DISCUSSION PAPER SERIES

No. 3303

ON ROBUST CONSTITUTION DESIGN

Emmanuelle Auriol and Robert Gary-Bobo

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP3303.asp

ISSN 0265-8003

ON ROBUST CONSTITUTION DESIGN

Emmanuelle Auriol, Université des Sciences Sociales de Toulouse
Robert Gary-Bobo, Université de Cergy-Pontoise and CEPR

Discussion Paper No. 3303
April 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3303

April 2002

ABSTRACT

On Robust Constitution Design*

We study a class of representation mechanisms, based on reports made by a random subset of agents, called representatives, in a collective choice problem with quasi-linear utilities. We do not assume the existence of a common prior probability describing the distribution of preference types. In addition, there is no benevolent planner. An individual who cannot be assumed impartial, a self-interested executive, will carry out decisions. These assumptions impose new constraints on Mechanism Design. A robust mechanism is defined as maximizing expected welfare under a vague prior probability distribution, and over a set of mechanisms which is at the same time immune from opportunistic manipulations by the executive, and compatible with truthful revelation of preferences by representatives. Robust mechanisms are characterized and their existence is shown. Sampling Groves mechanisms are shown to be robust.

JEL Classification: D70, D80 and H00
Keywords: collective choice, incomplete information, mechanism design and representative democracy

Emmanuelle Auriol
Institut D'Economie Industrielle
Université des Sciences Sociales
21 Allée de Brienne
31000 Toulouse
FRANCE
Tel: (33 5) 61 12 86 2
Email: eauriol@cict.fr

Robert Gary-Bobo
Université de Cergy-Pontoise
THEMA
33, boulevard du Port
95011 Cergy-Pontoise
Cedex
FRANCE
Tel: (33 1) 3425 6062
Fax: (33 1) 3425 6233
Email: robert.gary-bobo@eco.u-cergy.fr

* We thank Nabil Al Najjar, Jacques Drèze, Jean-Pierre Florens, Françoise Forges, Roger Guesnerie, Jean-Jacques Laffont, John Ledyard, Philippe Mongin and Thomas Palfrey for their help and comments, as well as seminar audiences in Dublin, Toulouse, Stockholm, University of Wisconsin, Madison, University of Illinois at Urbana-Champaign, Northwestern University, Boston College, Caltech, UCLA and CEPR (Public Policy Program) for useful remarks.

Submitted 13 March 2002

1 Introduction

Practitioners and applied economists express some dissatisfaction with the literature on Mechanism Design and the implementation of collective-decision rules, in particular when they discuss applications to public decision-making. There are at least two lines of attack. One is simplicity, or *realism*: it is commonly argued that the mechanisms exhibited by pure theory are often too complex or contrived to be reasonably understood by economic agents and implemented in practice. The other one is *robustness*: theoretical mechanisms would be too fragile, too dependent on some form of fine tuning to resist mistakes, collective manipulations of various kinds, or deviations from rationality made by agents. These criticisms have been recognized by students of Mechanism Design since a long time. However, because of the genuine difficulty of the task, the concepts of simplicity and robustness have not yet acquired well-defined meanings in the theoretical literature. In the present contribution, we focus on the notion of *robust mechanism*, and put the emphasis on one aspect of robustness.

There are various approaches to robustness in the literature. Some authors have considered the ability of agents to communicate before the (collective-decision) game is played, and, if communication cannot be controlled, have studied the problem of *communication-proof* implementation[4]. Other approaches have investigated the resistance of mechanisms with respect to renegotiation phenomena, when, in essence, the number of stages or inter-agent bargaining cannot be fully controlled. These considerations lead to the concept of *durable* or *renegotiation-proof* implementation[5]. The possibility of collusion between agents is of course an important aspect of the problem, and there is a recent line of research on *collusion-proofness*[6]. Other avenues can be explored, radically questioning the game-theoretic concepts themselves, by considering the idea of robustness with respect to imperfections of rationality, knowledge, and common knowledge. The notion that the social planner knows the relevant characteristics of a prior probability distribution describing the population of agents, or that this prior is common to the agents, has also been questioned. Weakening these assumptions can lead to dramatic changes in the results obtained[7].

The definition of robustness that we consider below is inspired by the specific problems of public decision-making, and by the idea of an application to the analysis of political institutions in a representative democracy. We propose a normative theory of collective decisions under the assumption that there is no benevolent planner on which to rely, and without assuming that probabilistic beliefs are common, and common knowledge. In other words, we assume that all decisions are made by "real" (as opposed to fictive) individuals, that these "real" individuals are endowed with preferences, thus pursuing some private in-

[4]See Palfrey and Srivastava (1991), and for a general discussion, Palfrey and Srivastava (1993).

[5]See, among many other contributions, Holmstrom and Myerson (1983), Maskin and Moore (1999).

[6]See, again among other contributions, Laffont and Martimort (2000).

[7]A notion of *robust trading mechanism* is defined in Hagerty and Rogerson (1987); it combines the requirement of dominant strategies with *ex post* rationality.

terest, and we do not assume that the structure of preferences is well known by agents, in the form of a common prior. The spirit of the present contribution is to remove the benevolence assumption in normative analysis[8], while simultaneously removing the Harsanyian assumption that a description of society, taking the form of a prior probability distribution, is common knowledge. In contrast, in our theory, the purpose of public decision mechanisms will be to produce relevant, but otherwise non-existent information at some cost. Due to the absence of benevolence, the way information is processed and the mechanism itself should resist opportunistic manipulations. This context imposes additional constraints on mechanism design, on top of the classic incentive compatibility constraints.

In the following, we consider an economy with a finite number of agents. Some public decision must be made, such as producing a public good. Utilities are quasi-linear, allowing for possibilities of compensation by means of money transfers. The agents' private valuations for the public decision are unknown. The distribution of valuations in the economy can be described by a probability measure which is not assumed to be common knowledge. Some information on preferences must be produced by the institutions, and information production is costly. To capture this, we assume that individuals need to pay some fixed cost to be able to transmit their valuation functions[9].

We consider a family of mechanisms, called *representation mechanisms*, in which some agents are drawn at random in the population and pay the fixed cost to uncover and to report their preferences. These individuals are called *representatives*. Society is then divided into three subsets: (i) the representatives, who transmit information about preferences, (ii) the *passive citizens* who are simple taxpayers, and (iii), the *executive*, which is the agent in charge of executing public decisions. No agent is benevolent. The fact that representatives are not benevolent imposes incentive compatibility constraints, to ensure truthful reporting of preferences; the fact that the executive is not benevolent will impose an additional requirement, called *political robustness*.

The classic framework of Mechanism Design is a particular case of the above described setting in which, (i), all individuals can transmit at zero cost their personal characteristic or type, (ii), all agents are representatives, and (iii), there implicitly exists a fully reliable and benevolent executive body.

Now, given that our goal is to propose a normative analysis, we will assume that the constitution, a set of rules obeyed by all agents, is designed and written

[8] In *The Making of Economic Policy*, A. Dixit (1996, p. 8) writes, "As a crude but effective caricature, one can say that normative policy analysis began by supposing that the policy was made by an omnipotent, omniscient, and benevolent dictator. The work on second best removed the omnipotence. That on information removed the omniscience. However, the assumption of benevolence and dictatorship have remained unaffected by all these improvements in our understanding of the limits on instruments and information.(...) The normative approach continues to view policy-making as a purely *technical* problem."

[9] Reporting a valuation function for a public decision generally involves a certain amount of work, like reading files, listening to experts, etc., to understand the issue at hand and its consequences on the welfare of individuals. Participating in representative institutions, where bargaining and voting take place, is a time-consuming activity, and time has a non-negligible opportunity cost.

down by fictitious individuals called the *Founding Fathers,* behind the veil of ignorance. In contrast to the "real" agents, the Founding Fathers are assumed benevolent, utilitarian and Bayesian welfare maximizers; but they don't know the probability distribution describing the agents' preferences, and they know that they don't know it. In addition, the Founding Fathers know that nobody will be impartial or benevolent. In spite of all this, they want their institutions to function equally well in all possible future societies, the characteristics of which are unknown to them. Their most difficult problem can therefore be called an *informational robustness* problem. The Founding Fathers feel a tension between their desire to limit opportunistic manipulation possibilities in the course of public decision-making, which imposes a form of rigidity of rules, and at the same time, the need to adapt institutions to changing patterns of the citizens' preferences, which requires flexibility.

We define as *robust*, a mechanism which is immune from manipulations (i.e., politically robust), and at the same time, maximizes expected welfare under a vague prior probability distribution representing the Founding Father's ignorance (i.e, is informationally robust). We provide a characterization of politically robust mechanisms, and show that a form of Groves mechanism[10], applied to the subset of representatives, is a fully robust solution, that is, both politically and informationally robust. We finally illustrate the theory by means of an example.

In the following, Section 2 describes the model and states the assumptions. Section 3 is devoted to the notions of robustness. Section 4 contains the statement of our formal results. Theorem 1 provides a characterization of politically robust mechanisms. Theorem 2 shows the existence of robust mechanisms. Section 5 is devoted to an illustrative example.

2 The Model

We consider an economy composed of $N+1$ agents, indexed by $i = 0, 1, ..., N$. A public decision denoted x should be chosen in a set X containing l elements, i.e., $|X| = l$. Agent i will pay a tax denoted t_i , which must be interpreted as a subsidy if it is negative. Each agent's utility function, denoted u_i, depends on the public decision and the tax.

Assumption 1. (Quasi-linearity) Utilities are quasi-linear, and defined as $u_i(x, t_i) = v_i(x) - t_i$, where $v_i : X \longrightarrow \Re$, is a private valuation function.

These valuation functions are formally vectors $v_i = (v_i(x))_{x \in X}$ belonging to $\Re^l$, and each of these vectors can be viewed as a drawing in some unknown probability distribution P, with support $V \subset \Re^l$.

Assumption 2. (Stochastic Independence) For all i, the v_i are independent drawings from the same distribution P. The distribution P has a well-defined mean.

[10]On Groves and Clarke-Groves mechanisms, see Clarke (1971), Groves (1973), Green and Laffont (1977), Holmström (1979), Moulin (1988).

Let E_P denote the expectation operator, taken with respect to P. For all $x \in X$, we denote $E_P(v_i(x)) = \overline{v}_P(x)$, the mean valuation of alternative x in the population described by P. Agent i cannot reveal or transmit a type v_i, unless he (or she) pays a fixed cost[11] F. Agent i is endowed with a probabilistic belief denoted P_i on the set of possible preferences, describing what he or she believes about the preferences of other agents. These probabilities can be viewed as purely subjective.

Our goal is now to provide a formal description of a constitution in this economy. Because of the representation cost F, we focus on representation mechanisms. To be more precise, in a representation mechanism, a subset of agents, of size $n \leq N$, represents the entire society. Direct democracy (i.e., $n = N$) is a particular case. By definition, the *representatives* are the agents labelled i, with $i = 1, ..., n$. They, (i), pay a fixed cost F, and (ii), transmit a message $\widehat{v}_i \in V$ to the mechanism. It cannot be taken for granted that reports are truthful, and it follows that there will be a revelation problem. The inputs of the public decision rule are the representatives' reports, a vector denoted $\widehat{v} = (\widehat{v}_1, ..., \widehat{v}_n) \in V^n$, and some other inputs, called "parameters", and denoted λ. Parameters are chosen in some set Λ. Without any loss of generality, we can assume that Λ has a product space structure, with a finite number of factors Λ_b indexed by index b in some index set B, that is, $\Lambda = \prod_{b \in B} \Lambda_b$, and $\lambda = (\lambda_1, ..., \lambda_{|B|})$.

Definition 1. *A representation mechanism is an array* (f, t, Λ^0)*, where* $\Lambda^0 \subset \Lambda$ *is an admissible subset of parameters,* f *is the public decision rule, and* $t = (t_0, t_1, ..., t_N)$ *is a vector of taxes. The vector* (f, t) *maps inputs of the decision rule into the set of public decisions and tax allocations; formally,*

(i) $(f, t) : V^n \times \Lambda^0 \longrightarrow X \times \Re^{N+1}$

$$(\widehat{v}, \lambda) \mapsto (f(\widehat{v}, \lambda), t(\widehat{v}, \lambda)),$$

and,

(ii) $\sum_{i=0}^{i=N} t_i(\widehat{v}, \lambda) \equiv 0.$

Point (ii) in Definition 1 is of course the public budget constraint. It follows that budget imbalance is not permitted by the constitutions considered here. For the sake of completeness, we can formally define a *constitution* as the specification of a representation mechanism for each n, with $0 \leq n \leq N$[12].

In the economy described here, public decisions must be executed by at least one agent, called the executive. The executive is conventionally indexed by $i = 0$. We assume that a single individual is enough to execute decisions. Becoming an executive has an opportunity cost $F_0 \neq 0$. All messages $\widehat{v}$ are transmitted to agent 0, in charge of execution. He (or she) has the potentially

[11] The fixed cost F can be viewed as the opportunity cost of becoming a representative. It can also be viewed abstractly as the cost of transmitting a message about preferences. Another possible interpretation is that the agents do not know their own preferences, unless they exert some costly effort, and that at a fixed cost F, they learn their valuation function completely. The model is then as if we had assumed the existence of a simplified *information acquisition* technology.

[12] But n itself is not fixed in the constitution.

important power to set the value of the parameters λ. By definition, a parameter λ is said to be *specified in the constitution* if $\Lambda^0 = \{\lambda\}$.

Assumption 3. (Subsidiarity of the executive) *Any input of the public decision rule which is not provided by the representatives or specified in the constitution is chosen by the executive.*

This assumption says that, a parameter value being needed to choose $x = f(\widehat{v}, \lambda)$, either λ is carved in constitutional marble, that is, $\Lambda^0 = \{\lambda\}$, or agent 0 chooses $\lambda \in \Lambda^0$. This aims at capturing the idea of bureaucratic manipulation power. We assume that the executive has unknown preferences v_0, drawn at random from the probability distribution P. Being an agent, the executive will exploit any possibilities to distort the public decision in the direction of his or her private preferences, if given the possibility to do so. He or she would choose the value of $\lambda \in \Lambda^0$ so as to maximize $v_0(f(\widehat{v}, \lambda)) - t_0(\widehat{v}, \lambda)$. We add the following simple assumption,

Assumption 4. (Separation of Powers) *The executive cannot be a representative.*

The model should now describe the way of choosing representatives. Our goal is to construct a simple and stylized model of an ideal representative democracy. We therefore assume the existence of a perfect, costless, and non-manipulable representation technology: the n representatives are a random sample of preferences.

Assumption 5. (Perfect random representation) *The n representatives are independent random drawings in the probability distribution P.*

This assumption says that there exists a non-manipulable way of creating an unbiased sample of the agents' preferences. A "perfect" electoral system would do this job, and our assumption has the merit of black-boxing the electoral process completely. This is of course a strong assumption. In the real world, most representation systems are presumably biased in the sense that the subset of representatives is not a reduced mirror image of the people's preferences. Our focus will not be the problem of the choice of representatives, even if this problem is important, but to reveal and aggregate relevant private information in a robust way, while keeping some democratic principles at work. Our goal being to construct a normative theory of collective decision mechanisms, the idea of unbiased random representation captures an important democratic ideal. In this perspective, Assumption 5 provides a desirable simplification, and we will adhere to this naive view of "perfect representation".

We add the following assumption.

Assumption 6. (Weak equality) *Indistinguishable individuals are treated equally by the constitution.*

Representative i's tax schedule t_i will be used to create revelation incentives, but individuals indexed by $i = 0$ and $i = n + 1, ..., N$, are indistinguishable. More precisely, in the economy under study, an agent can only be distinguished from the other by his (her) preferences. These preferences being unobserved, there is no basis for differential tax treatment of those whose preferences will never be revealed. It immediately follows from this that the same tax t_0 will be paid by any agent who does not belong to the subset of representatives. The budget balance constraint thus writes,

$$(N+1-n)t_0 + \sum_{i=1}^{n} t_i = 0. \tag{1}$$

To sum up, society has been partitioned into 3 subsets; the executive, indexed 0, the representatives, indexed from 1 to n, and the passive citizens, indexed from $n+1$ to N. The representatives, who pay the fixed cost F, are the only citizens which can transmit information about their preferences. Representatives will be subject to revelation incentives, while passive citizens (and agent 0) balance the budget. The fixed cost explains why direct democracy is not necessarily optimal.

3 Notion of Robustness

To give formal content to the idea of an impartial and benevolent point of view on society and its constitutional rules, we assume the existence of fictitious agents called the Founding Fathers. The Founding Fathers (hereafter the FF) choose the constitution. They are assumed utilitarian, Bayesian, and benevolent. In addition, they know that they don't know the true distribution of preferences P, and they know that, once the set of rules embodied in the constitution will be put into use, there will not exist a single impartial or benevolent individual to carry out public decisions. The FF want their constitution to be the best among the politically robust ones. Optimality here, refers to maximization of an expected sum of utilities, and political robustness, informally, means that the FF want their constitution to resist manipulations. These requirements are problematic without the help of the common knowledge assumption, since the true distribution of preferences is unknown, and the agents' probabilistic beliefs P_i are unknown (i.e., information is completely decentralized).

3.1 The Founding Father's beliefs

There are two classic ways of modelling a problem of decision under ignorance. One is to use a non-probabilistic representation of ignorance such as the maximin principle, that is, in the present context, maximize the minimum of expected welfare, where the minimum is taken over some set of possible probability distributions representing society; the other is to remain in the realm of Bayesian decision theory and to choose mechanisms which are optimal against some vague or uninformative prior. We have chosen to follow the latter route. As is well known from the Bayesian statistical literature, the recourse to non-informative priors can be a delicate matter. To this end, the Founding Fathers' system of beliefs will be assumed to belong to a family containing non-informative priors as limit points.

Let Q denote the FF's system of probability judgments on preferences, it is a probability distribution on V^{N+1}. The vector $(v_0, ..., v_N)$ is a sequence of random valuation vectors, each belonging to $V \subset \Re^l$. If we assume that the sequence $(v_0, ..., v_N)$ is *exchangeable*, i.e., that the probability of $(v_0, ..., v_N)$ is the same under Q as that of any permutation $(v_{\pi(0)}, ..., v_{\pi(N)})$, then, extensions of de Finetti's Theorem yield a representation for Q of the following form: there exists a probability measure M on the space $\mathcal{P}$ of all distributions P on V such that,

$$Q(v_0, ..., v_N) = \int_{\mathcal{P}} \prod_{i=0}^{N} P(v_i) dM(P). \tag{2}$$

In other words, the v_i are independent, conditionally on the knowledge of a distribution P. If P is unknown, then, the FF's probabilistic beliefs can be described with the help of a prior M on the set of possible distributions P, and the beliefs Q are a mixture of *a priori* possible Ps. The model described by (2) is too general to be easily tractable in our economic framework. But if we were ready to assume a little bit more about the FF's beliefs, we would get a parameterized representation with a nice structure (see Bernardo and Smith (1994), Chap. 4). To avoid useless sophistication, we directly assume that the FF have a multivariate normal and fully exchangeable view of each valuation vector v_i.

Assumption 7. (Normality of the Founding Fathers' Conditional Beliefs.) *Let Q denote a possible probabilistic belief of the Founding Fathers. There exists a prior distribution M on the parameter space A, such that*

$$Q(v_0, ..., v_N) = \int_{A} \prod_{i=0}^{N} \mathcal{G}(v_i; \alpha) dM(\alpha), \tag{3}$$

where $\mathcal{G}(v_i; \alpha)$ is a multivariate normal distribution with parameter $\alpha \in A$. Parameter α writes $\alpha = (\mu, T)$, where $\mu \in \Re^l$ is a mean vector and T is an $l \times l$ positive definite precision matrix.

In the eyes of the FF, individual valuations v_i are *conditionally independent.* The parameter space A contains all pairs (μ, T) with a vector in $\Re^l$ and a full rank, symmetric, positive definite matrix, which has $l(l+1)/2$ distinct elements lying above or on its main diagonal[13]. Given that our goal is to provide a convenient representation of vague prior knowledge, Assumption 7 is not very restrictive. The choice of the prior distribution M is essentially unrestricted, and normal distributions put probability weight on every region of the space $V = \Re^l$.

[13] The precision matrix of a multivariate normal distribution is defined as the inverse of its variance-covariance matrix.

We define a prior probability distribution M on the set A as *diffuse* if every open set in A has a positive probability. Consider now the marginal distribution $\overline{M}$ of the random parameter μ. The precision matrix of $\overline{M}$ is the inverse of its covariance matrix. Let $\overline{T}$ denote this matrix, i.e., $\overline{T} = [Cov_{\overline{M}}(\mu)]^{-1}$. Intuitively, a prior distribution of μ is vague if its precision matrix is close to zero. In the following, we will consider sequences of diffuse prior distributions (M_r) with associated sequences of beliefs (Q_r) and precision matrices $(\overline{T}_r)$. The FFs' prior will be said to become *vague* when the associated sequence of precision matrices $(\overline{T}_r)$ vanishes, i.e., $\overline{T}_r \to \mathbf{0}$, as $r \to +\infty$, where $\mathbf{0}$ is the matrix each of whose elements is 0.

We now turn to a formal definition of a robust mechanism. Our robustness notion is twofold. It has a political, and an informational aspect.

3.2 Political robustness

Political robustness, which entails the idea of immunity from political and bureaucratic manipulations, encompasses more than the usual incentives constraints. In a nutshell, the notion captures all the consequences of not having an omniscient, omnipotent and benevolent planner at hand. Consider first the revelation problem.

The representation mechanism should induce truthful revelation of the representatives' private information. The revelation principle applies to the game restricted to the n representatives, once they have paid the fixed cost F. We therefore concentrate on direct revelation mechanisms. Let E_i denote the expectation taken with respect to representative i's subjective probability distribution P_i. Let $U_i(v;\lambda) = v_i [f(v;\lambda)] - t_i [v;\lambda]$ be representative i's utility under the mechanism. By definition, the representation mechanism (f,t) is revealing (in the Bayesian sense), if for all $i = 1, ..., n$, and all $v_i \in V$,

$$E_i [U_i(v;\lambda) \mid v_i] \geq E_i [U_i(\widehat{v}_i, v_{-i};\lambda) \mid v_i], \text{ for all } \widehat{v}_i \in V, \tag{4}$$

where the usual notation $v = (v_i, v_{-i})$ is employed. But, as soon as we have written this definition, we see that the notion of a revealing mechanism (in the Bayesian sense) cannot be very useful, because such a mechanism is typically parameterized by the agents' subjective beliefs P_i, which are unknown to the FF, and to the executive. It then seems that the only useful mechanisms are those which are revealing, whichever the agents' beliefs. A representation mechanism is then defined as *universally revealing*, if it is revealing (in the Bayesian sense) for all vector of subjective beliefs $(P_0, ..., P_N)$. In particular, this should be true for beliefs concentrated on a single point v_{-i} for each i. Rewriting (4) under this assumption, we get the equivalent condition, for all $i = 1, ..., n$, all $v_i \in V$, and all $v_{-i} \in V^{n-1}$,

$$U_i(v;\lambda) \geq U_i(\widehat{v}_i, v_{-i};\lambda), \text{ for all } \widehat{v}_i \in V, \tag{5}$$

which is nothing but the definition of a dominant strategy revelation mechanism. Our political robustness notion should therefore embody the requirement that (f,t) is revealing in dominant strategies.

This step can be given a formal justification. We rephrase a result of Ledyard (1978) as saying that there are no Bayesian incentive compatible *and non-parametric*[14] mechanisms which are not at the same time revealing in dominant strategies. An incentive compatible mechanism is non-parametric, if it doesn't depend on parameters characterizing a particular economy, such as a prior probability distribution or its moments. If we insist on this, then, nothing can be gained by replacing the requirement of dominant strategy equilibrium with the weaker notion of Bayesian Nash equilibrium. It will be reassuring to discover below that, in our particular setting, the restriction to dominant strategies has no additional welfare cost, given that we rely on a representation mechanism.

On top of this, we define a representation mechanism as *immune from executive opportunism*, if it simply does not leave any margin of manoeuvre to the executive, all parameters being specified in the constitution, i.e., $\Lambda^0 = \{\lambda\}$. In a more general setting, complete immunity from executive opportunism could be socially costly, but in the present model, the hands of the executive can be easily and costlessly tied by fixing mechanism parameters in the constitution. This notion has consequences, as will appear below, since it conflicts with the requirement of informational robustness, and forbids the use of some otherwise "natural" mechanisms.

Finally, we add a technical requirement. A representation mechanism is said to satisfy the condition of *non-imposition*, if for all $x \in X$, there exists a $\widehat{v} \in V^n$ such that $x = f(\widehat{v})$. This condition is almost innocuous, for if some x will never be chosen, then it may be excluded from the set X[15]. We conclude this discussion with a definition.

Definition 2. *A representation mechanism* (f, t, Λ^0) *is* politically robust *(hereafter PR) if it is universally revealing, immune from executive opportunism and satisfies the condition of non-imposition.*

3.3 Informational robustness

The Founding Fathers, willing to design a constitution that can adapt to the largest possible set of societies, as represented by the true, but unknown, probability distribution P, do not want the representation mechanism to be taylored too closely to a particular society, for it would not resist changes in the distribution of preferences. Indeed, if many citizens are consistently hurt by collective decisions, they will fight the system in one way or another. Informational robustness is a form of stability requirement.

The chosen mechanism should, as much as possible, be independent of a particular distribution of preferences, but this task might be impossible to accomplish if the mechanism must at the same time satisfy optimality properties. Intuitively, parameters, including prior information on the distribution of preferences, should be fixed in the constitution, to avoid possible manipulation by

[14] In the sense of Hurwicz (1972).

[15] The condition of non-imposition has some power, however, if it is assumed that there are at least three elements in X.

the executive, but, at the same time, the constitution should be flexible enough to adapt (and therefore remain optimal) in a large class of possible societies. To illustrate this tension, let us assume for a moment that the distribution P is common knowledge, and known to the FF.

The welfare function is defined as,

$$W_n(f) = -(nF + F_0) + \sum_{i=0}^{N} v_i(f). \tag{6}$$

It is the difference between the utilitarian measure of welfare, and the costs of constitutional organisation, including the variable social cost nF of the representatives, and the fixed cost of the executive. Given the knowledge of P, the best that the FF can do is to choose (f, t) so as to maximize $E_P\left[W_n\left(f\right)\right]$. Now, remark that,

$$E_P\left[W_n\left(f\right)\right] = E_P\left[E_P(W_n\left(f\right) \mid \widehat{v})\right].$$

Given the assumed independence of valuations (Assumption 2), and dropping the parameter λ to lighten notation, we get,

$$E_P\left[W_n\left(f\right)\right] = -(nF + F_0) + E_P\left\{\sum_{i=1}^{n} \widehat{v}_i\left[f(\widehat{v})\right] + (N + 1 - n)\overline{v}_P\left[f(\widehat{v})\right]\right\}, \tag{7}$$

(recall that $\overline{v}_P(x) = E_P\left[v_i\left(x\right) \mid x\right]$ for all $x \in X$). It follows from (7) that the first-best optimal decision rule, denoted $f_P(\widehat{v})$, solves the problem,

$$\max_{x \in X}\left\{\sum_{i=1}^{n} \widehat{v}_i\left(x\right) + (N + 1 - n)\overline{v}_P\left(x\right)\right\},$$

for all $\widehat{v} \in V^n$, that is, maximizes the term between curly brackets in (7) pointwise. With our assumptions, and in particular Assumption 3, if the executive was in charge of, and free to choose $\overline{v}_P$, he or she could simply set

$$\overline{v}_P\left(x\right) = \frac{-1}{N + 1 - n}\left(\sum_{i=1}^{n} \widehat{v}_i\left(x\right) - v_0(x) + t_0\left(\widehat{v}\right)\right),$$

implying that the rule f_P maximizes $v_0 - t_0$, i.e., the executive's private utility function. This is why a constitutional specification of the parameter $\overline{v}_P$ is needed. This mechanism becomes immune to executive opportunism if $\Lambda = V$ and $\Lambda^0 = \{\overline{v}_P\}$, i.e., if the mean valuation $\overline{v}_P$ is specified in the constitution. But then, the representation mechanism trivially depends on a particular preference distribution P, which contradicts the requirement of informational robustness. Unless the FF agree on this particular distribution, or, in other words, unless their prior M concentrates all probability on P, the "usual" first best rule f_P

cannot be part of a robust constitution[16].

A reasonable solution to this problem is to define a mechanism as *informationally robust*, if it maximizes expected welfare, when the expectation is taken with respect to some vague or uninformative prior M. A representation mechanism is then defined as simply *robust*, if it is informationally robust, subject to the political robustness constraint. To do this rigorously, we must take care of the well-known difficulties associated with improper prior distributions in Bayesian statistics. Recall that A is the set of pairs (μ, T); that $\overline{M}$ is the marginal distribution of μ and that $\overline{T}$ denotes the precision matrix of $\overline{M}$, i.e., $\overline{T} = [Cov_{\overline{M}}(\mu)]^{-1}$.

Definition 3. (Robust Representation Mechanisms). *A representation mechanism* (f^*, t^*, Λ^{0*}) *is* robust*, if it is PR, and if there exists a sequence of* diffuse *prior distributions* (M_r)*, with an associated sequence of probabilistic beliefs* (Q_r) *(derived from* (M_r) *using expression (3) above*), *such that,*
(a) the associated sequence of precision matrices $\left(\overline{T}_r\right)$ *vanishes, i.e.,* $\overline{T}_r \to \mathbf{0}$, as $r \to +\infty$*;*
(b) for each integer r*, there exists a PR representation mechanism* (f_r, t_r, Λ_r) *such that* (f_r, t_r) *maximizes the expected welfare* $E_{Q_r}[W_n(f)]$*, on the set of PR representation mechanisms;*
(c) and $(f_r, t_r) \to (f^*, t^*)$*, as* $r \to +\infty$*, pointwise.*

In essence, Definition 3 says that a representation mechanism is robust if it is the limit of a converging sequence of politically robust representation mechanisms, each element in the sequence maximizing the expression of expected welfare obtained with increasingly vague and diffuse priors. The requirements of robustness are quite strong; yet, we will exhibit an interesting member of the species below.

[16]If P was unknown to the FF, but was at the same time common knowledge among the agents, then, one could imagine a complete information revelation procedure to reveal P, and then use P as an input in the decision process, as proposed above. But we do not take this common knowledge properties for granted, and did not assume the existence of a common prior. In our setting, the goal of the public decision process is not to reveal hidden information to an asymmetrically informed planner, it is, literally, to *produce* pieces of otherwise non-existent information, by means of a costly sample of representatives.

4 Characterization and Existence of Robust Representation Mechanisms

Given the requirements of political robustness, our representation mechanisms must assume a simple form, as stated in the following theorem.

Theorem 1 *Under assumptions 1-6, $V = \Re^l$, and $|X| = l \geq 3$, let n be fixed. Then, (f, t, Λ^0) is politically robust if and only if, (a),*

$$f(\widehat{v}) \in \arg\max_{x \in X} \left\{ \sum_{i=1}^{n} \widehat{v}_i(x) + k(x) \right\},$$

for all $\widehat{v} \in V^n$, where k is an arbitrary function $X \to \Re$; (b), for all $i = 1, ..., n$,

$$t_i(\widehat{v}) = -\sum_{j \neq i, j=1}^{n} \widehat{v}_j \left[f\left(\widehat{v}\right) \right] - k \left[f\left(\widehat{v}\right) \right] - h\left(\widehat{v}_{-i}\right) \tag{8}$$

where h is an arbitrary mapping, depending on the $\widehat{v}_j$, $j \neq i$ only; t_0 is given by the budget constraint, i.e.,

$$t_0(\widehat{v}) = \frac{-1}{N+1-n} \sum_{i=1}^{n} t_i(\widehat{v});$$

and (c), the parameter space writes $\Lambda = V \times \mathcal{H}$, where $\mathcal{H}$ is the set of mappings $V^{n-1} \to \Re$, and parameters are fully specified in the constitution, i.e., $\Lambda^0 = \{(k, h)\}$.

For proof, see the appendix.

The strongest part of the above statement is the "only if" part. It is a classic exercise to show that a decision rule with the form given by Theorem 1 can be implemented in dominant strategies with the help of the transfers defined by expression (8), whatever the choice made for the arbitrary function k. These tranfers are a particular form of the classic Groves scheme. Expression (8) then provides us with a PR mechanism if k and h (the latter function being interpreted as the representatives' "base wage"), cannot be manipulated by the executive. To achieve this goal, it is sufficient to specify (k, h) in the constitution. The difficult part of the proof is to show that a PR representation mechanism can only be of this form. To this end, a deep result of K. Roberts (1979) provides us with a major step.

Note that under this representation mechanism, the budget is always balanced. This is in turn due to the definition of t_0 and the fact that there is at least one passive citizen in the revelation game, i.e., $n < N + 1$, there is at least one individual (i.e., the executive) who does not report a preference type. If the mechanism was a "direct democracy", in which every agent sends a message, revelation in dominant strategies could be incompatible with budget balance (e.g. Green and Laffont (1979)).

We have shown that PR mechanisms make decisions which maximize some sum of utilities, but the arbitrary function k appearing in expression (8) can be chosen in such a way that anything ranging from "representative democracy" (i.e. $k = \mathbf{0}$), and sheer dictatorship (i.e., take $|k|$ arbitrarily large relative to the sum of the representatives' utilities) is implementable in this sense. Thus, the Founding Fathers' influence on the future behaviour of institutions can be overwhelming, just by their choice of a small weight placed on the representatives' preferences in the decision-making process. Dictatorship being ordinarily defined as the imposition of a single *actual* person's preferences on everybody else, a crushing weight of k more precisely represents a form of predetermination coming from above, which can be called the "Reign of Tradition".

An interesting question is now to choose the value of k "optimally," that is here, in a robust way. According to Definition 3 above, to this end, we should pick a robust mechanism in the larger set of PR mechanisms. The next result shows that $k = \mathbf{0}$, an intuitively "focal" and *democratic* choice, is *robust*, thereby proving that the set of robust mechanisms is not empty.

Theorem 2 *Under assumptions 1-7, $V = \Re^l$, and $|X| = l \geq 3$, the following mechanism is robust: the mapping f satisfies,*

$$f(\widehat{v}) \in \arg\max_{x \in X} \left\{ \sum_{i=1}^{n} \widehat{v}_i(x) \right\}; \tag{9}$$

the transfers t are defined by expression (8) above with $k = \mathbf{0}$, and h is specified in the constitution, i.e., $\Lambda = V \times \mathcal{H}$, and $\Lambda^0 = \{(\mathbf{0}, h)\}$.

For proof, see the appendix.

We have proved the existence, but not the uniqueness of robust mechanisms. However, if there are robust choices other than $k = \mathbf{0}$, they must be very contrived, and are most likely uninteresting. Our solution puts all the weight on observations, i.e., the sample of representatives' preferences, and a zero weight on the Founding Fathers' prior estimates of willingnesses to pay. This result is obtained with a standard Bayesian approach: the robust decision rule is the limit of a sequence of optimal Bayesian decision rules obtained when prior knowledge becomes increasingly vague in an appropriate way. If prior probabilities are extremely vague, prior knowledge is not very reliable, and more weight should be given to sample observations (i.e., the representatives' preferences). Therefore, intuitively, k must converge towards $\mathbf{0}$.

On top of robustness, we can show that $k = \mathbf{0}$ has several other appealing properties, which cannot be satisfied by any other choice. First of all, $k = \mathbf{0}$ is the only choice which is unbiased in the following sense: it is the only PR mechanism whose decisions are based on an unbiased estimate of the true expected welfare, and this, for all probability distribution P on V. Formally, one can define a given k as *uniformly unbiased* if for all P, there exists a number

$\gamma > 0$, such that

$$\gamma E_P \left[\sum_{i=1}^{n} v_i(x) + k(x) \right] = E_P \left[\frac{1}{N+1} \sum_{i=0}^{N} v_i(x) \right].$$

This clearly implies $\gamma = 1/n$ and $k = \mathbf{0}$, for otherwise, k would depend on P, which is not permitted. But unbiasedness is not desirable *per se*. In decision or statistical problems in which a particular utility function should be used to rank decision rules, some biased estimates can be more desirable than unbiased ones, if they lead, say, to more precise evaluations of relevant parameters. This is why we needed a more radical, decision-theoretic definition of robust mechanisms; but we get uniform unbiasedness as an interesting by-product of robustness.

Our robust decision rule can also be called democratic because it depends on information about preferences that has been transmitted by the people only, through the representation process. The other PR rules, with non-zero k functions, would always bear the risk of being biased towards some preference, and would therefore be less "purely democratic".

On more intuitive grounds, $k = \mathbf{0}$ is also the only fully impartial and democratic choice, among the PR mechanisms in that it solves an imaginary dispute among Founding Fathers. Assume that there are several of these FF, that each FF is a rational and benevolent expected utility maximizer. The FF have the same utility function (which is the utilitarian sum of valuations), but they have differing probabilistic beliefs P. Assume in addition that all the FF are extremely stubborn. Each of them believes that he knows the true distribution and irreducibly disagrees with the others. They still need to reach an agreement. Now, assume that the FF are offered to inspect a sample of the citizen's preferences before a decision is made. To solve their bargaining problem, before the sample comes to be known, wouldn't they agree to let the final decision depend on sample information only? (For comments on a similar "story," see Savage (1972), chap. 10). A reasonable decision rule, which must be chosen by the FF before sample information is disclosed, would then depend only (and thus put all the weight) on sample information.

With our robust mechanism, not only are representatives provided with incentives to report truthfully information on their preferences, but also, once they are subjected to the Groves transfer scheme defined above, sample members *unanimously support* the robust decision rule (9). To see this, note that the utility of a representative writes,

$$u_i = \sum_{j=1}^{n} v_j(f(v)) + h(v_{-i}) - F,$$

and it follows that all representatives share the same objective, and would like to see f implemented. More precisely, given v, none of them would like f to be changed for another decision rule.

We have not proposed a particular choice for the function h, which remains arbitrary, given that it plays no role in the determination of x and in the expression of social welfare. If we were to consider seriously the question of *individual* *rationality*, and more precisely the willingness of representatives to participate, we should adjust the function h to make sure that they do not lose, in expected terms, from accepting the job and paying the cost F. This can be done at the expense of the passive taxpayers. An appealing, non-parametric choice for h is the Clarke pivotal mechanism, (which is a particular case of the Groves mechanism). The Clarke transfer function guarantees the best minimal utility levels, as shown by a result of Moulin (1986)[17], it is defined by setting:

$$h(v_{-i}) = F + \max_{x \in X} \sum_{j \neq i} v_j(x).$$

To sum up, we have exhibited a representation mechanism which solves, in the simplified world considered here, some important problems of collective decision-making in a representative democracy: the sampling procedure reduces the cost of collective decision-making, while at the same time ensuring honest representation, and adherence of representatives to the pursuit of general interest.

Some important issues remain however. The paper does not consider the question of the choice and election of representatives. The random sampling procedure is a shortcut, a way of black-boxing the electoral process and the party system, allowing us to study the structure of the representation problem under the assumption that an ideally unbiased electoral system can be found and implemented.

We do not defend random representation as a practical solution for democratic societies. This question has been discussed by others (see, among other writings, Dahl (1970)). The topic has also attracted the interest of the Public Choice school, e.g., Mueller et al. (1972), Tullock (1977). Representation by lot did exist in the Ancient Greece (e.g., Hansen (1991)), and in some other societies of the past. There are probably good reasons for which these systems have not survived in modern times (e.g., Manin (1995)), except for some institutions like criminal court jurys.

Viewed as a revelation mechanism, our robust representation mechanism can be called a *sampling Groves mechanism*, because it is applied to a sample of agents only. This random decision rule has some nice properties, one of them being that it is a good second-best approximation to the first-best Pareto optimum. The quality of this approximation increases when the dispersion of preferences decreases, and for a given sampling rate n/N, when the total population size N grows large, due to the Law of Large Numbers[18].

Finally, we should address the problem of the number n of representatives. A large value of n reduces sampling errors, but increases the total cost of representation nF. The optimal n should trade off these benefits and costs. In our view, n cannot be fixed in the constitution, because it typically depends on

[17] See also Moulin (1988), chap. 8.

[18] Given that the budget is balanced, revelation in dominant strategies and Pareto optimality are known to be incompatible in this context, and approximate optimality is the best that we can hope for. For more details on the properties of sampling Groves mechanisms see Gary-Bobo and Jaaidane (2000).

prior information about the dispersion of preferences, as will be shown in the next section. This is why we do not propose a "robust" theory of this number. In the real world, the number of representatives is not fixed in constitutions. In the United States, the number of seats in the House of Representatives has increased during the 19th century and reached a ceiling of 435 in 1910, which has been fixed by statute in 1929[19]. In France, this number is specified by an *organic act* of Parliament, which is higher in the hierarchy of norms than an ordinary act, but below the constitution. To follow the evolution of society, and population growth in the first place, the number of representatives can be changed without changing the constitution.

5 An Illustrative Example

To illustrate the concepts introduced above and the workings of the theory, we now compute an example of our model. Explicit computations make it clear then, that, to discuss the optimal number of representatives requires some prior information on the dispersion of preferences.

A utility function v_i can be viewed as a vector of $\Re^l$, that is, $v_i = (v_i(x))_{x \in X}$. If probability distribution P is known, the v_i are independent normal vectors. We further assume that for each i, the coordinates of v_i are uncorrelated with the same fixed variance σ^2, and more precisely, that $v_i \backsim \mathcal{N}(\mu, \sigma^2 I)$, where $\mu = (\mu(x))_{x \in X} \in \Re^l$ and I is the identity matrix of dimension l. To specify the FF's prior M, assume that μ itself is normally distributed with mean $\overline{\mu}$, and that the coordinates of μ are independent with variance z^2, that is, $\mu \backsim \mathcal{N}(\overline{\mu}, z^2 I)$. Under this specification, strictly speaking, M is not *diffuse* in the sense of Definition 3, because we view σ^2 as non-random, but this simplification is, in fact, harmless. Define $\widehat{V}(x) = \sum_{i=1}^n v_i(x)$ and denote $\widehat{V}$ the vector with coordinates $\widehat{V}(x)$, $x \in X$.

Under these normality assumptions, it is a well-known result from Bayesian statistical theory (see De Groot (1970), chap. 9.5), that, denoting the sample of utilities $\widehat{v} = (v_1, ..., v_n)$,

$$E_Q[\mu \mid \widehat{v}] = \frac{\sigma^2 \overline{\mu} + z^2 \widehat{V}}{\sigma^2 + z^2 n}. \tag{10}$$

Introducing the variance (or precision) ratio $\zeta = \sigma^2 / z^2$, we get the expression $E_Q[\mu \mid \widehat{v}] = (\zeta + n)^{-1}(\zeta \overline{\mu} + \widehat{V})$, and it is easy to see that, if the prior $\overline{M}$ becomes increasingly dispersed, then $\zeta \to 0$, and

$$E_Q[\mu \mid \widehat{v}] \to \frac{\widehat{V}}{n}, \tag{11}$$

the posterior expected mean of the distribution converges to the *arithmetic average* of the observed utility vectors.

[19]See O'Connor and Sabato (1993), p. 191.

We can now compute the expected welfare, from the point of view of the FF, given that the decision rule depends on $\widehat{v} = (v_1, ..., v_n)$ only, that is, $x = f(\widehat{v})$. By definition, and the law of conditional expectation,

$$\begin{aligned}
E_Q\left[W_n(f)\right] &= -(nF + F_0) + E_Q\left[\sum_{i=0}^{N} v_i(f)\right] \\
&= -(nF + F_0) + E_Q\left[E_Q\left[\sum_{i=0}^{N} v_i(f) \mid \ \widehat{v}\right]\right] \\
&= -(nF + F_0) + E_Q\left[\widehat{V}(f) + \sum_{j \notin \{1,...,n\}} E_Q\left[v_j(f) \mid \widehat{v}\right]\right] \quad (12)
\end{aligned}$$

Now, given that the v_j are independent, and using the law of conditional expectations again, we get for all $j = n+1, ..., N$ and $j = 0$,

$$\begin{aligned}
E_Q\left[v_j(f) \mid \widehat{v}\right] &= E_Q\left[E_Q\left[v_j(f) \mid \widehat{v},\ \alpha\right] \mid \widehat{v}\right] \\
&= E_Q\left[\mu(f) \mid \widehat{v}\right], \quad (13)
\end{aligned}$$

so that, finally, using the above assumptions and results,

$$\begin{aligned}
E_Q\left[W_n(f)\right] &= -(nF + F_0) + E_Q\left[\widehat{V}(f) + (N + 1 - n)E_Q\left[\mu(f) \mid \widehat{v}\right]\right] \\
&= -(nF + F_0) + \frac{\zeta + N + 1}{\zeta + n} E_Q\left\{\widehat{V}(f) + k(f)\right\}, \quad (14)
\end{aligned}$$

where,

$$k(x) = \frac{(N + 1 - n)\zeta}{\zeta + N + 1}\overline{\mu}(x). \quad (15)$$

We know from Theorem 1 that f is politically robust if it chosen so as to maximize $\widehat{V}(x) + k(x)$ with respect to x for each $\widehat{v}$, and that any politically robust rule is of this form for some k. The robust decision rule f^*, which maximizes $\widehat{V}(x)$, is clearly obtained by taking the limit of $k(x)$ when $\zeta \to 0$, that is, as shown by (15), $k(x) \equiv 0$.

We now substitute the robust rule f^* in the expression of expected welfare (14). Remark that with the robust rule $f = f^*$, we get by definition,

$$E\left[\widehat{V}(f^*)\right] = E\left[\max_{x \in X}\left\{\widehat{V}(x)\right\}\right]. \quad (16)$$

To push our computations further, let us simplify the model. Assume that the number of alternatives is the minimum necessary, that is, $l = 3$: the set X contains three objects denoted x_k with $k = 1, 2, 3$. Assume also that the coordinates of $\overline{\mu}$ have the same mean m, i.e., $\overline{\mu} = (m, m, m)$. Denote the arithmetic average of utilities $\overline{V} = (1/n)\widehat{V}$. Then, clearly, $\overline{V}$ has a normal distribution with mean (m, m, m), and has independent coordinates. The variance of each $\overline{V}(x)$ can be computed as follows, using the classic variance decomposition formula:

$$\begin{aligned} Var(\overline{V}(x)) &= Var\left[E\left(\overline{V}(x) \mid \mu, \sigma^2\right)\right] + E\left[Var\left(\overline{V}(x) \mid \mu, \sigma^2\right)\right] \\ &= Var(\mu(x)) + E\left(\frac{\sigma^2}{n}\right) = z^2 + \frac{\sigma^2}{n}. \end{aligned} \tag{17}$$

On the other hand, we know (e.g., Johnson et al. (1994)), that if the random variables Y_k are independent and $Y_k \backsim \mathcal{N}(0,1)$, then,

$$E\left[\max\left\{Y_1, Y_2, Y_3\right\}\right] = \frac{3}{2\sqrt{\pi}} \tag{18}$$

From this we derive the following,

$$\begin{aligned} E_Q\left[\max_x\left\{\widehat{V}(x)\right\}\right] &= nE_Q\left[\max_x\left\{\overline{V}(x)\right\}\right] \\ &= nm + n\sqrt{z^2 + \frac{\sigma^2}{n}}E\left[\max_x\left\{\frac{\overline{V}(x) - m}{\sqrt{z^2 + \sigma^2/n}}\right\}\right] \\ &= nm + n\sqrt{z^2 + \frac{\sigma^2}{n}}E\left[\max\left\{Y_1, Y_2, Y_3\right\}\right] \\ &= nm + \frac{3n}{2\sqrt{\pi}}\sqrt{z^2 + \frac{\sigma^2}{n}}. \end{aligned} \tag{19}$$

Using this result, we can obtain the expression of expected welfare. After some straightforward algebra, we get from (14), (15), (16) and (19),

$$E_Q\left[W_n(f)\right] = -(nF + F_0) + (N+1)m + \phi(n, \sigma^2, z^2), \tag{20}$$

where by definition,

$$\phi(n, \sigma^2, z^2) = \frac{3}{2\sqrt{\pi}}\left[\sigma^2 + (N+1)z^2\right]\sqrt{\frac{n}{\sigma^2 + nz^2}}. \tag{21}$$

Expression (20) is not well defined when $\zeta = 0$, because it happens that $\phi(n, \sigma^2, z^2) \to +\infty$ as $z^2 \to +\infty$[20]. The optimal number of representatives should be adapted to circumstances, and typically depends on the dispersion of preferences. Any interesting computation of this number will therefore depend on prior information on the distribution of preferences. We define as optimal the value of n which maximizes expected welfare for given m, σ and z.

It is easy to show that ϕ is a strictly concave function of n, viewed as a positive real variable. It follows that if the optimal *real* solution is interior, i.e., $1 < n^* < N$, then it must solve the following necessary and sufficient condition,

$$\frac{\partial\phi(n^*, \sigma^2, z^2)}{\partial n} = F. \tag{22}$$

[20] The function $E_Q\left[W_n(f)\right]$ is therefore not well-defined at $\zeta = 0$.

The optimal number is then the integer which is nearest to the solution n^* of (22). Remark that n^* does not depend on the prior mean m, but on the variances σ^2 and z^2. It can also be shown that the marginal value of an additional representative increases with the dispersion of preferences, and that the optimal number of representatives decreases with the cost of representation F, and increases with both the dispersion of preferences and the size of the population N. These predictions are reassuring[21].

In general, our robust approach has no reason to coincide with what would have been a standard approach to this problem, that is, assuming that some probability distribution P is common knowledge, and choosing f so as to maximize $\widehat{V}(x) + (N+1-n)\overline{v}_P(x)$, as shown by (7) above. A substitution of $f = f_P$ in the expression for expected welfare will typically not lead to the same optimal number of representatives as (22).

6 Conclusion

We have considered a class of representation mechanisms, based on reports made by a random subset of agents, in a collective choice problem with quasi-linear utilities. The absence of a benevolent planner, combined with the absence of common prior probability distributions, has led us to define new notions of robustness. A politically robust mechanism must be revealing in dominant strategies (whatever the probabilistic beliefs of the agents) and immune from opportunistic manipulations on the part of the executive, who is not assumed benevolent. An informationally robust mechanism maximizes expected welfare under a vague prior distribution, representing the Founding Father's ignorance as to the real profile of individual preferences. A representation mechanism is then called simply robust if it is informationally robust in the set of politically robust mechanisms. We have characterized politically robust mechanisms as a variant of Groves mechanisms, and shown the existence of robust mechanisms in this set. The robust mechanism that we exhibit "puts all the weight" on the representatives' reports while computing the collective decision. It is fully non-parametric in the sense that it does not depend on prior information about preferences, and is therefore not taylored to a particular society. In contrast, the usual approach to this problem would have made the optimal mechanism depend on the benevolent planner's prior probabilistic beliefs. These results can be viewed as a step in an attempt to construct a normative theory of public-decision making under very weak informational assumptions, and by giving a formal content to the idea that there is no benevolent planner.

[21] For results on the optimal number of representatives, see Auriol and Gary-Bobo (2000).

7 Appendix: proofs

Proof of Theorem 1.

(If) It is a standard result to show that (f, t), as defined in the statement of Theorem 1, is revealing in dominant strategies, and therefore universally revealing, with any function k and any function $h \in \mathcal{H}$. Given that the functions k and h are fixed in the constitution, the executive cannot manipulate them. It follows that the mechanism is also immune from executive opportunism (hereafter IEO), and therefore PR.

(Only if) This part of the proof uses a characterization result of Kevin Roberts (1979).

Theorem (Roberts). *If* $f : V^{N+1} \to X$ *is a collective decision rule which satisfies NI and is implementable in dominant strategies, then, there exists a vector of weights* $\lambda \in \Re^{N+1}, \lambda \geq 0, \lambda \neq 0$, *and a determinate real-valued function* $k : X \to \Re$ *such that for all* $v \in V^{N+1}$,

$$f(v) \in \left\{ x \in X : k(x) + \sum_{i=0}^{N} \lambda_i v_i(x) \geq k(y) + \sum_{i=0}^{N} \lambda_i v_i(y) \quad \forall \;\; y \in X \right\}.$$

Given this result, condition IEO imposes $\lambda_0 = 0$ and $\lambda_j = 0$ for all $j = n+1, ...N$.

If this was not the case, by Assumption 3 (subsidiarity of the executive), Agent 0 would have to provide for the missing information (v_j), for $j = 0$, and $j = n+1, ..., N$, in all the cases in which $n < N$, thus obviously contradicting IEO.

Now, if the weights λ_1 to λ_n where not all equal, i.e., $\lambda_1 = ... = \lambda_n$, the decision rule would violate Assumption 6, the principle of weak equality, for all agents are *ex ante* indistinguishable[22]. Without loss of generality, then, all weights can be set equal to one. It is also a well-known result (cf. Roberts (1979)) that f can only be implemented with the help of transfer schemes belonging to the "extended Groves family", that is, transfers of the form

$$t_i(\widehat{v}) = -\frac{1}{\lambda_i} \left[\sum_{j \neq i} \lambda_j \widehat{v}_j(f) + k(f) \right] - h_i(\widehat{v}_{-i})$$

The function h_i appearing in the transfer formula must be the same for every representative, because of Assumption 6 again. We therefore get transfers of the form given by (8), in the statement of Theorem 1.

Finally, the mechanism under scrutiny is IEO, and therefore PR, only if the functions k and h are fixed in the constitution. If this was not the case, using

[22] Representatives being chosen by lot, they cannot be labeled in advance, because their weight is attached to their function as representative, and not to a particular individual in society. This means that the Executive would be in charge of labeling the representatives. Now, the weights λ_i should be attributed to representatives independently of their message $\widehat{v}_i$. If these weights are not equal but fixed in the Constitution, the Executive could still choose the agent's label *ex post* and manipulate the decision. The anonymity condition $\lambda_1 = ... = \lambda_n$ forbids any such manipulation on the part of the Executive.

the budget balance condition (1), and the definition of transfers (8), we would get,

$$t_0 = \frac{1}{N+1-n}\left(\sum_{i=1}^{n} h(\widehat{v}_{-i}) + nk(f) + (n-1)\sum_{i=1}^{n}\widehat{v}_i(f)\right).$$

Now, t_0 being the tax paid by the executive, he (she) could manipulate k and h, to his (her) own advantage. Choosing small values of h would reduce the tax, while choosing the value of k can obviously bias the public decision toward the executive's private preferences v_0. Thus, to obtain a PR mechanism, k and h must be fixed in the constitution.
Q.E.D.

The Normal-Wishart Distribution

In the following proof, we use a particular specification for the prior probability distribution M of the Founding Fathers. M is a distribution of the parameters (μ, T), where by definition, $T = \Sigma^{-1}$ is the precision matrix of the multivariate normal distribution of v_i, i.e., given that we have assumed $v_i \backsim \mathcal{N}(\mu, \Sigma)$ for all i. The mean vector μ belongs to $\Re^l$ and T is a positive definite, symmetric matrix of dimension $l \times l$, which is an element of $\Re^l \times \Re^l_+ \times \Re^{l(l-1)/2}$.

From the point of view of the FF, the precision matrix T is random as well as the mean vector μ. This random matrix T is said to have a Wishart distribution with p degrees of freedom and precision matrix Γ if its density can be written,

$$g(T \mid p, \Gamma) = c\,|\Gamma|^{p/2}\,|T|^{(p-l-1)/2}\exp\left[-(1/2)\mathbf{tr}\ (\Gamma T)\right],$$

where $\mathbf{tr}(A)$ denotes the trace of matrix A, i.e., the sum of its diagonal elements[23], and c is the appropriate normalisation constant[24]. We have $E(T) = p\Gamma^{-1}$, and the distribution is diffuse.

Definition: *M is said to be Normal-Wishart with parameters $(p, \overline{\mu}, \Gamma, \beta)$, if the joint distribution of (μ, T) is as follows: the conditional distribution of μ knowing that $T = R$ is a multivariate normal distribution with mean vector $\overline{\mu}$ and precision matrix βR, $\beta > 0$, and the marginal distribution of the precision matrix T is a Wishart distribution with p degrees of freedom and precision matrix Γ, such that $p > l-1$.*

The following result can be proved.

Lemma 1 (see De Groot (1970), 5.6, 9.11). *If M is a Normal-Wishart distribution with parameters $(p, \overline{\mu}, \Gamma, \beta)$, then, the marginal distribution of μ is a multivariate Student t distribution with $p-l+1$ degrees of freedom, precision matrix $\beta(p-l+1)\Gamma^{-1}$, and $E(\mu) = \overline{\mu}$.*

Consider now the sample of preferences $\widehat{v} = (v_1, ..., v_n)$. It happens that the Normal-Wishart family is a *conjugate* family of distributions when v_i is

[23] The *determinant* of matrix A is denoted $|A|$.
[24] The Wishart distribution is a multivariate extension of the χ^2 distribution.

normally distributed, i.e., the posterior distribution of $(\mu, T \mid \widehat{v})$ is also of the Normal-Wishart type. More precisely, we can state the following result.

Lemma 2 (see De Groot (1970), 9.9-9.11). *Suppose that $\widehat{v} = (v_1, ..., v_n)$ is a random sample from a multivariate normal distribution with an unknown value of the mean vector μ and the precision matrix T. Suppose that the prior joint distribution of (μ, T) is Normal-Wishart with parameters $(p, \overline{\mu}, \Gamma, \beta)$. Then, the posterior marginal distribution of μ knowing $\widehat{v}$ is a multivariate Student t distribution with $p + n - l + 1$ degrees of freedom and*

$$E(\mu \mid \widehat{v}) = (\beta + n)^{-1} \left(\beta\overline{\mu} + \widehat{V}\right). \tag{23}$$

Proof of Theorem 2.
First compute the expected welfare, from the point of view of the FF, given that the decision rule f depends on $\widehat{v} = (v_1, ..., v_n)$ only, that is, $x = f(\widehat{v})$. By definition,

$$E_Q\left[W_n(f)\right] = -(nF + F_0) + E_Q\left[\sum_{i=0}^{N} v_i(f)\right] \tag{24}$$

By the laws of conditional expectation,

$$\begin{aligned} E_Q\left[\sum_{i=0}^{N} v_i(f)\right] &= E_Q\left[E_Q\left[\sum_{i=0}^{N} v_i(f)|\widehat{v}\right]\right] \\ &= E_Q\left[\widehat{V}(f) + \sum_{j\notin\{1,\ldots,n\}} E_Q\left[v_j(f) \mid \widehat{v}\right]\right], \end{aligned} \tag{25}$$

where $\widehat{V}(x) \equiv \sum_{i=1}^{n} v_i(x)$. Now, given that the v_j are independent, and using the law of conditional expectation again, we get for all $j = n+1, ..., N$ and $j = 0$,

$$\begin{aligned} E_Q\left[v_j(f(\widehat{v})) \mid \widehat{v}\right] &= E_Q\left[E_Q[v_j(f(\widehat{v})) \mid \widehat{v}; \mu, T] \mid \widehat{v}\right] \\ &= E_Q\left[\mu(f(\widehat{v})) \mid \widehat{v}\right], \end{aligned} \tag{26}$$

Using this result, we can write,

$$E_Q\left[W_n(f)\right] = -(nF + F_0) + E_Q\left[\widehat{V}(f) + (N + 1 - n)E_Q\left[\mu(f) \mid \widehat{v}\right]\right]. \tag{27}$$

We now construct a sequence of diffuse prior distributions, denoted (M_r) with the properties required by Definition 3. Let each element in the sequence be a Normal-Wishart distribution with parameters $(p_r, \overline{\mu}_r, \Gamma_r, \beta_r)$ and assume that the sequence $(p_r, \overline{\mu}_r, \Gamma_r)$ converges toward some finite limit $(p_\infty, \overline{\mu}_\infty, \Gamma_\infty)$, that $p_r > l$ and $|\Gamma_r| \neq 0$ for all r, and that $\beta_r \to 0$ as $r \to +\infty$. By Lemma 1, the precision matrix of the prior marginal distribution of μ is $\beta_r(p_r - l + 1)\Gamma_r^{-1}$ and converges toward $\mathbf{0}$ as $r \to \infty$. Point (a) in Definition 3 is satisfied.

We now compute the expected welfare under distribution Q_r, derived from M_r using expression (3) above. By Lemma 2, we get,

$$E_{Q_r}[\mu \mid \widehat{v}] = \frac{\beta_r \overline{\mu}_r + \widehat{V}}{\beta_r + n}. \tag{28}$$

Substituting (28) into (27) yields, after some straightforward algebra,

$$E_{Q_r}[W_n(f)] = -(nF + F_0) + \frac{N + 1 + \beta_r}{\beta_r + n} E_{Q_r}\left[\widehat{V}(f) + k_r(f)\right], \tag{29}$$

where,

$$k_r(x) = \frac{(N + 1 - n)\beta_r}{(N + 1 + \beta_r)} \overline{\mu}_r(x). \tag{30}$$

A glance at expression (29) now shows that a decision rule f_r satisfying,

$$f_r(\widehat{v}) \in \arg\max_{x \in X} \left\{\widehat{V}(x) + k_r(x)\right\}, \tag{31}$$

for all $\widehat{v}$, achieves the unconstrained maximum of $E_{Q_r}[W_n(f)]$ with respect to $f : V^n \to X$. In addition, by Theorem 1, there exists a vector of transfers t_r with which f_r can be implemented in dominant strategies. The transfers must assume the form given by expression (8) above with $k = k_r$, and with any given function h. The value of expected welfare does not depend on these transfer functions and m. Choose $\Lambda_r^0 = \{(k_r, h)\}$. The representation mechanism (f_r, t_r, Λ_r^0) is PR, and therefore, it also maximizes $E_{Q_r}[W_n(f)]$ on the set of PR mechanisms. Point (b) of Definition 3 is satisfied.

It is then easy to check that, by (30), $k_r \to \mathbf{0}$, because $\beta_r \to 0$, and $\overline{\mu}_r$ remains bounded, as $r \to \infty$. Given that X is a finite set, taking a converging subsequence of f_r if necessary, we get $f_r \to f^* \in \arg\max_x\{\widehat{V}(x)\}$, as $r \to \infty$, for all $\widehat{v}$. It follows that point (c) of Definition 3 is satisfied.
Q.E.D.

8 References

Auriol, E., and R.J. Gary-Bobo (2000), "On the Optimal Number of Representatives", Discussion paper n° 1286, *Center for Mathematical Studies in Economics and Management Science*, Northwestern University, Evanston, Illinois.

Bernardo, J., and A.F.M. Smith (1994), *Bayesian Theory*, Wiley and Sons, New York.

Clarke, E.H. (1971), "Multipart Pricing of Public Goods", *Public Choice,* 8, 19-33.

Dahl, R. A. (1970), *After the Revolution: Authority in a Good Society*, Yale University Press, New Haven.

DeGroot, M. H. (1970), *Optimal Statistical Decisions*, McGraw Hill, New York.

Dixit, A. (1996), *The Making of Economic Policy: A Transaction-Cost Politics Perspective*, MIT Press, Cambridge, Massachusetts.

Gary-Bobo, R. J., and T. Jaaidane (2000), "Polling Mechanisms and the Demand Revelation Problem", *Journal of Public Economics*, 76, 203-238.

Green, J. and J-J. Laffont (1977), "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica*, 45, 427-438.

Green, J. and J-J. Laffont (1979), *Incentives in Public Decision-Making*, North Holland, Amsterdam.

Groves, T. (1973), "Incentives in Teams", *Econometrica,* 41, 617-631.

Hagerty, K. M., and W. P. Rogerson (1987), "Robust Trading Mechanisms", *Journal of Economic Theory*, 42, 94-107.

Hansen, M. H. (1991), *The Athenian Democracy in the Age of Demosthenes. Structures, Principles and Ideology,* Blackwell, Oxford.

Holmström, B. (1979), "Groves Schemes on Restricted Domains", *Econometrica*, 47, 1137-1147.

Holmström, B., and R.B. Myerson (1983), "Efficient and Durable Decision Rules with Incomplete Information", *Econometrica*, 51, 1799-1819.

Hurwicz, L. (1972), "On Informationally Decentralized Systems", in McGuire and Radner eds., *Decision and Organization (in Honor of J. Marschak)*, 297-336, North-Holland, Amsterdam.

Johnson, N.L., S. Kotz and N. Balakrishnan (1994), *Continuous Univariate Distributions*, Volume 1, second edition, Wiley series in probability and statistics, Wiley and Sons, New York.

Laffont, J-J., and D. Martimort (2000), "Mechanism Design with Collusion and Correlation", *Econometrica*, 68, 309-342.

Ledyard, J. O. (1978), "Incentive Compatibility and Incomplete Information", *Journal of Economic Theory*, 18, 171-189.

Manin, B. (1995), *Principes du governement représentatif*, Calmann-Lévy, Paris; english transl. (1997), *The Principles of Representative Government*, Cambridge University Press, Cambridge.

Maskin E, and J. Moore (1999), "Implementation and Renegotiation", *Review of Economic Studies*, 66, 39-56.

Moulin, H. (1986), "Characterizations of the Pivotal Mechanism", *Journal of Public Economics*, 31, 53-78.

Moulin, H. (1988), *Axioms of Cooperative Decision-Making*, Cambridge University Press, Cambridge.

Mueller, D. C., Tollison, R. D. and T. D. Willett (1972), "Representative Democracy via Random Selection", *Public Choice*, 12, 57-68.

O'Connor, K. and L. Sabato (1993), *American Government, Roots and Reform*, McMillan, New York.

Palfrey T. and S. Srivastava (1991), "Efficient Trading Mechanisms with Pre-Play Communication", *Journal of Economic Theory*, 55, 17-40.

Palfrey T. and S. Srivastava (1993), *Bayesian Implementation*, Fundamentals of Pure and Applied Economics Series, Harwood Academic Publishers, Chur, Switzerland.

Roberts, K. (1979), "The Characterization of Implementable Choice Rules", chap. 18 in :J-J. Laffont ed., *Agregation and Revelation of Preferences*, North Holland, Amsterdam.

Savage, L. (1972), *The Foundations of Statistics*, Dover Publications, New York; first edition, (1954), Wiley and Sons, New York.

Tullock, G. (1977), "Practical Problems and Practical Solutions", *Public Choice*, 39, 27-35.