compatibility constraints. Except for the change of measure, similar martingale properties for marginal values also hold in many dynamic insurance models.[11]

Our model differs from them in that, in the dynamic insurance models, the problem is that agents have an incentive to claim to be poorer than they actually are. Instead, we face a situation were agents have an incentive to claim their type is more extreme than it actually is. In fact, the symmetry of the problem around type $\frac{1}{2}$, along with the tilting of the optimal allocation toward the middle to curb the players' incentives to exaggerate their preferences, implies that relevant direction in which the Incentive Compatibility constraints bind depends on whether the players's favorite action is above or below $\frac{1}{2}$. The relevant constraints for players whose favorite action is above $\frac{1}{2}$ are those that ensure they don't want to lie upwards. Conversely, for the case in which the players favorite action is below $\frac{1}{2}$, the relevant constraints are the local downward constraints. Therefore, the multipliers on the First Order Condition counterparts of ICi and ICn change sign at $\frac{1}{2}$. The change of measure needed for the property to hold in our setting follows from this point. Moreover, the dynamic insurance models either deal with the case in which there is a single agent or there is a continuum of them. In our setting, privacy of information would not pose a problem if there were just one agent, nor it is a problem in the case in which there is a continuum of agents. In the one agent case, there is no incentive problems for the agent because his report will not be weighted with any other reports. In the continuum of agents case, any report has no effect on the allocation and hence, there is no incentive to lie either. In our environment, incentives are harder to provide since there is a finite number of agents.

It can be shown (this is done in the appendix) that, under the optimal mechanism, for any given $\{w_i\}_{i=1}^{n-1}$ with w_i in $(\underline{v}, \overline{v})$ for all i – i.e., whenever $0 < \gamma_i < \infty$ – continuation values vary from period to period for at least one of the agents. Unlike the insurance models, variation in the continuation values is not necessary in order to provide insurance. For example, in Thomas and Worral (1990), when $\delta = 0$, there is no way the Principal can provide any insurance to the agent that gets a low income realization. In our model,

[10] In comparison to an agent's real utility, the virtual utility incorporates two terms related to the effects an action schedule has on incentives. First, the term $-\frac{d\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)d\theta_i} u(a(\theta), \theta_i)$ captures how tempting it is, for a given agent i, to deviate locally when his preference shock is θ_i. Second, the term $-\frac{\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)} \frac{\partial u(a(\theta), \theta_i)}{\partial \theta_i}$ captures how tempting it is for types *other* than θ_i to report that their preference shock is θ_i.

[11] For examples those studied by Green (1987), Thomas and Worral (1990), and Atkenson and Lucas (1993).

instead, since agents don't know how aligned their interests are it is possible even in the static case to have an allocation which depends on the Agents' types.[12] Nonetheless, in an optimal scheme it will always be efficient to have continuation values varying over time. The intuition for this is similar to the insurance models. Continuation values allow for agents with an extreme type in the current period (poor agents in the insurance models) to get more weight in the current allocation choice (higher current consumption) in exchange for forgoing decision rights (consumption) in the future.

The proof of Theorem 1 (see below) follows similar arguments to those in Thomas and Worral (1990). We first note that the Martingale Convergence Theorem implies that, for all i, $\frac{dV(v)}{dv_i}$ must converge to a random variable. Then, we show by contradiction that this random variable cannot have positive density for any value in $(-\infty, 0)$.[13] Therefore, eventually, the action taken will place weight only to one of the players. Alternatively, eventually, either

$$w_i(\theta) \to \overline{v} \text{ } for \text{ } some \text{ } i, \text{ } or \text{ } V(w(\theta)) \to \overline{v}$$

with probability 1.

Theorem 1 (Dictatorship in the limit) *The provision of intertemporal incentives necessarily leads to a dictatorial mechanism: In the limit as $t \to \infty$, either v_i converges to $\overline{v}$ almost surely for some i, or $V(v)$ converge to $\overline{v}$ almost surely.*

Whenever an agent is promised continuation values of $\overline{v}$, it must be the case that his favorite action is taken from then on. In other words, $\overline{v}$ is an absorbing state. Therefore, eventually, a single player will be given all bargain power over decisions to be taken. Dictatorship is an ex-post consequence of an optimal mechanism in repeated decision taking settings. It is worth pointing out that although Sonnenschein and Jackson's (2007) budgeting mechanism does not have this long run implication, it can lead to even lower values in long run. This happens when the set of reports left to an agent is very different from the distribution of types.

In environments with endogenous participation constraints, such as Fuchs and Lippi (2006), the threat of abandoning the partnership puts a limit on the extent to which one of the agents can dominate the decision process. We believe that incorporating these considerations is an interesting avenue for future research.

References

[1] Abreu, D., Pearce, D. and Stacchetti, E., 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica*, Vol. 58(5), pp. 1041-1063.

[2] Arrow, K., 1979. "The Property Rights Doctrine and Demand Revelation under Incomplete Information", *Economics and Human Welfare.* Academic Press.

[3] Athey, S., and Bagwell, K., 2001, "Optimal Collusion with Private Information", *Rand Journal of Economics,* 32 (3): 428-465.

[4] Atkeson, A., and Lucas, R.E., 1993, "On Efficient Distribution with Private Information", *Review of Economic Studies*, 59, pp.427-453.

[12] See Carrasco and Fuchs (2008) for a detailed analysis of this case when there are two players.

[13] Convergence to a value would violate the property that continuation values must vary to provide incentives.

[5] Ausubel, L. and Deneckere, R., 1993. “A Generalized Theorem of the Maximum” , *Economic Theory*, Vol. 3, No. 1, pp. 99-107.

[6] Casella, A., 2005. "Storable Votes", *Games and Economic Behavior*, 51, pp. 391-419.

[7] Carrasco, V., and Fuchs, W., 2008, "Dividing and Discarding: A Procedure for taking decisions with non-transferable utility", *mimeo.*

[8] d'Apresmont, C. and Gerard-Varet, L., 1979. "Incentives and Incomplete Information", *Journal of Public Economics* 11, pp. 25-45.

[9] Dobb, J., 1953. *Stochastic Processes,* John Wiley and Sons, Inc., New York, N. Y.

[10] Fuchs, W., and Lippi, F., 2006. "Monetary Union with Voluntary Participation", *Review of Economic Studies,* 73, pp. 437-457.

[11] Green, E. "Lending and the smoothing of uninsurable income", in Contractual Arrangements for Intertemporal Trade (E. C. Prescott and N. Wallace, Eds.), Press, Minnesota, 1987.

[12] Jackson, M. and Sonnenschein, H., 2007. "Overcoming Incentive Constraints by Linking Decisions", *Econometrica*, Vol. 75(1), pp. 241-258.

[13] Kolmogorov, A., and Fomin, S., *Introductory Real Analysis.* New York, Dover Publications, 1970.

[14] Milgrom, P, and I. Segal, 2002, "Envelope Theorems for Arbitrary Choice Sets", *Econometrica*, 70 (2) March, 583-601.

[15] Myerson, R., 1981. “Optimal Auction Design”, *Mathematics of Operations Research*, Vol. 6(1) , pp. 58-73.

[16] Myerson, R., 1984. “Cooperative games with imcomplete information”, *International Journal of Game Theory,* Vol. 13 (2) , pp. 69-96.

[17] Luenberger, D., 1969, *Optimization by Vector Space Methods,* John Wiley and Sons, Inc.

[18] Skrzypacz, A. and Hopenhayn, H., 2004. “Tacit Collusion in Repeated Auctions.” Journal of Economic Theory 114 (1), pp. 153-169.

[19] Stokey, N., Lucas, R.E., and Prescott, E., 1989, *Recursive Methods in Economic Dynamics,* Harvard University Press.

[20] Thomas, J. and Worrall, T., 1990. "Income Fluctuations and Asymmetric Information: An Example of a Repeated Principal-Agent Problem", *Journal of Economic Theory*, 51, pp. 367-390.

4 APPENDIX: The Dictatorship Result

We first prove that the function $V(.)$ is strictly concave. This will allow us to make use of Lagrangian methods. In order to do so, we start by pointing out that, since the players' preferences satisfy a single crossing condition, Incentive Compatibility can be replaced by a first order condition for truthtelling and a monotonicity condition.

Lemma 1 *A contract* $(a(.), w_1(.), ...w_{n-1}(.))_\theta$ *is Incentive Compatible if, and only if, it satisfies*

$$E_{\theta_{-i}}\left(\left[\frac{du(a(\theta),\theta_i)}{da}\frac{da(\theta)}{d\theta_i}\right]+\delta\frac{d}{d\theta_i}w_i(\theta)\right)=0,\ i\neq n \qquad \text{(IC Local i)}$$

$$E_{\theta_{-n}}\left(\frac{du(a(\theta),\theta_n)}{da}\frac{da(\theta)}{d\theta_n}+\delta\frac{d}{d\theta_n}V(w_1(\theta),...,w_{n-1}(\theta))\right)=0,\ i=n. \qquad \text{(IC Local n)}$$

$$E_{\theta_{-i}}\left[\frac{\partial u(a(\tau,\theta_{-i}),\theta_i)}{\partial\theta_i}\right]\ is\ non-decreasing\ in\ \tau\ for\ all\ i. \qquad \text{(Expected Monotonicity)}$$

Proof. Standard given that the players' instantaneous utility satisfies a single crossing condition. ■

As a first step toward showing that $V(.)$ is strictly concave, we have

Lemma 2 *For any* $(w_i)_{i=1}^{n-1}\in[\underline{v},\overline{v}]^{n-1}$*, define, for a given* $V_0(.)$ *strictly concave, the sequence* $\{V_k(.)\}_{k\geq 1 w\in[\underline{v},\overline{v}]^{n-1}}$ *recursively as follows*

$$V_k(w)=\max_{\{a(\theta),w(\theta)\}_\theta} E_\theta[(1-\delta)u(a(\theta),\theta_n)+\delta V_{k-1}(w(\theta))]$$

subject to

$$\begin{aligned}
E_\theta[(1-\delta)u(a(\theta),\theta_1)+\delta w_1(\theta)] &= w_1\\
&\vdots\\
E_\theta[(1-\delta)u(a(\theta),\theta_{n-1})+\delta w_{n-1}(\theta)] &= w_{n-1}
\end{aligned}$$

$$E_{\theta_{-i}}[(1-\delta)u(a(\theta),\theta_i)+\delta w_i(\theta)]\geq E_{\theta_{-i}}\left[(1-\delta)u\left(a\left(\widehat{\theta}_i,\theta_{-i}\right),\theta_i\right)+\delta w_i\left(\theta_{-i},\widehat{\theta}_i\right)\right]\ for\ all\ \theta_i,\hat{\theta}_i,\ i\neq n$$

$$E_{\theta_{-n}}[(1-\delta)u(a(\theta),\theta_n)+\delta V_{k-1}(w(\theta_n))]\geq E_{\theta_{-n}}\left[(1-\delta)u\left(a\left(\widehat{\theta}_n,\theta_{-n}\right),\theta_n\right)+\delta V_{k-1}\left(w\left(\widehat{\theta}_n,\theta_{-n}\right)\right)\right]\ for\ all\ \theta_n,\hat{\theta}_n$$

$$w(\theta)\in[\underline{v},\overline{v}]^{n-1}\ for\ all\ \theta.$$

Then, there exists a $\overline{\delta}<1$ *such that, if* $\delta\geq\overline{\delta}$, $V_k(.)$ *is strictly concave for all k*

Proof. We make an induction argument. By hypothesis, $V_0(.)$ is strictly concave. Assume $V_{k-1}(.)$ is strictly concave.

Let $(a_1(\theta), w_1(\theta))$ and $(a_2(\theta), w_2(\theta))$ be, respectively, solutions of the problem in the statement of the Lemma (the one that defines $V_k(,)$) when the promise keeping constraint is indexed by $w_1=(w_{11},...,w_{1n-1})$ and $w_2=(w_{21},...,w_{2n-1})$. Denote $\alpha w_{1j}+(1-\alpha)w_{2j}$ by w_j^α.

If it were feasible to implement $a^\alpha(\theta)=\alpha a_1(\theta)+(1-\alpha)a_2(\theta)$, where $\alpha\in(0,1)$, with continuation values $w^\alpha(\theta)=\alpha w_1(\theta)+(1-\alpha)w_2(\theta)$, we would have, for any player $j\neq n$,

$$\begin{aligned}
& E_\theta\left[(1-\delta)u(a^\alpha(\theta),\theta_j)+\delta w_j^\alpha(\theta)\right] \qquad \text{(Ineq)}\\
> \ & \alpha E_\theta[(1-\delta)u(a_1(\theta),\theta_j)+\delta w_{1j}(\theta)]+(1-\alpha)E_\theta[(1-\delta)u(a_2(\theta),\theta_j)+\delta w_{2j}(\theta)]\\
= \ & \alpha w_{1j}+(1-\alpha)w_{2j}\equiv w_j^\alpha,
\end{aligned}$$

where the first inequality follows from the strict concavity of $u(.,\theta_2)$, and the equality follows from the definition of $(a_1(\theta), w_1(\theta))$ and $(a_2(\theta), w_2(\theta))$.

We consider two cases:

Case 1: There exists $\epsilon > 0$ such that $w_j^\alpha(\theta) - \epsilon > \underline{v}$ for all j and $\theta \in [0,1]^n$ and

$$E_{\theta_{-i}}\left[\frac{du\left(a^\alpha\left(\tau,\theta_{-i}\right),\theta_i\right)}{d\theta_i}\right] \text{ non-decreasing in } \tau \text{ for all } i.$$

Since the inequality in 1 is strict, we can find, for all $j = 1, ..., n-1$ and for some $\overline{w}_j \geq w_j^\alpha$, a non-negative function $g_j(\theta_j, \theta_n; \delta) \equiv h_j(\theta_j; \delta) + h_{nj}(\theta_n; \delta)$ and a $\overline{\delta}$ such that, if $\delta \geq \overline{\delta}$,

$$E_\theta\left[(1-\delta)\, u\left(a^\alpha(\theta), \theta_j\right) + \delta\left[w_j^\alpha(\theta) - g_j(\theta_j, \theta_n; \delta)\right]\right] = \overline{w}_j,$$

$$w_j^\alpha(\theta) - g_j(\theta_j, \theta_n; \delta) > \underline{v}$$

and, at the same time, fo $j = 1, ..., n-1$,

$$-\frac{\delta}{(1-\delta)}\frac{dh_j(\theta_j)}{d\theta_j} = E_{\theta_{-j}}\left[\begin{array}{c} \alpha\frac{du(a_1(\theta),\theta_j)}{da}\frac{da_1(\theta)}{d\theta_j} + (1-\alpha)\frac{du(a_2(\theta),\theta_j)}{da} \\ -\frac{du(a^\alpha(\theta),\theta_j)}{da}\left[\alpha\frac{da_1(\theta)}{d\theta_j} + (1-\alpha)\frac{da_2(\theta)}{d\theta_j}\right] \end{array}\right] \tag{2}$$

and, for agent n,

$$E_{\theta_{-n}}\left[\frac{du\left(a^\alpha(\theta), \theta_n\right)}{da}\left[\alpha\frac{da_1(\theta)}{d\theta_n} + (1-\alpha)\frac{da_2(\theta)}{d\theta_n}\right]\right] + \frac{\delta}{1-\delta}E_{\theta_{-n}}\left[\frac{d}{d\theta_n}V_{k-1}\left(w^\alpha(\theta) - g(\theta;\delta)\right)\right] = 0. \tag{3}$$

Conditions 2 and 3 guarantee Incentive Compatibility. Indeed, with these continuation values, for all players, the first order condition for truthtelling in the above lemma is satisfied. Moreover, by assumption, $a^\alpha(\theta)$ satisfies expected monotonicity. Therefore, $a^\alpha(\theta)$, coupled with continuation values $w_i^\alpha(\theta) - g_i(\theta)$, is feasible when the promised value for player i is $\overline{w}_i \geq w_i^\alpha$.

We then have, for all $\delta \geq \overline{\delta}$,

$$\begin{aligned}
V_k(w^\alpha) &= V_k(\alpha w_1 + (1-\alpha) w_2) \geq V_k(\overline{w}) && \text{(Concavity)} \\
&\geq E_\theta\left[(1-\delta)\, u\left(a^\alpha(\theta), \theta_n\right) + \delta V_{k-1}\left(w^\alpha(\theta) - g(\theta)\right)\right] \\
&> \alpha\left[E_\theta\left[(1-\delta)\, u\left(a_1(\theta), \theta_n\right)\right]\right] + (1-\alpha)\left[E_\theta\left[(1-\delta)\, u\left(a_2(\theta), \theta_n\right)\right]\right] + \delta E_\theta\left[V_{k-1}\left(w^\alpha(\theta)\right)\right] \\
&\geq \left(\begin{array}{c} \alpha\left[E_\theta\left[(1-\delta)\, u\left(a_1(\theta), \theta_n\right)\right] + \delta V_{k-1}\left(w_1(\theta)\right)\right] \\ +(1-\alpha)\left[E_\theta\left[(1-\delta)\, u\left(a_2(\theta), \theta_n\right)\right] + \delta V_{k-1}\left(w_2(\theta)\right)\right] \end{array}\right) \\
&= \alpha V_k(w_1) + (1-\alpha) V_k(w_2),
\end{aligned}$$

where the first inequality follow from the fact that $V_k(.)$ is strictly decreasing, the second inequality follows from the fact that $a^\alpha(\theta)$ along with $w^\alpha(\theta) - g(\theta;\delta)$ is feasible when the promised value for players $j = 1, .., n-1$ are $\overline{w}_j$ and $\delta \geq \overline{\delta}$, the third inequality follows from strict concavity of Player n's instantaneous payoff and from the fact that $V_{k-1}(.)$ is decreasing, and the fourth inequality follows from the Concavity of V_{k-1}. It follows that $V_k(.)$ is strictly concave.

Case 2: $w_j^\alpha(\theta) = \underline{v}$ for some j and for all θ belonging to a (positive probability) set $A \subset [0,1]^n$ and/or $a^\alpha(\theta)$ does not satisfy expected monotonicity.

The same procedure (changing $a^\alpha(\theta)$) applies to both cases, so we focus on the situation in which $w_j^\alpha(\theta) = \underline{v}$ for some j and for all θ in some $A \subset [0,1]^n$.

Denote by J the set of all players for which $w_j^\alpha(\theta) = \underline{v}$ for some θ. Since the inequality in 1 is strict, we can find, for some $\overline{\delta}_1$ and $\overline{w}_j > w_j^\alpha, j \in J$, a function $l(\theta;\delta) = \sum_{j \in J} l_j(\theta j; \delta) + l_n(\theta_n; \delta)$ such that

$$E_\theta\left[(1-\delta)\, u\left(a^\alpha(\theta) + l(\theta), \theta_j\right) + \delta w_j^\alpha(\theta)\right] = \overline{w}_j,$$

$$E_{\theta_{-j}}\left[\frac{du\left(a^{\alpha}\left(\theta\right)+l\left(\theta\right),\theta_j\right)}{da}\left[\frac{d\left(a^{\alpha}\left(\theta\right)+l\left(\theta\right)\right)}{d\theta_j}\right]-\left[\alpha\frac{du\left(a_1\left(\theta\right),\theta_j\right)}{da}\frac{da_1\left(\theta\right)}{d\theta_j}-\left(1-\alpha\right)\frac{du\left(a_2\left(\theta\right),\theta_j\right)}{da}\frac{da_2\left(\theta\right)}{d\theta_j}\right]\right]=0,$$

and

$$E_{\theta_{-n}}\left[\frac{du\left(a^{\alpha}\left(\theta\right)+l\left(\theta\right),\theta_j\right)}{da}\left[\alpha\frac{da_1\left(\theta\right)}{d\theta_j}+\left(1-\alpha\right)\frac{da_2\left(\theta\right)}{d\theta_j}+\frac{dl_n\left(\theta_n\right)}{d\theta_n}\right]\right]+\frac{\delta}{1-\delta}E_{\theta_{-n}}\left[\frac{d}{d\theta_n}V_{k-1}\left(w^{\alpha}\left(\theta\right)\right)\right]=0,$$

for all $\delta\geq\overline{\delta}_1$.

For all players $i\notin J$, as in Case 1, we can find a $\overline{\delta}_2$ and non-negative function $h_i\left(\theta_i;\delta\right)$ such that, whenever $\delta\geq\overline{\delta}_2$,

$$E_{\theta}\left[\left(1-\delta\right)u\left(a^{\alpha}\left(\theta\right)+l\left(\theta;\delta\right),\theta_i\right)+\delta\left[w_i^{\alpha}\left(\theta\right)-h_i\left(\theta_i;\delta\right)\right]\right]=\overline{w}_i,$$

and

$$E_{\theta_{-j}}\left[\left(1-\delta\right)\frac{du\left(a^{\alpha}\left(\theta\right)+l\left(\theta;\delta\right),\theta_i\right)}{da}\left[\frac{d\left[a^{\alpha}\left(\theta\right)+l\left(\theta;\delta\right)\right]}{d\theta_i}\right]+\delta\frac{d\left[w_i^{\alpha}\left(\theta\right)-h_i\left(\theta_i;\delta\right)\right]}{d\theta_i}\right]=0,$$

for some $\overline{w}_i>w_i^{\alpha}$. Hence, $a^{\alpha}\left(\theta\right)+l\left(\theta;\delta\right)$ coupled with continuation values $\widetilde{w}_j\left(\theta\right)=w_j^{\alpha}\left(\theta\right)$ for $j\in J$, and $\widetilde{w}_i\left(\theta\right)=w_i^{\alpha}\left(\theta\right)-h_i\left(\theta;\delta\right)$ for the other players, is feasible when promised values are $\overline{w}>w^{\alpha}$ and $\delta\geq\max\left\{\overline{\delta}_1,\overline{\delta}_2\right\}$.

It then follows, that whenever $\delta\geq\overline{\delta}$ (which is defined below and larger than $\max\left\{\overline{\delta}_1,\overline{\delta}_2\right\}$),

$$\begin{aligned}V_k\left(w^{\alpha}\right)&=V_k\left(\alpha w_1+\left(1-\alpha\right)w_2\right)>V_k\left(\overline{w}\right)\\&\geq E_{\theta}\left[\left(1-\delta\right)u\left(a^{\alpha}\left(\theta\right)+l\left(\theta\right)\right)+\delta V_{k-1}\left(\widetilde{w}\left(\theta\right)\right)\right]\\&\geq E_{\theta}\left[\left(1-\delta\right)u\left(a^{\alpha}\left(\theta\right)\right)+\delta\left[\alpha V_{k-1}\left(w_1\left(\theta\right)\right)+\left(1-\alpha\right)V_{k-1}\left(w_2\left(\theta\right)\right)\right]\right]\\&\geq\left(\begin{array}{c}\alpha\left[E_{\theta}\left[\left(1-\delta\right)u\left(a_1\left(\theta\right),\theta_n\right)\right]+\delta V_{k-1}\left(w_1\left(\theta\right)\right)\right]\\+\left(1-\alpha\right)\left[E_{\theta}\left[\left(1-\delta\right)u\left(a_2\left(\theta\right),\theta_n\right)\right]+\delta V_{k-1}\left(w_2\left(\theta\right)\right)\right]\end{array}\right)\\&=\alpha V_k\left(w_1\right)+\left(1-\alpha\right)V_k\left(w_2\right),\end{aligned}$$

where the first inequality follows from $V_k\left(.\right)$ being strictly decreasing, the second follows because $a^{\alpha}\left(.\right)+l\left(.\right)$ and $\widetilde{w}\left(.\right)$ are feasible when promised values are $\overline{w}$. The third inequality, which is the key one, holds because, since V_{k-1} is strictly concave,

$$V_{k-1}\left(\widetilde{w}\left(\theta\right)\right)>\alpha V_{k-1}\left(w_1\left(\theta\right)\right)+\left(1-\alpha\right)V_{k-1}\left(w_2\left(\theta\right)\right),$$

so that there exists a $\overline{\delta}_3$ such that, if $\delta\geq\overline{\delta}_3$, the inequality holds.

Letting $\overline{\delta}=\max\left\{\overline{\delta}_1,\overline{\delta}_2,\overline{\delta}_3\right\}$, the result follows. ■

Proposition 2 *There exists a $\overline{\delta}<1$ such that, if $\delta\geq\overline{\delta}$, $V\left(.\right)$ is strictly concave.*

Proof. We prove the result in five steps.

Define

$$T\left(V\right)\left(w\right)=\max_{\left\{a\left(\theta\right),w\left(\theta\right)\right\}_{\theta}}E_{\theta}\left[\left(1-\delta\right)u\left(a\left(\theta\right),\theta_n\right)+\delta V\left(w\left(\theta\right)\right)\right]$$

subject to

$$\begin{aligned}E_{\theta}\left[\left(1-\delta\right)u\left(a\left(\theta\right),\theta_1\right)+\delta w_1\left(\theta\right)\right]&=w_1\\&\vdots\\E_{\theta}\left[\left(1-\delta\right)u\left(a\left(\theta\right),\theta_{n-1}\right)+\delta w_{n-1}\left(\theta\right)\right]&=w_{n-1}\end{aligned}$$

$$E_{\theta_{-n}}\left[(1-\delta)\,u\left(a\left(\theta\right),\theta_n\right)+\delta V_{k-1}\left(w\left(\theta_n\right)\right)\right]\geq E_{\theta_{-n}}\left[(1-\delta)\,u\left(a\left(\widehat{\theta}_n,\theta_{-n}\right),\theta_n\right)+\delta V_{k-1}\left(w\left(\widehat{\theta}_n,\theta_{-n}\right)\right)\right]\ for\ all\ \theta_n,\hat{\theta}_n$$

$$E_{\theta_{-i}}\left[(1-\delta)\,u\left(a\left(\theta\right),\theta_i\right)+\delta w_i\left(\theta\right)\right]\geq E_{\theta_{-i}}\left[(1-\delta)\,u\left(a\left(\theta_{-i},\widehat{\theta}_i\right),\theta_i\right)+\delta w_i\left(\theta_{-i},\widehat{\theta}_i\right)\right]\ for\ all\ \theta_i,\hat{\theta}_i,i=1,...,n-1.$$

$$w\left(\theta\right)\in\left[\underline{v},\overline{v}\right]^{n-1}\ for\ all\ \theta.$$

STEP 1: For any δ, the set of $\{a\left(\theta\right),w\left(\theta\right)\}_{\theta}$ that satisfies the constraints of the above problem is compact and upper hemi-continuous.

We prove compactness. The proof of upper hemi-continuity of the constraint set follows similar steps.

Note that $\{a\left(\theta\right),w\left(\theta\right)\}_{\theta}$ satisfies the Incentive Compatibility constraints if, and only if:

$$\begin{aligned}&(1-\delta)\,E_{\theta_{-i}}\left[u\left(a\left(\theta\right),\theta_i\right)\right]+\delta E_{\theta_{-i}}\left[w_i\left(\theta\right)\right] \qquad \text{(Envelope)}\\ =\ &(1-\delta)\,E_{\theta_{-i}}\left[u\left(a\left(0,\theta_{-i}\right),0\right)\right]+\delta E_{\theta_{-i}}\left[w_i\left(0,\theta_{-i}\right)\right]\\ &+(1-\delta)\int_0^{\theta_i}E_{\theta_{-i}}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i}\right),\tau\right)\right]d\tau,\ \ i=1,...,n-1\end{aligned}$$

and

$$\begin{aligned}&(1-\delta)\,E_{\theta_{-n}}\left[u\left(a\left(\theta\right),\theta_n\right)\right]+\delta E_{\theta_{-n}}\left[V\left(w\left(\theta\right)\right)\right] \qquad \text{(Envelope 1)}\\ =\ &(1-\delta)\,E_{\theta_{-n}}\left[u\left(a\left(0,\theta_{-n}\right),0\right)\right]+\delta E_{\theta_{-n}}\left[V\left(w\left(0,\theta_{-n}\right)\right)\right]\\ &+(1-\delta)\int_0^{\theta_n}E_{\theta_{-n}}\left[u_{\theta_n}\left(a\left(\tau,\theta_{-n}\right),\tau\right)\right]d\tau.\end{aligned}$$

and $E_{\theta_{-i}}\left[u_{\theta_i}\left(a\left(\tau,\theta_{-i}\right),\theta_i\right)\right]$ being non-decreasing in τ. (The envelope conditions follow from Milgrom and Segal (2002)). Using 1, after some integration by parts, the Promise Keeping constraints can be written as

$$(1-\delta)\,E_{\theta_{-i}}\left[u\left(a\left(0,\theta_{-i}\right),0\right)\right]+\delta E_{\theta_{-i}}\left[w_i\left(0,\theta_{-i}\right)\right]+(1-\delta)\,E_{\theta}\left[u_{\theta_i}\left(a\left(\theta_i,\theta_{-i}\right),\theta_i\right)\frac{\left(1-F\left(\theta_i\right)\right)}{f\left(\theta_i\right)}\right]=w_i.$$

Feasibility, in turn, calls for

$$w_i\left(\theta\right)\in\left[\underline{v},\overline{v}\right]\ \text{for all } i,\ \text{and } \theta$$

Now, take a sequence $\{a_k\left(.\right),w_k\left(.\right)\}_k$ satisfying all those constraints of the problem. We show that there exists a convergent subsequence.

Since $E_{\theta_{-j}}\left[u_{\theta_j}\left(a_k\left(\tau,\theta_{-j}\right),\tau\right)\right]$ is a sequence of non-decreasing and uniformly bounded functions, by Helly's Selection Theorem (Kolmogorov and Fomin, 1970, p. 373), there exists a subsequence $E_{\theta_{-j}}\left[u_{\theta_j}\left(a_{k_s}\left(\tau,\theta_{-j}\right),\tau\right)\right]$ that converges to a non-decreasing $E_{\theta_{-j}}\left[u_{\theta_j}\left(a\left(\tau,\theta_{-j}\right),\tau\right)\right]$, $j=1,...,n$. Moreover, for $i=1,...,n-1$,

$$\left[\begin{array}{c}(1-\delta)\,E_{\theta_{-i}}\left[u\left(a_k\left(\theta\right),\theta_i\right)\right]\\+\delta E_{\theta_{-i}}\left[w_{ki}\left(\theta\right)\right]\end{array}\right]=$$

$$\left[\begin{array}{c}(1-\delta)\,E_{\theta_{-i}}\left[u\left(a_k\left(0,\theta_{-i}\right),0\right)\right]+\delta E_{\theta_{-i}}\left[w_{ki}\left(0,\theta_{-i}\right)\right]\\+(1-\delta)\displaystyle\int_0^{\theta_i}E_{\theta_{-i}}\left[u_{\theta_i}\left(a_k\left(\tau,\theta_{-i}\right),\tau\right)\right]d\tau\end{array}\right]$$

and for n

$$\begin{bmatrix} (1-\delta) E_{\theta_{-n}} [u (a_k (\theta), \theta_n)] \\ +\delta E_{\theta_{-n}} [V (w_k (\theta))] \end{bmatrix} =$$

$$\begin{bmatrix} (1-\delta) E_{\theta_{-n}} [u (a_k (0, \theta_{-n}), 0)] + \delta E_{\theta_{-n}} [V (w_k (0, \theta_{-n}))] \\ + (1-\delta) \int\limits_0^{\theta_n} E_{\theta_{-n}} [u_{\theta_n} (a_k (\tau, \theta_{-n}), \tau)] \, d\tau \end{bmatrix}$$

We now argue that, for both expressions, the right hand side converges (possibly along subsequences). We show this for an arbitrary agent i (the analysis for player n is analogous).

Note that

$$(1-\delta) E_{\theta_{-i}} [u (a_k (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_{ki} (0, \theta_{-i})]$$

is a sequence of real numbers that lies in a compact set. Therefore, there exists a subsequence that converges. Let the convergent subsequence be

$$(1-\delta) E_{\theta_{-i}} [u (a_{k_s} (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_{k_s i} (0, \theta_{-i})]$$

and denote its limit by

$$(1-\delta) E_{\theta_{-i}} [u (a (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w (0, \theta_{-i})] .$$

Also, since there exists a subsequence of $E_{\theta_{-i}} [u_{\theta_i} (a_{k_s} (\tau, \theta_{-n}), \tau)]$ that converges to a non-decreasing $E_{\theta_{-i}} [u_{\theta_i} (a (\tau, \theta_{-n}), \tau)]$, one has, by the Dominated Convergence Theorem, that

$$(1-\delta) \int\limits_0^{\theta_n} E_{\theta_{-i}} [u_{\theta_i} (a_{k_s} (\tau, \theta_{-i}), \tau)] \, d\tau \to (1-\delta) \int\limits_0^{\theta_n} E_{\theta_{-i}} [u_{\theta_i} (a (\tau, \theta_{-i}), \tau)] \, d\tau.$$

Therefore, letting $a(.)$ and $w(.)$ be the functions for which

$$\lim_{s\to\infty} \begin{bmatrix} (1-\delta) E_{\theta_{-i}} [u (a_{k_s} (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_{k_s i} (0, \theta_{-i})] \\ + (1-\delta) \int\limits_0^{\theta_n} E_{\theta_{-n}} [u_{\theta_i} (a_{k_s} (\tau, \theta_{-n}), \tau)] \, d\tau \end{bmatrix}$$

$$= \begin{bmatrix} (1-\delta) E_{\theta_{-i}} [u (a (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_i (0, \theta_{-i})] \\ + (1-\delta) \int\limits_0^{\theta_n} E_{\theta_{-i}} [u_{\theta_i} (a (\tau, \theta_{-i}), \tau)] \, d\tau. \end{bmatrix},$$

it follows that 1 holds at $(a(.), w(.))$.

As for the Promise Keeping constraints, since, for all k and i,

$$(1-\delta) E_{\theta_{-i}} [u (a_k (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_{ki} (0, \theta_{-i})] + (1-\delta) E_\theta \left[u_{\theta_i} (a_k (\theta_i, \theta_{-i}), \theta_i) \frac{(1 - F(\theta_i))}{f(\theta_i)} \right] = w_i,$$

and, along a subsequence, the left hand side converges, one has, invoking again the Dominated Convergence Theorem, that

$$(1-\delta) E_{\theta_{-i}} [u (a (0, \theta_{-i}), 0)] + \delta E_{\theta_{-i}} [w_i (0, \theta_{-i})] + (1-\delta) E_\theta \left[u_{\theta_i} (a (\theta_i, \theta_{-i}), \theta_i) \frac{(1 - F(\theta_i))}{f(\theta_i)} \right] = w_i.$$

Finally, since, for all k,

$$w_k(\theta) \in [\underline{v}, \overline{v}]^{n-1}$$

one must have

$$w(\theta) \in [\underline{v}, \overline{v}]^{n-1}.$$

Therefore, the choice set is compact.

STEP 2: If $V(.)$ is continuous, $T(V)$ is also continuous

Given STEP 1, this follows from Theorem 2 of Ausubel and Deneckere (1993).

STEP 3: $T(.)$ is a contraction of modulus δ.

Denote by $C(w)$ the set of feasible actions and continuation values, $\{a(.), w(.)\}$, given current values w. We have

$$\begin{aligned} T(V_1) &= \max_{\{a(.),w(.)\}\in C(w)} E_\theta \left[(1-\delta) u(a(\theta), \theta_n) + \delta V_1(w(\theta))\right] \\ &= \max_{\{a(.),w(.)\}\in C(w)} E_\theta \left[(1-\delta) u(a(\theta), \theta_n) + \delta V_2(w(\theta)) + \delta \left[V_1(w(\theta)) - V_2(w(\theta))\right]\right] \\ &\leq \max_{\{a(.),w(.)\}\in C(w)} E_\theta \left[(1-\delta) u(a(\theta), \theta_n) + \delta V_2(w(\theta))\right] + \delta \|V_1 - V_2\| \\ &= T(V_2) + \delta \|V_1 - V_2\|, \end{aligned}$$

where $\|.\|$ is the sup norm.

Therefore,

$$\|T(V_1) - T(V_2)\| \leq \delta \|V_1 - V_2\|.$$

STEP 4: The sequence $\{V_k(.)\}_{k\geq 1 w\in[\underline{v},\overline{v}]^{n-1}}$, with $V_0(w) = 0$ for all w, converges to $V(.)$

This follows from the fact that T is a contraction and the set, $C[\underline{v}, \overline{v}]$, of continuous functions over $[\underline{v}, \overline{v}]$ endowed with the sup norm is a complete metric space..

STEP 5: There exists a $\overline{\delta} < 1$ such that, for all $\delta \geq \overline{\delta}$, $V(.)$ is strictly concave.

Pick the $\overline{\delta}$ such that the sequence of $\{V_k(.)\}$ is strictly concave for $\delta \geq \overline{\delta}$. As each element in the sequence $\{V_k(.)\}$ is strictly concave, the limit must be concave. Now, using the concavity of $V(.)$ and proceeding exactly as in the proof of the above Lemma, it is easy show that $V(.)$ must be, in fact, strictly concave. ■

A property of $V(.)$ that we will use to prove the dictatorship result is

Lemma 3 $V(.)$ *is continuously differentiable over* $(\underline{v}, \overline{v})^{n-1}$

Proof. Since $V(.)$ is concave, this follows from Corollary 2 in Milgrom and Segal (2002). ■

Now, ignoring Expected Monotonicity, we construct the Lagrangian by assigning multipliers to the local IC $(\lambda_i(\theta_i))_{i=1,\dots,n,\theta_i\in[0,1]}$, and PK $(\gamma_i)_{i=1}^{n-1}$ constraints:

$$V(v) = \max_{\{a(.),w_j(.)\}} \left[\begin{array}{c} E_\theta \left[(1-\delta) u(a(\theta), \theta_n) + \delta V(w(\theta))\right] \\ \sum_{i=1}^{n-1} \left[\gamma_i \left(E_\theta \left[(1-\delta) u(a(\theta), \theta_i) + \delta w_i(\theta)\right] - v_i\right)\right] \\ + \sum_{i=1}^{n-1} \int_0^1 \left[\lambda_i(\theta_i) \left(E_{\theta_{-i}} \left(\begin{array}{c} (1-\delta) \left[\frac{du(a(\theta),\theta_i)}{da} \frac{da(\theta)}{d\theta_i}\right] \\ +\delta \frac{d}{d\theta_i} \left[w_i(\theta)\right] \end{array} \right)\right)\right] d\theta_i \\ + \int_0^1 \left[\lambda_n(\theta_n) \left(E_{\theta_{-n}} \left(\begin{array}{c} \left[(1-\delta) \frac{du(a(\theta),\theta_n)}{da} \frac{da(\theta)}{d\theta_n}\right] \\ +\delta \frac{d}{d\theta_n} V(w(\theta)) \end{array} \right)\right)\right] d\theta_n \end{array} \right]$$

As is standard (see Theorems 1 and 2 in sections 8.3-8.4 of Luenberger (1969))[14], $\{a^*(\theta), w_i^*(\theta)\}_\theta$ – with $a^*(.)$ satisfying expected monotonicity strictly – is optimal if, and only if, there are multipliers $\{\lambda_i(\theta_i), \gamma_i\}_{i=1,..,n,\theta_i}$ for which $\{a^*(\theta), w_i^*(\theta)\}_\theta$ maximizes the above Lagrangian.

4.1 The Lagrangian Representation and the Result

Some rounds of integration by parts allow us to write our program of interest as

$$V(v) = \max_{\{a(.),w(.)\}} \left[\begin{array}{c} E_\theta\left[(1-\delta)u(a(\theta),\theta_n) + \delta V(w(\theta))\right] \\ \sum_{i=1}^{n-1}\left[\gamma_i\left(E_\theta\left[(1-\delta)u(a(\theta),\theta_i) + \delta w_i(\theta)\right] - v_i\right)\right] \\ +\sum_{i=1}^{n-1}\left(\lambda_i(\theta_i)\left[(1-\delta)E_{\theta_{-i}}\left[u(a(\theta),\theta_i)\right]\right] + \delta E_{\theta_{-i}}\left[w(\theta)\right]|\right)_{\theta_i=0}^{\theta_i=1} \\ -\sum_{i=1}^{n-1}\left(\begin{array}{c} \int_0^1 \left[\frac{d\lambda_i(\theta_i)}{d\theta_i}\left[(1-\delta)E_{\theta_{-i}}\left[u(a(\theta),\theta_i)\right]\right]\right] d\theta_i \\ +\int_0^1 \lambda_i(\theta_i)\left[(1-\delta)E_{\theta_{-i}}\left[\frac{du(a(\theta),\theta_i)}{d\theta_i}\right]\right] d\theta_i \end{array} \right) \\ -\delta\sum_{i=1}^{n-1}\left[\int_0^1 \left[\frac{d\lambda_i(\theta_i)}{d\theta_i}E_{\theta_{-i}}\left[w_i(\theta)\right]\right] d\theta_i\right] \\ +\left(\lambda_n(\theta_n)\left[(1-\delta)E_{\theta_{-n}}\left[u(a(\theta),\theta_n)\right]\right] + \delta E_{\theta_{-n}}\left[V(w(\theta))\right]|\right)_{\theta_n=0}^{\theta_n=1} \\ -\int_0^1 \left[\frac{d\lambda_n(\theta_n)}{d\theta_n}\left[(1-\delta)E_{\theta_{-n}}\left[u(a(\theta),\theta_n)\right]\right]\right] d\theta_n \\ -\int_0^1 \lambda_n(\theta_n)\left[(1-\delta)E_{\theta_{-n}}\left[\frac{du(a(\theta),\theta_n)}{d\theta_n}\right]\right] d\theta_n \\ -\delta\int_0^1 \left[\frac{d\lambda_n(\theta_n)}{d\theta_n}E_{\theta_{-n}}\left[V(w(\theta))\right]\right] d\theta_n. \end{array} \right]$$

As suggested by Myerson (1984), it is convenient to think about the Lagrangian as representing the weighted sum of the agents' virtual utilities.[15]

Indeed, if, for agents $i = 1, ..., n-1$, one defines new multipliers

$$\widetilde{\lambda}_i(\theta_i) = \frac{\lambda_i(\theta_i)}{\gamma_i},$$

[14] The concavity of $V(.)$ in our setting plays the role Proposition 1 in section 8.3 plays in Theorems 1 and 2 of Luenberger (1969).

[15] See also Myerson's notes on virtual utility at http://home.uchicago.edu/~rmyerson/research/virtual.pdf

the FOC wrt $a(.)$, and $w_j(.)$ are, for $\theta \in (0,1)^n$, respectively,

$$\begin{aligned} &\left(\begin{array}{c} \left[\frac{\partial u(a(\theta),\theta_n)}{\partial a} f(\theta_n) - \frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{\partial u(a(\theta),\theta_n)}{\partial a} - \lambda_n(\theta_n)\frac{\partial^2 u(a(\theta),\theta_n)}{\partial\theta_n\partial a}\right]\prod_{i\neq n} f(\theta_i) \\ +\sum_{i=1}^{n-1}\left[\left[\gamma_i\frac{\partial u(a(\theta),\theta_i)}{\partial a} f(\theta_i) - \frac{d\lambda_i(\theta_i)}{\gamma_i d\theta_i}\frac{\partial u(a(\theta),\theta_i)}{\partial a} - \frac{\lambda_i(\theta_i)}{\gamma_i}\frac{\partial^2 u(a(\theta),\theta_i)}{\partial\theta_i\partial a}\right]\prod_{j\neq i} f(\theta_j)\right] \end{array}\right) \\ &\left(\begin{array}{c} \left[\frac{\partial u(a(\theta),\theta_n)}{\partial a} f(\theta_n) - \frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{\partial u(a(\theta),\theta_n)}{\partial a} - \lambda_n(\theta_n)\frac{\partial^2 u(a(\theta),\theta_n)}{\partial\theta_n\partial a}\right]\prod_{i\neq n} f(\theta_i) \\ +\sum_{i=1}^{n-1}\gamma_i\left[\left[\frac{\partial u(a(\theta),\theta_i)}{\partial a} f(\theta_i) - \frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\frac{\partial u(a(\theta),\theta_i)}{\partial a} - \widetilde{\lambda}_i(\theta_i)\frac{\partial^2 u(a(\theta),\theta_i)}{\partial\theta_i\partial a}\right]\prod_{j\neq i} f(\theta_j)\right] \end{array}\right) \\ = \;& 0, \end{aligned}$$

and

$$\left(\begin{array}{c} \frac{dV(w(\theta))}{dw_j}\prod_i f(\theta_i) + \gamma_i\prod_i f(\theta_i) - \frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\prod_{i\neq j} f(\theta_i) \\ -\frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{dV(w(\theta))}{dw_j}\prod_{i\neq n} f(\theta_i) \end{array}\right) = 0.$$

Defining agent i's ($i \neq n$) virtual instantaneous utility as being

$$u(a(\theta),\theta_i) - \frac{d\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)\,d\theta_i} u(a(\theta),\theta_i) - \frac{\widetilde{\lambda}_i(\theta_i)}{f(\theta_i)}\frac{\partial u(a(\theta),\theta_i)}{\partial\theta_i}$$

and agent n's virtual instantaneous utility as being

$$u(a(\theta),\theta_n) - \frac{d\lambda_n(\theta_n)}{f(\theta_n)\,d\theta_n} u(a(\theta),\theta_n) - \frac{\lambda_n(\theta_n)}{f(\theta_n)}\frac{\partial u(a(\theta),\theta_n)}{\partial\theta_n},$$

it can be readily seen from the FOC for $a(.)$, the optimal mechanism maximizes the weighetd sum of the agents' virtual instantaneous utilities, with the weight to agent n being one, and the weight to agent $i \neq 1$ being equal to γ_i.

The first order conditions for continuation values lead to the following result:

Lemma 4 (Martingale Lemma) *There exists a measure $\mathcal{Q}^i$ such that Player 1´s marginal value (with respect to player i's continuation value) follows a martingale, i.e.*

$$E^{\mathcal{Q}^i}\left[\frac{dV(w(\theta))}{dw_i}\right] = \frac{dV(w)}{dw_i}$$

Proof. The FOC wrt $w_i(.)$ is

$$\left(\begin{array}{c} \frac{dV(w(\theta))}{dw_i}\prod_j f(\theta_j) - \frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{dV(w(\theta))}{dw_i}\prod_{j\neq n} f(\theta_j) \\ +\gamma_i\prod_j f(\theta_j) - \frac{d\lambda_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_i) \end{array}\right) = 0.$$

This can be re-written as

$$\frac{dV(w(\theta))}{dw_i}\left(\prod_j f(\theta_j) - \frac{d\lambda_n(\theta_n)}{d\theta_n}\prod_{j\neq n} f(\theta_j)\right) = -\gamma_i\left(\prod_j f(\theta_j) - \frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)\right),$$

or

$$\frac{dV(w(\theta))}{dw_i}\frac{\left(\prod_j f(\theta_j)-\frac{d\lambda_n(\theta_n)}{d\theta_n}\prod_{j\neq n} f(\theta_j)\right)}{\left(\prod_j f(\theta_j)-\frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)\right)}=-\gamma_i.$$

Hence,

$$E_\theta\left[\frac{dV(w(\theta))}{dw_i}\frac{\left(\prod_j f(\theta_j)-\frac{d\lambda_n(\theta_n)}{d\theta_n}\prod_{j\neq n} f(\theta_j)\right)}{\left(\prod_j f(\theta_j)-\frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)\right)}\right]=-\gamma_i,$$

or

$$E^{\mathcal{Q}^i}\left[\frac{dV(w(\theta))}{dw_i}\right]=-\gamma_i,$$

where, as suggested by the notation, $\mathcal{Q}^i$ is distribution associated with the density

$$\frac{\left(\prod_j f(\theta_j)-\frac{d\lambda_n(\theta_n)}{d\theta_n}\prod_{j\neq n} f(\theta_j)\right)}{\left(\prod_j f(\theta_j)-\frac{d\widetilde{\lambda}_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)\right)}\prod_j f(\theta_j)\,^{16}.$$

By the Envelope Theorem,

$$\frac{dV(v)}{dv_i}=-\gamma_i.$$

Hence,

$$\frac{dV(v)}{dv_i}=E^{\mathcal{Q}^i}\left[\frac{dV(w(\theta))}{dw_i}\right]$$

as claimed. ■

Lemma 5 (Spreading of Values) *Assume that $v_i \in (\underline{v},\overline{v})$ for all i. Then, for each i, there is positive probability of both $\frac{dV(w(\theta))}{dw_i} > \frac{dV(v)}{dv_i}$ and $\frac{dV(w(\theta))}{dw_i} < \frac{dV(v)}{dv_i}$.*

Proof. Assume toward a contradiction that, for some i,

$$\frac{dV(w(\theta))}{dw_i}\geq\frac{dV(v)}{dv_i}$$

for almost all θ (the other case is analogous).

Since

$$\frac{dV(v)}{dv_i}=E^{\mathcal{Q}^i}\left[\frac{dV(w(\theta))}{dw_i}\right],$$

[16] A proper normalization of $\widetilde{\lambda}_i(\theta_i)$ can be made so to guarantee that this integrates to 1.

it must be the case that
$$\frac{dV(v)}{dv_i}=\frac{dV(w(\theta))}{dw_i}\text{ for almost all }\theta.$$

Plugging this in the first order conditions for $w_i(.)$, we get

$$\frac{dV(w)}{dw_i}\prod_j f(\theta_j)-\frac{dV(w)}{dw_i}\prod_j f(\theta_j)-\frac{d\lambda_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)-\frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{dV(w)}{dw_i}\prod_{j\neq n} f(\theta_j)=0,$$

(where we have used $\frac{dV(v)}{dv_i}=-\gamma_i$) or

$$-\frac{d\lambda_i(\theta_i)}{d\theta_i}\prod_{j\neq i} f(\theta_j)=\frac{d\lambda_n(\theta_n)}{d\theta_n}\frac{dV(w)}{dw_i}\prod_{j\neq n} f(\theta_j).$$

Dividing both sides by $\prod_j f(\theta_j)$, we have

$$-\frac{d\lambda_i(\theta_i)}{d\theta_i}\left[\frac{1}{f(\theta_i)}\right]=\frac{d\lambda_n(\theta_n)}{d\theta_n}\left[\frac{1}{f(\theta_n)}\right]\frac{dV(w)}{dw_i}.$$

Since the left hand side just depends on θ_i and the right hand side on θ_n, the above can hold for almost all (θ_i,θ_n) only if
$$\frac{d\lambda_i(\theta_i)}{d\theta_i}=\frac{d\lambda_n(\theta_n)}{d\theta_n}=0\text{ for almost all }(\theta_i,\theta_n)\ .$$

Moreover, since for all $s\in\left[\frac{1}{2},1\right]$ and j, $\lambda_j\left(\frac{1}{2}-s\right)=-\lambda_j\left(\frac{1}{2}+s\right)$ – this follows because the problem is symmetric around $\frac{1}{2}$ – , one must have $\lambda_i(\theta_i)=0$, and $\lambda_n(\theta_n)=0$ for all θ_i,θ_n.

Plugging $\lambda_n(\theta_n)=0=\frac{d\lambda_n(\theta_n)}{d\theta_n}$ for all θ_n in the FOC for $w_k(\theta)$ for $k\neq i$, we get

$$\left(\frac{dV(w(\theta))}{dw_k}\prod_j f(\theta_j)+\gamma_k\prod_j f(\theta_j)-\frac{d\lambda_k(\theta_k)}{d\theta_k}\prod_{j\neq k} f(\theta_j)\right)=0.$$

Dividing through by $\prod_j f(\theta_j)$, one gets

$$\frac{dV(w(\theta))}{dw_k}+\gamma_k=\frac{d\lambda_k(\theta_k)}{d\theta_k}\left[\frac{1}{f(\theta_k)}\right]\text{ for almost all }\theta.$$

Since, by the Envelope Theorem, $\gamma_k=-\frac{dV(w)}{dw_k}$, one must have

$$\frac{dV(w(\theta))}{dw_k}-\frac{dV(w)}{dw_k}=\frac{d\lambda_k(\theta_k)}{d\theta_k}\left[\frac{1}{f(\theta_k)}\right]\text{ for almost all }\theta.$$

Now taking expectations over both sides of the above equality, we have

$$E^{\mathcal{Q}^k}\left[\frac{dV(w(\theta))}{dw_k}-\frac{dV(w)}{dw_k}\right]=E^{\mathcal{Q}^k}\left[\frac{d\lambda_k(\theta_k)}{d\theta_k}\left[\frac{1}{f(\theta_k)}\right]\right]$$

Using the martingale property, it must be the case that

$$E^{\mathcal{Q}^k}\left[\frac{d\lambda_k(\theta_k)}{d\theta_k}\left[\frac{1}{f(\theta_k)}\right]\right]=0.$$

Since $\frac{d\lambda_k(\theta_k)}{d\theta_k}$ does not change sign, it must be the case that $\frac{d\lambda_k(\theta_k)}{d\theta_k} = 0$ for almost all θ_k, which implies – as $\lambda_k\left(\frac{1}{2} - s\right) = -\lambda_k\left(\frac{1}{2} + s\right)$ for all s, since the problem is symmetric around $\frac{1}{2}$ – that $\lambda_k(\theta_k) = 0$ for all k, and for all θ_k.

Plugging this in the FOC for $w_k(\theta)$,

$$\frac{dV(w(\theta))}{dw_k} - \frac{dV(w)}{dw_k} = 0 \text{ for almost all } \theta \text{ and for all } k,$$

which implies, given the strict concavity of $V(.)$, that, for all k, $w_k(\theta) = w_k$ for almost all θ

Using the fact that, for all i and for all θ_i, $\frac{d\lambda_i(\theta_i)}{d\theta_i} = \lambda_i(\theta_i) = 0$, the FOC for a reads

$$\frac{\partial u(a(\theta), \theta_n)}{\partial a} + \sum_{i=1}^{n-1} \gamma_i \frac{\partial u(a(\theta), \theta_i)}{\partial a} = 0.$$

It is easy to see that the $a(.)$ implicitly defined by the above equation is not IC when continuation values are constant, unless $\gamma_i = 0$ for all i, or $\gamma_i = \infty$ for some i; that is, unless the dictatorship holds. Dictatorship, however, contradicts $w_j \in (\underline{v}, \overline{v})$ for all j ■

We are now able to show:

Theorem 1 *The provision of intertemporal incentives necessarily leads to a dictatorial mechanism: In the limit as $t \to \infty$, either v_j converges to $\overline{v}$ almost surely, $j = 1, ..., n-1$, or $V(v)$ converge to $\overline{v}$ almost surely.*

Proof. Since $\frac{dV(v)}{dv_i}$ is a non-positive martingale, by Dobb's convergence Theorem (see Dobb (1953)), it converges almost surely to some random variable, R_i. Next we show by contradiction that R_i cannot have any positive likelihood for values in $(0, \infty)$. Hence, all the probability is concentrated where $R_i = 0$ or $R_i = -\infty$. Since this must be true for all i, that implies that either $V(v)$ goes to $\overline{v}$, or v_j converges to $\overline{v}$ for some i, i.e., one of the players becomes a dictator in the limit.

In search of a contradiction, suppose there existed, for some i, a positive probability of finding a path $\frac{dV(v)}{dv_i}$ with the property that $lim_{t\to\infty} \frac{dV(v)}{dv_i} = C$, where $0 < C < \infty$. Since $\frac{dV(v)}{dv_i}$ is continuous for any $v \in (\underline{v}, \overline{v})$, the sequence v_t converges. Denote its limit by $lim_{t\to\infty}\ v_t = v' \in (\underline{v}, \overline{v})$. Let $W(w, \theta)$ denote the next period's continuation value given the current promised value w and reported state θ. For w_t to converge it must be that $W(w', \theta) = w'$ for all θ. This however contradicts Lemma (5). ■

4.2 The Approximate Efficiency Result

Proof of Approximate Efficiency. We define the ex-ante efficient allocation as

$$a^*(\theta) = \arg\max_a E_\theta\left[\sum_{i=1}^n u(a, \theta_i)\right].$$

We prove that, for any $\epsilon > 0$, there exists $\overline{\delta} \in (0,1)$ such that for $\delta > \overline{\delta}$, the sum of the players' equilibrium payoffs is within ϵ of the payoff associated with

$$v^{FB} = E_\theta\left[\sum_{i=1}^n u(a^*(\theta), \theta_i)\right].$$

We do so by constructing continuation values that replicate as closely as possible the expected payments of the expected externality mechanism proposed by Arrow (1979), and d'Aspremont and Gerard-Varet (1979), that guarantee efficiency in a standard (static) Mechanism Design problem.

Define

$$\xi_i(\theta_i) = E_{\theta_{-i}}\left[\sum_{j\neq i} u(a^*(\theta), \theta_j)\right],$$

and consider, for $i = 1, ..., n$, the following candidates for continuation values

$$v_i(\theta) = \left(\begin{array}{c} \left(\frac{1-\delta}{\delta}\right)\left[\left(\sum_{j\neq i} E_{\theta_{-i}}[u(a^*(\theta), \theta_j)]\right) - E_{\theta_{-i}}\left[\frac{1}{n-1}\left(\sum_{j\neq i}\xi_j(\theta_j)\right)\right]\right] \\ +E_\theta[u(a^*(\theta), \theta_i)] - \frac{\kappa(\delta)}{\delta} \end{array} \right),$$

where $\kappa(\delta)$ is given by

$$\kappa(\delta) = (1-\delta)\left[\sum_{j\neq i}\max_\theta E_{\theta_{-i}}[u(a^*(\theta), \theta_j)] - \frac{1}{n-1}\left[\sum_{j\neq i}\sum_{k\neq j}\min_\theta E_{\theta_i}[u(a^*(\theta), \theta_j)]\right]\right] \equiv (1-\delta)\overline{d},$$

and $\overline{d}$ is a finite number.

Note that $\kappa(\delta)$ is strictly positive and just depends on δ. It is chosen so to guarantee that, for all θ, $\{v_i(\theta)\}_i$ are feasible values. Note, moreover, that

$$\begin{aligned} & -\frac{1}{n-1}E_{\theta_{-i}}\sum_{i=1}^{n}\sum_{j\neq i}\xi_j(\theta_j) \\ = & -\frac{1}{n-1}\sum_{i=1}^{n}(n-1)\xi_i(\theta_i) \\ = & -\sum_{i=1}^{n}\xi_i(\theta_i) = -\sum_{i=1}^{n}\sum_{j\neq i}E_{\theta_{-i}}[u(a^*(\theta), \theta_j)]. \end{aligned}$$

Hence, upon inducing truthfulness from the players – so that $a^*(\theta)$ can be implemented in the first period in an Incentive Compatible way –,

$$\sum_{i=1}^{n} v_i(\theta) = \left(\sum_i E_\theta[u(a^*(\theta), \theta_i)]\right) - \frac{n\kappa(\delta)}{\delta}$$

so that the sum of the players' expected payoffs when these continuation values are used is $v^{FB} - \frac{n\kappa(\delta)}{\delta}$.

We now proceed by showing that, with these continuation values, one can implement $a^*(\theta)$ in an incentive compatible way. We then show that we can make $\kappa(\delta)$ arbitrarily small as $\delta \to 1$.

Note that, if players other than i are being truthful, player i's problem, if a^* is implemented and he faces $v_i(\theta)$ as a continuation value, is

$$\max_{\widehat{\theta}_i}(1-\delta)E_{\theta_{-i}}\left(u\left(a^*\left(\widehat{\theta}_i, \theta_{-i}\right), \theta_i\right)\right) + \delta E_{\theta_{-i}}\left[v_i(\widehat{\theta}_i, \theta_{-i})\right]$$

which has the same solution as the one associated with the program[17]

$$\max_{\widehat{\theta}_i}(1-\delta)E_{\theta_{-i}}\left(u\left(a^*\left(\widehat{\theta}_i, \theta_{-i}\right), \theta_i\right)\right) + (1-\delta)\sum_{j\neq i}\left[E_{\theta_{-i}}\left[u\left(a^*\left(\widehat{\theta}_i, \theta_{-i}\right), \theta_{-i}\right)\right]\right].$$

[17] All other terms do not not affect incentives.

Since

$$a^*(\theta) = \arg\max_a \sum_{i=1}^{n} u\left(a, \theta_i\right),$$

the announcement $\widehat{\theta}_i = \theta_i$ is optimal.

Now, pick $\epsilon > 0$. Consider the $\overline{\delta}$ that solves

$$\frac{n\kappa\left(\overline{\delta}\right)}{\overline{\delta}} = \frac{n\left(1-\overline{\delta}\right)\overline{d}}{\overline{\delta}} = \epsilon.$$

It is easy to see that

$$\overline{\delta} = \frac{n\overline{d}}{n\overline{d}+\epsilon} < 1.$$

Moreover, for $\delta > \overline{\delta}$, the sum of the players' equilibrium payoff is within ϵ of v^{FB}, when one implements $a^*\left(\theta\right)$ with continuation values $\left\{v_i\left(\theta\right)\right\}_{i=1}^{n}$. Since the optimal contract can not do worse than $\left\{a^*\left(\theta\right), v_i\left(\theta\right)\right\}_i$, the result follows. ■