

# On the optimal number of representatives

Emmanuelle Auriol\* and Robert J. Gary-Bobo †

September 2010; revised 28 March 2011.

## Abstract

We propose a normative theory of the number of representatives based on a model of a representative democracy. We derive a formula giving the number of representatives as proportional to the square root of total population. Simple tests of the formula on a sample of a 100 countries yield good results. We then discuss the appropriateness of the number of representatives in some countries. It seems that the United States has too few representatives, while France and Italy have too many. The excess number of representatives matters: it is positively correlated with indicators of red tape and barriers to entrepreneurship.

*Keywords:* *Representative Democracy, Number of Representatives, Constitution Design, Incentives.* JEL No: D7, H11, H40.

---

\*Toulouse School of Economics, ARQADE and IDEI, email: emmanuelle.auriol@tse-fr.eu

†CREST, GENES, 15 boulevard Gabriel Péri, 92245, Malakoff cedex, France. email: robert.gary-bobo@ensae.fr.

# 1 Introduction

The production of public goods affects the well-being of large number of citizens, whereas a typically much smaller number of individuals is in charge of public decisions. This is true at almost all levels of society: there are parliaments at the national level, councils at the local levels and committees within public or private organizations. The presence of costs associated with the acquisition of information and with the preparation of decisions plays a major role in this concentration of power. The forces driving the division of labor help understanding the emergence of political representation. As a counterpart, protection against the opportunistic behavior of these representatives becomes a major justification for collective decision rules. This paper studies the trade-off between the need to economize on decision costs, suggesting that a small number of individuals should specialize in public decision-making, and the democratic requirement that decisions should reflect the citizens' true preferences. We derive a theory of the optimal number of representatives, and we find that a preliminary look at political data does not contradict its predictions.

We adopt a two-stage approach to constitutional design,<sup>1</sup> with a constitutional and a legislative stage, to derive the optimal number of representatives. In contrast to most of the recent work on constitution design, we completely black-box elections and voting and construct what could be called a reduced-form theory of representative democracy. The legislators' assembly is modeled as a random sample of preferences, drawn from the population of citizens. The randomly chosen representatives do not vote; they use a nonmanipulable, revealing mechanism instead. This mechanism reveals the representatives' preferences and efficient public decisions are carried out by a self-interested executive. During the preliminary constitutional stage, fictitious Founding Fathers choose decision rules behind the veil of ignorance, so as to maximize the expected total sum of citizens' utility. The Founding Fathers know that no agent is benevolent. It follows from this that the executive's hands must be tied as much as possible and that representatives must be provided with incentives to reveal preferences truthfully. In addition, our Founding Fathers know that they don't

---

<sup>1</sup>On this question, see the survey in Mueller (2003), and the discussion of some recent contributions below.

know the distribution of preferences that will prevail in society: we do not assume that this distribution is common knowledge. A robust mechanism is therefore required, in the following particular sense: among nonmanipulable preference-revealing mechanisms, the Founding Fathers pick a decision rule that maximizes expected utility against a vague (or noninformative) prior relative to citizens' preferences.<sup>2</sup> Robustness in this sense can be understood as a political stability requirement. The Founding Fathers know that society is going to evolve, but they cannot anticipate in which way. A constitution could not last for more than 200 years if it was tailored too closely to a particular preference profile.

Our model singles out a well-defined robust mechanism, that happens to be a Sampling Groves mechanism. Statistical sampling properties then yield an optimal sample size, trading off the direct and opportunity costs of representatives for the welfare loss induced by representation (*i.e.*, the loss due to the fact that a subset of citizens make decisions). A "square-root formula" for the optimal number of representatives directly follows from this stylized model of representation. The rule is then tested with a sample of more than 100 countries, and we find that our square-root theory is almost true and reasonably robust. Observations collected on the size of legislatures from around the globe are well-approximated by a number of national representatives proportional to  $N^{0.4}$ , where  $N$  is the country's total population. We also identify the United States, France and Italy as outliers. The former lie below the regression line; the latter two much above it. Indeed, constitutional History shows that the representation ratio has been decreasing during more than 200 years in the United States.<sup>3</sup> The number of seats in the House of Representatives reached a ceiling of 435 in 1910.<sup>4</sup> According to our results, the US Lower and Upper Houses should have a total of 800 members.

We finally check for correlation of the number of representatives with some indices

---

<sup>2</sup>Using a well-known technique from Bayesian statistics, a limiting argument is used to derive the effect of the Founding Fathers' ignorance on the optimal mechanism. The most technical aspects of our approach are presented in Auriol and Gary-Bobo (2007).

<sup>3</sup>Tocqueville (1835, part I, Chap. VIII, p 190, footnote) already noted the fact that the representation ratio decreased from 1 representative for every 30,000 inhabitants in 1792, to 1 over 48,000 in 1832. This trend has not been reversed ever since, the ratio reaching a record low of 1 over 611,000 in the recent years.

<sup>4</sup>This number has been fixed by statute in 1929. See O'Connor and Sabato (1993: 191). The number of seats in US State legislatures also seems to be characterized by institutional rigidity.

measuring the costs of setting up a new firm (i.e., "red tape") and the degree of state interference in markets.<sup>5</sup> The results are clearly that the number of representatives matters: it is positively and significantly correlated with state interference and red tape. More precisely, we cannot reject the fact that it is the *excess* number of representatives (i.e., the actual number less the number predicted by the  $N^{0.4}$  formula) which in fact matters for red tape and the degree of state interference.

As far as we know the problem of the optimal number of legislators has been studied by a handful of economists only<sup>6</sup>. In contemporary writings, Buchanan and Tullock (1962) are clearly the forerunners of the approach followed here<sup>7</sup>. Thinking about constitutional design, they developed a theory of the optimal constitution based on four variables: rules for choosing representatives; rules for deciding issues in assemblies; the degree of representation (i.e., the proportion of total population elected); and the basis of representation (i.e., for instance, the geographical basis). Buchanan and Tullock's approach is clearly normative, insofar as the goal of the analysis is to fix the four variables in order to minimize the expected sum of decision costs and *external* costs of institutions. Another forerunner is Stigler (1976), who sketched a theory of the degree of representation and reported some regression work on the number of representatives in relation to total population in the US states.

A small (but influential) number of authors belonging to the Public Choice school has played with the ideas emphasized here more than 40 years ago: following Dahl (1970), Mueller *et al.* (1972) discuss random representation. Tullock (1977) went as far as to ponder

---

<sup>5</sup>We use indices constructed by Barro and Lee (1994), and Djankov *et al.* (2002).

<sup>6</sup>This problem is essentially distinct from that of fair representation or apportionment, that was studied quite extensively, *e.g.* Balinski and Young (2001). Our theory is not related to L. S. Penrose's (1946) well-known square-root formula. This formula determines the size of a country's delegation in supra-national institutions like the UN or EU, not the number of representatives itself. The question of the appropriate number of seats in US Parliament was posed long ago by the founding fathers and opponents of the American Constitution. James Madison addressed the question in a famous passage of Federalist no 10 (see, Madison, *Federalist 10*; in Pole 1987: 155). The Anti-Federalist writers have emphasized a related point: "The very term, representative, implies, that the person or body chosen for this purpose, should resemble those who appoint them (...). Those who are placed instead of the people, should (...) be governed by their interests, or, in other words, should bear the strongest resemblance of those in whose room they are substituted. (...) Sixty-five men cannot be found in the Unites States, who hold the sentiments, possess the feelings, or are acquainted with the wants and interests of this vast country" (*Essays of Brutus*, III, 1787, in Storing 1981: 123).

<sup>7</sup>For more recent developments, see *e.g.*, McCormick and Tollison (1981), Weingast *et al.* (1981).

the possibility of using pivotal mechanisms in the US Congress to make public decisions. In the present paper, our intention is not to advocate recourse to random choice of legislators, or Groves mechanisms in practice, but to propose a model of representative democracy in reduced form and to derive a formula for the optimal number of representatives<sup>8</sup>.

There has been a recent revival of interest in the normative method among writers in political economy, voting theory and mechanism design. Our normative approach does not rely on the existence of a benevolent planner and our self-interested executives are clearly in line with the citizens-candidate approach of Osborne and Slivinski (1996) and Besley and Coate (1998). The two-stage approach to constitutional design recently received further impetus from Aghion and Bolton (2003), Barbera and Jackson (2004) and Gersbach (2009). Some contributions explore voting rules, or alternative collective decision procedures, with the idea of improving efficiency through a better expression of the intensity of preferences (e.g., Casella 2005). On strategic behavior and information aggregation in polling mechanisms, see, among other contributions, Gary-Bobo and Jaaidane (2000) and Morgan and Stocken (2008).

Our approach is also related to the emerging literature on the design of committees and recent trends in the theory of mechanisms. Early work on information acquisition and voting is due to Gersbach (1995). Condorcet's Jury Theorem has been reconsidered under the assumption of strategic voting by Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1998). Subsequent work has studied strategic behavior in jury or committee models with costly information acquisition<sup>9</sup>. Other contributions have studied costly information acquisition in mechanism design, assuming that agents have incomplete knowledge of their own preferences or valuations, for public or private goods<sup>10</sup>.

---

<sup>8</sup>We are not the first to adopt a "reduced-form approach" to model politics. For instance, in Becker (1983), political parties and voting receive little attention because "they are assumed mainly to transmit the pressure of active groups". More recent contributions in which a common agency model is used to study public policy-making can also be viewed as employing a reduced-form methodology (see, e.g., Dixit *et al.* 1997).

<sup>9</sup>On voting with costly participation, see also Palfrey and Rosenthal (1985), Osborne *et al.* (2000) and Börgers (2004). On committees, see, e.g., Li (2001), Persico (2004), Gerardi and Yariv (2008).

<sup>10</sup>On Bayesian incentive-compatible mechanisms, see Bergemann and Välimäki (2002); on auctions, see for instance Matthews (1984), Compte and Jehiel (2007). In a preliminary version of the present paper (Auriol and Gary-Bobo (1999)), we have considered sampling Groves mechanisms with information acquisition. In

In the following, Section 2 presents our basic assumptions; Section 3 develops our model of representation; Section 4 derives the robust representation mechanism and the square-root theory of the optimal number of representatives. Section 5 presents the empirical results: econometric tests of the square-root theory in the world and among the US State legislatures; it also discusses the empirical relevance of the number of representatives by showing its impact on red tape and other indicators of state interference. A few technical results are proved in the appendix.

## 2 The model

### 2.1 Basic assumptions

We consider an economy composed of  $N + 1$  agents, indexed by  $i = 0, 1, \dots, N$ . A public decision, denoted  $q$ , must be chosen from a set  $Q$ . Agent  $i$  will pay a tax denoted  $t_i$ . This tax must be interpreted as a subsidy if it is negative. Each agent's utility depends on the public decision and the tax.

**Assumption 1.** (Quasi-linearity) Utilities are quasi-linear, and defined as  $v_i(q) - t_i$ , where  $v_i$  is a private valuation function. Valuation functions belong to a set  $V$ . The set  $V$  is a closed and convex subset of a metric space.

These valuation functions can be viewed as random draws from some probability distribution  $P$  on the set of admissible valuation functions  $V$ . Distribution  $P$  is not common knowledge.

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

Society comprises three types of individuals. Agent  $i = 0$ , called the executive, is in charge of implementing the collective decision  $q$ . After some relabelling if necessary, agents  $i = 1, \dots, n$  are representatives; and agents  $i = n + 1, \dots, N$  are passive citizens. The task of representatives is to transmit information on preferences.

---

these models, an increase in the number of jury or committee members, analogous to an increase in the number of representatives in our model, causes a dilution of individual influence and reduces the individual incentives to acquire information.

The set of representatives essentially is a random sample of  $n \leq N$  agents (or, equivalently, a random sample of preferences  $v = (v_1, \dots, v_n)$ ).

**Assumption 3.** (Perfect Representation) The valuations of the  $n$  representatives are independent random drawings in the probability distribution  $P$ .

In practice, it is doubtful that voting mechanisms would produce an unbiased random sample of preferences. On the one hand, Assumption 3 might seem rather naïve, but can be defended if our goal is to construct a normative theory of representative democracy and to determine the optimal number of representatives. On the other hand, the idea of unbiased random representation provides a desirable simplification, putting the entire electoral process in a black box. Representatives being a random sample, there is a risk that some minorities will not be represented, and therefore the welfare loss is also random. The optimal representation problem is a tradeoff between expected losses and the costs of a larger representation. The permanent representation biases induced by some voting systems cannot be studied with the simplest form of this model. We will nevertheless continue to work with this convenient idealization. Representation by lot existed in some societies of the past (see Hansen 1991, Manin 1997); it has been discussed by political scientists (Dahl 1990) and is still used to select juries in some countries.<sup>11</sup> We also assume the following.

**Assumption 4.** (Cost of Representation) Each representative pays a fixed cost  $F$ , i.e., if  $i$  is a representative, then  $i$ 's utility is  $v_i(q) - t_i - F$ .

This cost can be viewed as the sum of direct and opportunity costs of becoming a representative or, alternatively, as an elementary form of information-acquisition cost paid by agent  $i$  to obtain information about one's own preferences  $v_i$ . Under the former interpretation, citizens use resources to transmit information to the collective decision system. Under the latter interpretation, individuals do not know their own utility function and must incur costs

---

<sup>11</sup>The ancient Greeks, in Athens, used random drawings to choose their legislators. The Athenian people's assembly itself, with its 6000 members, was in fact a random sample of the citizen population. Each citizen attending a session of this Assembly would receive the equivalent of a worker's daily wage. Socrates was sentenced to death by a jury of 501 randomly drawn citizens (see Hansen 1991).

to become aware of their own preferences. The two interpretations are compatible.<sup>12</sup>

Each representative will report to a representation mechanism. Individual  $i$ 's report, denoted  $\hat{v}_i$  is chosen from the set  $V$ .

**Definition 1** (Representation Mechanism). A representation mechanism is an array of functions  $(f, t)$ , where  $f$  is a collective decision rule mapping representatives' reports about preferences  $\hat{v} = (\hat{v}_1, \dots, \hat{v}_n)$  into  $Q$ , i.e.,  $q = f(\hat{v})$ , and a list of tax functions denoted  $t = (t_0, t_1, \dots, t_N)$ , satisfying the budget constraint  $\sum_{i=0}^N t_i = 0$ .

By definition, the constitution specifies  $(f, t)$  for every possible value of  $n$ , but  $n$  itself is not fixed in the constitution.

## 2.2 The first-best optimum

We can now compute the first-best optimum in the above defined economy. The standard Utilitarian, first-best Bayesian decision relies on the assumption that the distribution of preferences  $P$  is common knowledge. This first-best decision maximizes the function

$$EW = E_P \left\{ \sum_{i=0}^N (v_i(q) - t_i) \mid (\hat{v}_1, \dots, \hat{v}_n) \right\} - nF, \quad (1)$$

with respect to  $q$  in  $Q$ , subject to the budget constraint  $\sum_{i=0}^N t_i = 0$ , where  $E_P$  denotes the expectation with respect to probability  $P$ . Given that individual preferences are independent draws in probability distribution  $P$ , this is equivalent to solving the problem:

$$\max_{q \in Q} \left\{ (N + 1 - n)E_P(v(q)) + \sum_{i=1}^n \hat{v}_i(q) - nF \right\}, \quad (2)$$

where  $E_P(v(\cdot))$  is the average utility function in the population. To understand what this first-best optimum looks like, assume for example that preferences are quadratic, with a single-dimensional parameter  $\theta$ , i.e.,  $v_i(q) = \theta_i q - q^2/2$  and that  $q$  is a nonnegative real number. Assume in addition that  $P$  is such that  $E(\theta) = \mu$  and  $Var(\theta) = \sigma^2$ . With these

---

<sup>12</sup>It is of course possible to extend the model to take coordination costs into account. A straightforward generalization would be to let the "fixed" cost  $F$  become an increasing function of  $n$ .

specifications, representative  $i$ 's report is a real number denoted  $\hat{\theta}_i$  and (2) becomes,

$$\max_{q \in Q} \left\{ q \left[ \sum_{i=1}^n \hat{\theta}_i + (N+1-n)\mu \right] - (N+1)\frac{q^2}{2} - nF \right\}. \quad (3)$$

This immediately yields the optimal decision

$$q^* = f^*(\hat{\theta}_1, \dots, \hat{\theta}_n) = \frac{1}{N+1} \left( \sum_{i=1}^n \hat{\theta}_i + (N+1-n)\mu \right); \quad (4)$$

Substituting (4) into  $EW$ , taking the expectation with respect to the distribution of  $\theta_i$ , yields the ex ante expected welfare associated with the optimal decision rule  $f^*$ . After some easy computations, we obtain

$$EW(f^*) = \frac{n\sigma^2}{2(N+1)} + \frac{(N+1)\mu^2}{2} - nF, \quad (5)$$

where we make use of the fact that the  $\hat{\theta}_i$  are i.i.d. This function being linear with respect to  $n$ , we can state the following result.

**Proposition 1.** *Assume that the distribution of preferences is common knowledge, then, with quadratic preferences, the first-best optimum has two possible values: either  $n^* = N+1$ , if  $\sigma^2 > 2(N+1)F$ , (i.e., a Direct Democracy), or  $n^* = 0$ , if  $\sigma^2 \leq 2(N+1)F$ , (i.e., a "Reign of Tradition").*

The interpretation of Proposition 1 is easy. If the dispersion of preferences is large enough with respect to costs of representation, then direct democracy is first-best optimal. In other words, if  $F$  is small, or if the number of citizens is small, then democracy must be direct. The only other case is not a democratic constitution: we call this "Reign of Tradition" because it is not dictatorship (which would correspond to  $n = 1$ ). In the Reign of Tradition, no citizen is endowed with the power of deciding on behalf of others and we can view the public decision as being the result of "Tradition", i.e.,  $f^* = \mu$ . Another equivalent view is that the decision is made by a disembodied benevolent planner. This arrangement is optimal only if the dispersion of preferences is small or if the population is large and if, in addition, the prior mean of preference parameters  $\mu$  is common knowledge. Proposition 1 is disappointing, because it never prescribes a representative democracy, in which the solution

would be interior, i.e.,  $0 < n^* < N + 1$ . The most likely case is one in which  $F$  is small but nonnegligible,  $N$  is very large, and tastes do not differ in an extreme way, which seems to indicate that the Reign of Tradition would often be the recommended solution for reasonably homogenous societies.<sup>13</sup> This failure to pick a representative democracy as a solution is not essentially due to the fact that expected welfare is linear with respect to  $n$  under quadratic preferences (and to the fact that total representation costs  $nF$  are linear). It stems from the assumption that the distribution of preferences is common knowledge. Indeed, if this is the case, if in addition  $N$  is large and if the dispersion of tastes is reasonable, by the Law of Large Numbers,  $\mu$  is an excellent estimator of the true population-mean of individual valuations and it is not useful to ask citizens about their taste parameters. Our claim is that there is something wrong with the above definition of the optimum, because the model describes a world in which information is not really decentralized. The model is that of an abstract planner, assumed to be benevolent, endowed with prior knowledge of the distribution of preferences (i.e.,  $(\mu, \sigma)$  in the quadratic example), but in a large economy with quadratic preferences, if the planner knew  $\mu$ , he would know the only useful parameter: Democracy would then be useless.

In Section 4 below, we propose a different model in which information is fully decentralized, the distribution of tastes is not common knowledge and democratic representation is a useful (and only) way of producing information. Section 3 will first provide some basic definitions and pose the representatives' incentive compatibility problem.

### 3 Representation and incentives

We now study the constitutional stage. To give formal content to the idea of an impartial and benevolent point of view on society, we assume the existence of fictitious agents called the Founding Fathers (hereafter the FF). The FF are in charge of writing the constitution; they are assumed benevolent, Bayesian, and Utilitarian, and they do nothing in the economy, apart from setting constitutional rules. These FF know that, once the set of rules embodied in the

---

<sup>13</sup>A large number of representatives is in contrast justified by large heterogeneities regarding ethnicity, religion and language in a given country, since then  $\sigma$  is of considerable size.

constitution will be applied, there will not exist a single omniscient, impartial and benevolent individual to carry out public decisions. A disembodied “social planner” is not assumed to play an active role. This imposes restrictions on the set of admissible mechanisms, described in sub-section 3.1. The ensuing preference revelation problem is studied in sub-section 3.2.

### 3.1 Basic constitutional principles

The FF apply some important principles. First, Separation of Power holds: the executive cannot be a representative. Second, a Subsidiarity Principle applies. According to Definition 1 above, a representation mechanism is an array of functions  $(f, t)$ . To work in practice, such a mechanism needs to be fully specified and this specification may depend on a number of controls or parameters. We need to allocate the power to choose the exact value of these parameters, and these choices may open some possibilities of manipulation. This motivates the following definition.

**Definition 2** (Subsidiarity Principle). With the exception of the number of representatives  $n$  itself, if the parameters needed to fully pin down and implement mechanism  $(f, t)$  are not specified in the constitution and are not provided for by the representatives according to constitutional rules, then they are chosen by the executive.

The Subsidiarity Principle simply says that the executive will fill all the gaps in the public decision process. It can of course be dangerous to let the executive choose crucial parameters freely, because this executive is endowed with unknown preferences ( $v_0$  is a random draw in  $P$ ) and would be tempted to pursue private goals.

Third, the FF also apply a principle of “Anonymity” (or “Equality” in a weak sense), which requires equal treatment of indistinguishable individuals. This forces equal tax treatment of all passive citizens, because their preferences are unknown (and there is no basis for discrimination among them). Let  $t_0$  denote the tax of agents  $i = n + 1, \dots, N$  and  $i = 0$ . The budget constraint can thus be rewritten as follows:

$$\sum_{i=1}^n t_i + (N + 1 - n)t_0 = 0. \tag{6}$$

### 3.2 Incentive compatibility

The decision rule  $f$ , as well as taxes  $t$ , should be immune to manipulations of the representatives and of the executive. Appealing to the Revelation Principle, we require the representation mechanism  $(f, t)$  to be direct and revealing. But the agents' beliefs about others' preferences are not common knowledge and are unknown to the FF. Mechanism  $(f, t)$  must therefore be revealing whatever the beliefs of the representatives. In this context, it almost immediately follows that  $(f, t)$  must be revealing in dominant strategies (see Ledyard (1978)), i.e., for all  $i = 1, \dots, n$ , for all  $v_i, \widehat{v}_i$ , and  $v_{-i}$ , we must have

$$v_i(f(v)) - t_i(v) \geq v_i(f(\widehat{v}_i, v_{-i})) - t_i(\widehat{v}_i, v_{-i}),$$

where, as usual, we denote  $v_{-i} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$  and  $v = (v_i, v_{-i})$ .

Because of the subsidiarity principle, the self-interested executive could choose the free parameters of  $(f, t)$  to maximise his (her) own utility  $v_0$ . These parameters must therefore be fixed in the constitution. In our simple model, revelation in dominant strategies plus "mast-tying" of the executive, put together, define non-manipulability.

**Definition 3** (Non-Manipulability). A representation mechanism  $(f, t)$  is nonmanipulable if it is revealing in dominant strategies and if all its parameters are specified in the constitution.

This definition means that, in addition to the revelation property, there are no free parameters that the executive could manipulate. It is possible to prove (see the appendix, for comments and a formal statement), that under the separation-of-powers, subsidiarity and anonymity principles, nonmanipulable mechanisms must assume the following form: the decision rule  $f(\cdot)$  must maximize an objective which is the sum of an arbitrary function  $k$  and of the utilities reported by representatives, i.e.,

$$f(\widehat{v}) \in \arg \max_{q \in Q} \left\{ k(q) + \sum_{i=1}^n \widehat{v}_i(q) \right\}. \quad (7)$$

And for all  $i = 1, \dots, n$ , representatives must be bound by the following transfer schedules:

$$t_i(\widehat{v}) = - \sum_{j \neq i} \widehat{v}_j(f(\widehat{v})) - k(f(\widehat{v})) + m(\widehat{v}_{-i}), \quad (8)$$

where  $m$  is an arbitrary fixed function that does not depend on  $\widehat{v}_i$ . Finally, arbitrary functions  $k$ , and  $m$  must be fixed in the constitution. Obviously, the choice of these crucial parameters cannot be left to the executive, because the choice of  $k$  can distort decisions radically, while the choice of  $m$  can distort transfers. We assume that the FF are constrained to choose  $f(\cdot)$  in this set of nonmanipulable mechanisms. When  $k \equiv 0$ , the class of nonmanipulable mechanisms boils down to the well-known class of Clarke-Groves mechanisms, but restricted to a random subset of agents called the representatives<sup>14</sup>.

Note that these mechanisms are budget-balanced by construction, because there is at least one citizen who is not a representative (i.e., at least agent 0 does not report about his (her) preferences). In other words, passive citizens form a sink used to finance the revelation incentives of the representatives<sup>15</sup>.

## 4 Robust representation mechanisms under decentralized knowledge

The novelty of our approach is that we have assumed that the FF do not know the probability distribution of citizens' preferences  $P$ , and they know that nobody knows it. We add the constraint of *decentralized knowledge* to the assumptions of asymmetric information and individual opportunism: the probability distribution of preferences  $P$  is not common knowledge.

The fact that the FF do not know the real  $P$  poses a problem because they cannot fully specify the expected (or average) welfare function that they would like to maximize by means of the choice of a constitution. There are several ways of modeling behavior under ignorance in decision theory. One is to use a non-probabilistic representation and a maximin principle or, some more sophisticated variant in which the decision-maker uses a set of probability

---

<sup>14</sup>On Clarke-Groves mechanisms, see Clarke (1971), Groves (1973), Green and Laffont (1979), Holmstrom (1979), Moulin (1986). On sampling Groves mechanisms, see, Green and Laffont (1977), Gary-Bobo and Jaaidane (2000).

<sup>15</sup>It follows that there are no inefficiencies due to budget imbalance (budget surplus), as in the usual theory of pivotal mechanisms. The only welfare losses are due to the fact that the information on preferences used by a representation mechanism is not exhaustive; in other words, social costs are caused by sampling errors. On these points see Gary-Bobo and Jaaidane (2000).

distributions. The constitution would then be chosen so as to maximize welfare against the worst-case scenario. Another approach is to choose decision rules that are optimal against a *non-informative*, or *vague* prior. In contrast, this is a purely Bayesian approach. We choose this latter route here. There is a mathematical difficulty in the representation of a decision maker’s complete prior ignorance because a uniform distribution on the real line (or on the set of integers) doesn’t exist.<sup>16</sup> It follows that a situation of complete prior ignorance can be approached by limiting arguments, letting the prior’s variance go to infinity.

## 4.1 Admissible decision rules

We assume that the FF restrict themselves to choosing a decision rule that satisfies “Weak Utilitarianism”.<sup>17</sup>

**Definition 4** (Weak Utilitarianism). For every array of reports  $\hat{v} = (\hat{v}_1, \dots, \hat{v}_n) \in V^n$ , the decision rule  $f$  should maximize the expected utility  $E_{P_0(\hat{v})}(v(q))$  with respect to  $q$  for *some* probability distribution  $P_0(\hat{v})$  with support included in  $V$ .

Imposing Weak Utilitarianism in the sense of Definition 4 means that the decision rule must maximize *some* weighted sum of utilities. Given that the FF are already assumed to be Utilitarians, this requirement is very weak, because  $P_0$  can be chosen arbitrarily and vary with  $\hat{v}$ . But the fact that the FF are utilitarians is of course important, because they will write the constitution in order to constrain representatives to pursue the common interest.<sup>18</sup>

We can now derive what we call robust mechanisms. It is easy to see that, under non-manipulability, the FF’s goal is essentially to choose the arbitrary function  $k$ . The weak utilitarianism requirement imposes further constraints on the choice of  $k$ . This arbitrary function must be of the form  $k(q) = bv_0(q)$ , where  $b$  is a nonnegative weight and  $v_0$  is a valuation function chosen in  $V$ . We prove the following Lemma,

---

<sup>16</sup>Bayesian statisticians have developed the theory of improper priors. See, e.g., Bernardo and Smith (1994).

<sup>17</sup>But the utilitarian principle could also be derived, in the manner of Harsanyi (1955), by assuming that the FF are rational decision-makers, and choose the objective function behind the veil of ignorance.

<sup>18</sup>Gersbach (2000) shows that more information in collective choice may harm some, a majority or even the entire electorate when voters or representatives pursue different objectives. Our setting can underestimate the need for representation insofar as it strongly relies on the commitment value of the constitution.

**Lemma 1.** *The nonmanipulable decision rule  $f$  satisfies weak utilitarianism if and only if  $k$  can be expressed as  $k = bv_0$  where  $b \geq 0$  is a scalar and  $v_0 \in V$ .*

*For proof, see the appendix.*

To sum up, the Founding Fathers apply the following principles: (i) Separation of Powers (the executive doesn't reveal preferences: this is the task of representatives); (ii) Subsidiarity (any input of the mechanism that is not provided by the representatives is chosen by the executive: hence the need to tie the executive's hands); (iii) Anonymity (taxes are the same for all the citizens that are not representatives); (iv) Non-manipulability (this forces the decision rule to assume a certain form, compatible with the revelation of preferences, but also to rigidly fix parameters such as  $k$  in the constitution); (v) Weak Utilitarianism (this further constrains the set of admissible decision rules by removing some arbitrariness). We now need a framework in which mechanism robustness can be precisely defined.

## 4.2 Definition of robust mechanisms

Formally, the social surplus function is defined as

$$W(f) = -nF + \sum_{i=0}^N v_i(f). \quad (9)$$

This function is the total sum of all the citizens' utilities. The FF would like to maximize the expected value of this social surplus with respect to decision rule  $f(\cdot)$ , subject to nonmanipulability and weak utilitarianism. In this perspective, we assume that they have a "prior on priors", i.e., a distribution  $B$  on possible priors  $P$ ; and we assume that  $B$  is uninformative — this represents the FF's lack of knowledge about the true distribution of citizens' preferences. Expected social welfare can be expressed as  $E_B E_P(W)$ , were  $W$  is defined by (9). The only problem is now to give formal content to the idea that the FF will choose a nonmanipulable  $f(\cdot)$  so as to maximize  $E_B E_P(W)$  under a vague (or non-informative) probability  $B$ . Such a decision rule will be called 'robust'. Intuitively, this can be done by a simple limiting argument, if  $P$  belongs to a family with a finite vector of parameters, by letting the precision of  $B$  converge towards zero (or equivalently, by letting the variance-covariance

matrix of  $B$  go to infinity). This definition is involved, but the intuition is simple: find the nonmanipulable mechanism that maximizes expected welfare under the “veil of ignorance”, using a non-informative prior.

Auriol and Gary-Bobo (2007) have studied the existence of robust mechanisms in this sense, assuming that the set of public decisions is finite, that individual preferences profiles can be any vector and that these vectors are multivariate normal (i.e.,  $P$  is multivariate normal, according to the Founding Fathers’ beliefs). Thus, the domain of preferences is general, but a normality assumption is used. As in portfolio theory, we can weaken the normality requirement, but will obtain a tractable model only if utility is assumed to be quadratic. We follow this direction here, because our theory can easily be illustrated in the classic quadratic-preference setting.

**Assumption 5.** (Quadratic preferences) Decision  $q$  is a real number and

$$V = \left\{ v(q) = \theta q - \frac{q^2}{2}, \theta \in \mathbf{R} \right\}. \quad (10)$$

In this simple setting, the true probability distribution  $P$  is just a one-dimensional distribution of the taste parameter  $\theta$ , with a finite mean  $\mu_P$ , and a finite variance  $\sigma_P^2$ . In this case, we also assume that the FF do not know  $(\mu_P, \sigma_P^2)$ , but that they are endowed with a prior  $B$  on possible pairs  $(\mu_P, \sigma_P^2)$ . In addition we assume the following:

$$E_B(\mu_P) = \hat{\mu}, E_B(\sigma_P^2) = \hat{\sigma}^2, \text{ and } Var_B(\mu_P) = \hat{z}^2, \quad (11)$$

where  $\hat{\mu}$ ,  $\hat{\sigma}^2$ ,  $\hat{z}^2$  are themselves finite, and where  $\hat{\mu}$  is the mean of the possible means,  $\hat{\sigma}^2$  is the mean of the possible variances, and  $\hat{z}^2$  is the variance of the possible means. The prior variance of  $\theta$ , from the FF’s point of view, is denoted  $Var_{FF}(\theta)$ , and admits the well-known decomposition,

$$\begin{aligned} Var_{FF}(\theta) &= Var_B[E(\theta|P)] + E_B[Var(\theta|P)] \\ &= \hat{z}^2 + \hat{\sigma}^2. \end{aligned}$$

We propose the following simple formal definition.

**Definition 5** (Robust Representation Mechanism). A mechanism  $(f, t)$  is robust if it is the limit of a sequence  $(f_k, t_k)$  of mechanisms, such that each  $(f_k, t_k)$  maximizes  $E_{B_k}(E_P W)$  on the set of nonmanipulable mechanisms, where  $(B_k)$  is a sequence of priors with the property that that  $\hat{z}_k^2$  goes to  $+\infty$ , while  $\hat{\sigma}_k^2/\hat{z}_k^2$  goes to zero.

To understand this definition, assume that all possible  $P$  distributions have the same variance  $\sigma_P^2 = \hat{\sigma}^2$ , but that their mean  $\mu_P$  is unknown to the FF. To approach complete ignorance, we let the variance of the possible means, i.e.,  $\hat{z}^2$ , go to infinity. As indicated above, a more general definition is of course possible, but would be more technical.

### 4.3 Derivation of the robust mechanism in the case of quadratic utility

Under Assumption 5, nonmanipulability and weak utilitarianism force us to choose a utility function  $v_0$  of the form  $v_0(q) = \alpha q - q^2/2$  with a weight  $\beta \geq 0$  and a decision rule  $f^{**}(\cdot)$ , such that

$$f^{**}(\hat{\theta}_1, \dots, \hat{\theta}_n) \in \arg \max_q \left\{ q \sum_{i=1}^n \hat{\theta}_i - \frac{nq^2}{2} + \beta \left( \alpha q - \frac{q^2}{2} \right) \right\}, \quad (12)$$

assuming that each representative  $i$  reports  $\hat{\theta}_i$ . We immediately find

$$f^{**}(\hat{\theta}_1, \dots, \hat{\theta}_n) = \frac{\sum_{i=1}^n \hat{\theta}_i + \alpha\beta}{n + \beta}. \quad (13)$$

Let now  $W_P(\alpha, \beta)$  be the expected surplus for a given distribution  $P$  and  $f^{**}$  as above. We have

$$W_P(\alpha, \beta) = E_P \left\{ f^{**}(\hat{\theta}) \sum_{i=0}^N \theta_i - \frac{(N+1)f^{**}(\hat{\theta})^2}{2} \right\} - nF. \quad (14)$$

We then compute the expected value of  $W_P$  with respect to the FF's prior  $B$ . Some computations yield the following formula.

**Lemma 2.**

$$\begin{aligned} E_B [W_P(\alpha, \beta)] &= \left( n + \beta - \frac{N+1}{2} \right) \frac{n\hat{\sigma}^2}{(n+\beta)^2} + \frac{b^2(N+1)}{2(n+\beta)^2} (2\alpha\hat{\mu} - \alpha^2) \\ &\quad + \frac{n(N+1)}{(n+\beta)^2} \left( \frac{n}{2} + \beta \right) (\hat{\mu}^2 + \hat{z}^2) - nF. \end{aligned} \quad (15)$$

For proof, see the appendix.

For given  $B$ , the best mechanism is obtained as a maximum of  $\overline{W} = E_B [W_P(\alpha, \beta)]$  with respect to  $(\alpha, \beta)$ . We find the following result.

**Lemma 3.** *For given  $B$ , the optimal values of  $\alpha$  and  $\beta$  are  $\alpha^{**} = \widehat{\mu}$ , and*

$$\beta^{**} = \frac{(N+1-n)\widehat{\sigma}^2}{\widehat{\sigma}^2 + (N+1)\widehat{z}^2}. \quad (16)$$

For proof, see the appendix.

This solution can be rewritten as a function of the ratio  $\zeta = \widehat{\sigma}^2/\widehat{z}^2$ . We immediately find the limit of  $\beta^{**}$  as  $\zeta \rightarrow 0$ ,

$$\lim_{\zeta \rightarrow 0} \beta^{**} = \lim_{\zeta \rightarrow 0} \frac{(N+1-n)\zeta}{\zeta + (N+1)} = 0.$$

Under decentralized knowledge, the only robust mechanism entails  $v_0(q) = \widehat{\mu}q - q^2/2$  and  $\beta^{**} = 0$  and therefore, the arbitrary function  $k$  must be set identically equal to 0. This mechanism is a *sampling Groves mechanism*. To make a public decision, it relies on the representatives' reports only. Formally, we have just proved the following result.

**Proposition 2.** *Under Assumptions 1-5, the only robust mechanism  $f^{**}(\widehat{v})$  maximizes  $\sum_{i=1}^n \widehat{v}_i(q)$ , with transfers  $t$  given by (8) above.*

Since preferences are assumed to be quadratic, we get  $q^{**} = f^{**}(\widehat{v}_1, \dots, \widehat{v}_n) = (1/n) \sum_{i=1}^n \widehat{\theta}_i$ . In fact, the same sampling Groves mechanism is robust in our sense with a much more general set of preferences, but at the cost of some normality assumption (on  $P$ , not on  $B$ ).<sup>19</sup>

The sampling Groves mechanism solves a number of difficult problems of a representative democracy simultaneously. It saves on the costs of producing information on preferences, captured by the fixed cost  $F$ , because of sampling; it ensures honest revelation of their preferences by representatives in a very strong sense (i.e., Groves mechanisms are revealing in dominant strategies); and finally, once subjected to the incentive transfer system (8) (see

---

<sup>19</sup>Normality is not required here. Again, see Auriol and Gary-Bobo (2007).

also Proposition A1 in the appendix), every representative adheres to the same social objective (i.e., every representative agrees with the objective of maximizing  $\sum_{i=1}^n \widehat{v}_i(q)$ ). The interpretation of this result is that the legislative bargaining process yields an approximate Pareto optimum, insofar as the representation is a correct mirror image of the population's preferences. Of course, this nice solution is obtained for a somewhat simplified economy with quasi-linear preferences (i.e., a public good economy with possibilities of compensation).

Remark that, if we let the prior's variance  $\widehat{z}^2$  go to zero instead, while  $\widehat{\sigma}^2$  remains bounded, then, we find  $\lim_{\zeta \rightarrow \infty} \beta^{**} = N + 1 - n$ . This means that the FF know the distribution of preferences in society for sure. In this case, the recommended solution is the standard Bayesian mechanism of sub-section 2.2, where  $v_0(q) = \widehat{\mu}q - q^2/2 = E_P(v(q))$  and  $N + 1 - n$  is the appropriate weight of  $v_0$  in the expected welfare function  $E[W | q, \widehat{v}_1, \dots, \widehat{v}_n]$  (and  $N + 1 - n$  is also the number of passive citizens). In this latter case, the sampled agents represent only themselves, while in the robust mechanism, sampled agents are truly representatives: they stand for the entire society. This is a major difference. We now show that in this setting, an optimal number  $n^{**}$  can be interior, i.e.,  $0 < n^{**} < N + 1$ , in sharp contrast with the standard Bayesian first-best analysis presented in sub-section 2.2.

#### 4.4 Optimal number of representatives

We can now compute the optimal number of representatives, denoted  $n^{**}$ . Substituting the robust decision rule  $f^{**}(\theta) = (1/n) \sum_{i=1}^n \theta_i$  in the expression for expected welfare yields

$$\overline{W} = \frac{N+1}{2}(\widehat{\mu}^2 + \widehat{z}^2) + \frac{\widehat{\sigma}^2}{2} - \left[ \frac{1}{n} - \frac{1}{N+1} \right] \frac{(N+1)\widehat{\sigma}^2}{2} - nF. \quad (17)$$

Define  $q_{N+1} = \frac{1}{N+1} \sum_{i=0}^N \theta_i$ . If we compute the first-best surplus in an economy with  $N + 1$  agents, using complete knowledge of the preference profile and then take expectations, we find

$$\begin{aligned} E_B E_P \left[ q_{N+1} \sum_{i=0}^N \theta_i - (N+1) \frac{q_{N+1}^2}{2} \right] - nF &= (N+1) E_B E_P \left[ \frac{q_{N+1}^2}{2} \right] - nF \\ &= \frac{\widehat{\sigma}^2}{2} + \frac{N+1}{2}(\widehat{\mu}^2 + \widehat{z}^2) - nF. \end{aligned} \quad (18)$$

Let  $r_n = \frac{1}{n} \sum_{i=1}^n \theta_i$ . Under the robust mechanism, we get the following expression of welfare,

$$\bar{W} = (N + 1)E_B E_P \left[ r_n q_{N+1} - \frac{q_n^2}{2} \right] - nF. \quad (19)$$

Taking the difference of expressions (18) and (19), we find the welfare loss (with respect to the complete information first-best) to be

$$L(n) = \frac{(N + 1)}{2} E_B E_P (q_{N+1} - r_n)^2. \quad (20)$$

It is then easy to check that

$$L(n) = \left[ \frac{1}{n} - \frac{1}{N + 1} \right] \frac{(N + 1)\hat{\sigma}^2}{2}; \quad (21)$$

and it follows that expression (17) is first-best surplus, minus the cost of representatives, minus the welfare loss due to the fact that some information on preferences is not reported. The optimal number of representatives  $n^{**}$  trades off the cost of an additional representative with the benefit of reducing the welfare loss, i.e.,  $n^{**}$  minimizes  $nF + L(n)$ . The representatives protect citizens against arbitrary public decisions, but there is a social cost of representation.

Observe that the social cost of representation  $nF + L(n)$  does not depend on  $\hat{z}^2$  (which can thus be arbitrarily large). It follows that if the FF had prior information on the variance of preferences  $\hat{\sigma}^2$ , they could compute the optimal number of representatives under the robust mechanism. At the time of the writing of the constitution, the FF may have had some knowledge of  $F$ ,  $N$  and  $\hat{\sigma}$ , but were well aware that these parameters vary with time. The constitution should therefore allow for changes in the optimal  $n$ . In other words, the number of legislative seats should not be fixed by the constitution.<sup>20</sup>

The first-order condition for a maximum of  $\bar{W}$  with respect to  $n$ , viewed as a real number, is easy to compute and yields  $-F + (1 + (N + 1)/2n^2)\hat{\sigma}^2 = 0$ . From this we derive the following result.

---

<sup>20</sup>This does not mean that the size of the legislature should be determined arbitrarily. In our stylized model, the rule for changing the number of seats could be fixed by the constitution, while the number itself is not. In practice, it is usually possible to change the number of representatives without amending the constitution. For instance, in France the number of representatives is determined by an "organic act" which is stronger than ordinary law but weaker than the constitution.

**Proposition 3.** *With quadratic preferences, the optimal number of representatives is 1 plus the integer part of*

$$n = \hat{\sigma} \sqrt{\frac{N+1}{2F}}. \quad (22)$$

If  $n$  is smaller than 1, we choose  $n^{**} = 1$ . This appears when  $F$  is very large, or  $\hat{\sigma}$  very small. In this case, a single person (a “technocrat”) will make the public decision.<sup>21</sup> If, on the contrary,  $F$  is small, or  $\hat{\sigma}$  is very large, we get  $n^{**} = N$  (everybody is a representative, except the executive), and we obtain a direct democracy. In this latter case, the first-best is almost implemented.<sup>22</sup>

Proposition 3’s formula suggests an econometric model of the form:

$$\log(n) = \log(\hat{\sigma}) + (1/2) \log(N+1) - (1/2) \log(2F) + \epsilon, \quad (23)$$

where  $\epsilon$  is a zero-mean, random error term. This formulation is simple and natural. The three factors determining the number of representatives are: the variance of preferences, the size of the population, and the costs of representation. This simple model fits the data remarkably well, as we now show.

## 5 Empirical assessment, on political data

To empirically predict the size of representative political institutions, we have assembled a data set for a sample of 111 countries that possess a parliament or representative assemblies. The total number of representatives,  $n$ , is expressed in numbers of individuals. It includes all representatives at the national (or federal) level, e.g., the sum of the members of the lower and upper houses, when a country has a bicameral legislature. We do not count the representatives in local governments, in the member states of a federation, or in the district or city-councils. Our point of view has been to study the determinants of the sizes of national legislatures. The population size, denoted  $N$  in the following, is expressed in millions of citizens. These two pieces of information were extracted from The Europa World Year Book

---

<sup>21</sup>But the technocrat is not a dictator, because, when  $\hat{\sigma}$  is small, preferences tend to be quite similar, and there is a consensus about the optimal decision.

<sup>22</sup>In the first-best case, strictly speaking, we have  $n^* = N + 1$  (see sub-section 2.2).

(1995). To fix ideas, the United States is in the sample with  $n = 535$  and  $N = 260.341$ . The United Kingdom has 651 representatives<sup>23</sup> (i.e., MPs). France has 898 representatives (*députés plus sénateurs*). We have estimated the same model separately with data relative to the 50 US state legislatures.

Our goal here is not to "test" a normative theory but to compare the prescriptions of this theory with what can be observed in the real world. Of course, there may be reasons for which a correspondence between observed facts and normative results exists. Some countries may have chosen and adjusted the number of representatives according to efficiency considerations, trading off costs for quality of representation in a certain way. Some other countries may have just imitated a more ancient and venerable system (for instance, Japan taking inspiration from Britain and the German Empire in 1889). Some groups, including the representatives (and politicians) themselves, may of course push for increases (or reductions) in the number of representatives to promote private goals, but possibly not enough to make the normative theory totally irrelevant. There is a need for further research on this point. It is in any case interesting to compare each country with the "international norm" or "average" revealed by the log-linear regression estimated below. Deviations from this international norm are also interesting in their own right, as we will see.

## 5.1 The square-root model with world data

To get a preliminary view of the empirical relevance of the theory, we have first regressed the total number of representatives  $n$  (expressed in numbers of individuals) on population size  $N$  (expressed in millions of citizens). A first regression of the form  $n = a + bN$  yields significant estimates of  $a$  and  $b$ , but with a poor goodness-of-fit statistic (the adjusted  $R^2$  is 0.27). By contrast, a much better adjustment is obtained when, as suggested by theory,  $\log(n)$  is regressed on  $\log(N)$  plus a constant (without any constraint). We find the following result,

$$\log(n) = 4.324 + 0.41 \log(N) \tag{24}$$

(75.26)      (17.63)

---

<sup>23</sup>In the UK case, adding some 1221 peers to 651 MPs (in 1995) would have created an outlier: so we decided not to add the Peers.

In the above regression,  $t$ -statistics are between brackets. The adjusted  $R^2$  is 0.74, and the global (Fisher)  $F$ -statistic is highly significant with a value of 311.23. Moreover, the estimated constant, 4.324, and the estimated coefficient, 0.41, are both relatively close to the theoretical predictions which are  $6.561 = (1/2)(\log(10^6)/2 - \log(2))$ , and 0.5, respectively. In particular, the estimated power of  $N$  is below  $1/2$ , but not much so. The estimated constant captures some of the effect of the omitted variables. But the result is surprisingly good for such a crude regression. See Figure 1, for a plot of  $n$  against  $N$  in the studied sample.

INSERT FIG 1 HERE.

According to the theory, a more heterogeneous population should lead to larger parliaments, and countries where the cost of representation is high should have smaller ones. It is difficult to capture population heterogeneity  $\sigma^2$  and the per capita (opportunity) cost of representation  $F$  in the regression. We can only hope to find proxies for  $\sigma$  and  $F$ . We were not able to find a database, or even international comparison studies on the social cost of maintaining a representative assembly. We have checked some national accounts in order to get a sense of the costs involved. They are quite large. In the United States, for example, funding for the legislative branch rose from USD 2.8 billion in 2001 to USD 4.3 billion, requested in 2007 (a 57% growth). The average annual cost of maintaining one representative can hence be estimated in 2006 to be around USD 8 million, or 210 times the US GNP per capita. In Australia, the cost of maintaining the elected representatives in federal parliament was estimated at AD 400 million in 2004. This puts the average annual cost of maintaining one representative around AD 2 million (i.e., USD 2.6 million), more than 100 times the Australian GNP per capita. In Canada, the total cost was CD 468 million in 2004-2005. The average annual cost per representative is then CD 5.5 million (i.e., USD 4.95 million), more than 200 times the Canadian GNP per capita. None of these amounts include the costs of holding elections (i.e., campaigning and administrative costs). It is obvious that there is some variance in the unit cost of representation: in GDP per capita terms, US and Canadian representatives cost twice as much as Australian representatives.<sup>24</sup> According to

---

<sup>24</sup>This is presumably due to the fact that, contrary to their US and Canadian counterparts, Australian

the theory, this should play a role in determining the number of representatives, given that  $F$  is in fact the sum of opportunity and direct costs per representative. To capture the impact of these costs on the size of legislative bodies, we rely on several proxies. We add the logarithm of GDP per capita to the regression. We also add the logarithm of the total national tax revenue, expressed as a percentage of GDP (denoted  $TAXREV$ ). The idea is that wealthier countries and wealthier governments will not find it difficult to maintain large assemblies. The expected sign of the the tax-revenue variable is thus positive. The expected sign of the per capita GNP is ambiguous: if it acts as a proxy for the opportunity costs of representation, the coefficient might as well be negative. We also add the logarithm of the average government wage (denoted  $GOVWAGE$ ), expressed as a percentage of GDP. This variable provides an indication on the representatives' wages, and that of their staffs. It is related to the per capita cost of maintaining the assembly. The expected sign of this wage variable is thus negative. Unfortunately, we have wage data for 62 countries only. We first add the three variables sequentially to the  $\log(n)$  regression. The results are presented in Table 1.

Column (1) is just the crude regression presented above. The quality of fit increases substantially with additional controls, as indicated by the adjusted  $R^2$  of columns (2)-(5), which is around 83%. The coefficients of the GNP per capita and of the tax revenue are positive; the coefficient of government wages is negative, as predicted by the theory. We next run a regression with the three variables simultaneously, reported in column (2) of Table 1. There are only 62 countries there because of missing wage data. We next run a regression without the GNP variable because it is not significant in the regression above and one without the wage variable, because it reduces our sample by half. The coefficient estimates are fairly robust, and particularly the coefficient on  $\log(N)$ : at 0.44, its values are closer to the theoretical prediction in columns (2) and (5), possibly because the simplest log-linear model has an omitted-variable bias.

INSERT TABLE 1 HERE

---

representatives do not set their own wages and benefits (they are fixed by an independent court). On this theme, and on the US legislatures, see McCormick and Tollison (1978).

In order to measure population heterogeneity  $\sigma$ , we have tried several different variables. We first added population density (inhabitants per square kilometer, divided by 10,000) to the basic regression. The intuition is that people who live far apart do not interact much, and may differ more. According to the theory, the sign of the density coefficient should then be negative. The data, which give the number of inhabitants per square kilometer in 1996, are from the World Bank (World development indicator, 1998). As predicted by the theory, the sign of the density coefficient is negative and significant, but it is quite small in absolute value.<sup>25</sup>

It also seems reasonable to assume that countries including many different linguistic and ethnic groups are more heterogeneous. We thus add the *ethno-linguistic fractionalization index*, denoted *ELF*, as an explanatory variable.<sup>26</sup> Higher ethno-linguistic fractionalization indices may signal more heterogeneous populations and the sign of the *ELF* coefficient should then be positive.

Another variable that might reflect population heterogeneity is the Gini coefficient for each country: a large coefficient signals an unequal distribution of income in the population. If we admit that more unequal societies are more heterogeneous, everything else being equal, they should have larger representative assemblies. The coefficient of the *Gini* variable should then be positive.<sup>27</sup> When added separately to the  $\log(n)$  equation, the *ELF* and *Gini* index coefficients are both significant and *negative*, which is the “wrong” sign, as shown in columns (3)-(5) in Table 1. This seemingly contradicts the theoretical result that more heterogeneity in the population should be associated with a larger representative body. However, when we consider the joint effect of income inequality and ethno-linguistic fragmentation (i.e., the interaction  $ELF \times Gini$ ), the coefficient of this variable is both significant and positive, which is the expected sign. The net effects of *Gini* and *ELF* are negative, but the effect of income

---

<sup>25</sup>To fix ideas, a decrease of the density variable by 1 unit in the United States would result in one additional representative only.

<sup>26</sup>This index varies between 0 and 1 and measures the probability that two randomly chosen individuals do not speak the same language. The index is known for 93 of the 111 countries considered. See, Easterly and Levine (2003).

<sup>27</sup>For the sake of comparison with the *ELF* coefficient, the data, which are from World Development Indicators (1998), are divided by 100 to be normalized between 0 and 1.

inequality dampens as the level of ethno-linguistic fragmentation rises. Symmetrically, the effect of higher ethno-linguistic fragmentation is mitigated as the level of income inequality rises.

It is not clear that the variables used as regressors are satisfactory proxies for the degree of preference heterogeneity in a given country. It might well be that *Gini* and *ELF* poorly capture the relevant aspects of heterogeneity. Clearly, other variables should be included in the regression to better control for the effect of preference heterogeneity on the size of legislatures.<sup>28</sup> Further details on this regression can be found in the working paper version of this article<sup>29</sup>. In these regressions, we did not include a measure of the degree of bicameralism, but we know that this degree is likely to also play a rôle (on this point, see, e.g., McCormick and Tollison 1981, Bradbury and Cain 2001, 2002).

To sum up, the results indicate that the number of representatives  $n$  is not determined by a constant  $n/N$  (i.e., a constant sampling rate): it increases less than proportionately with the size  $N$  of the population, since according to our estimates on the 111 countries,  $n \approx \exp(4.32)N^{0.4}$  (when  $N$  is the count of inhabitants in millions). This formula yields 474 representatives per 100 million people. This finding has been shown to be robust, and supports fairly well (it is indeed a close variant of) the square-root theory of the optimal number of representatives derived above.

Our empirical exercise can also be viewed as a first step to compare the political systems of different countries, in terms of legislature sizes. Figure 2 plots the actual number  $n$  of representatives (denoted *REPRE*) against the predicted number of representatives  $\hat{n}$  (denoted *REPREF*).<sup>30</sup> This plot has been drawn with the results of regression 4 in Table 1.

---

<sup>28</sup>From the statistical point of view, *Gini* and *ELF* are also potentially endogenous variables, at least in the long run, but an instrumentation of these variables is outside the scope of the present paper. Another possibility is of course that countries with more unequal income distributions (and higher *ELF* indices) are characterized by a form of power capture by the richest (and (or) by some ethnic groups). Some of the political regimes considered in the sample are far from being ideal democracies.

<sup>29</sup>See Auriol and Gary-Bobo (2007b). The article can be downloaded from the authors' websites, <http://ces.univ-paris1.fr/membre/Gary-Bobo/>.

<sup>30</sup>Table A2, in the working paper version (Auriol and Gary-Bobo 2007b), gives the list of countries, the observed values of  $n$  and their fitted values  $\hat{n}$ , computed with the help of regression 4 in Table 1.

INSERT FIG. 2 HERE

India and Israel do not have "enough" representatives (i.e., they lie below the regression line), which contrasts with recently democratized eastern European countries that appear to have too many of them. We also find that France and Italy have "too many" representatives, whereas the United States does not have enough of them. In fact, both France and Italy have more representatives than the United States. According to our results, the US Lower and Upper House should have 807 members rather than 535.

## 5.2 Number of representatives in the US state legislatures

Using the data provided by McCormick and Turner (2001), for US state legislatures in 1996, we have tested the square-root model on the 50 US states, adding the numbers of state senators and representatives together to form the  $n$  variable. The state population (in millions) is for 2005. We find that the crude log-linear regression yields

$$\log(n) = \underset{(52.35)}{4.696} + \underset{(3.32)}{0.172} \log(N), \quad (25)$$

the adjusted  $R^2$  is equal to 0.21, and the global  $F = 14.16$ , with exactly 49 observations ( $t$ -statistics are shown in parentheses). Adjustment quality is mediocre as compared to the results obtained with global data. Among the US states, New Hampshire, Nebraska and Nevada are outliers. New Hampshire has  $n = 400$  representatives and we have removed this state from the sample.<sup>31</sup> Removing another outlier will not change results much (this will not increase the coefficient on  $\log(N)$  substantially).

We then added the population density in 2000 and the representatives' salaries, averaged for 1995-97 (known for 40 states from McCormick and Turner 2001). Without New Hampshire, this yields the following regression, which is a little closer to our square-root model:

$$\log(n) = \underset{(41.13)}{4.723} + \underset{(3.09)}{0.218} \log(N) + \underset{(2.47)}{5.11(10^{-4})} \text{Density} - \underset{(-2.03)}{0.218(10^{-6})} \text{Salary}. \quad (26)$$

---

<sup>31</sup>If we take the 50 States, the simple log-linear regression above yields a coefficient of 0.14 on  $\log(N)$ , with a  $t$  of 2.55, and  $R^2 = 0.12$ .

The adjusted  $R^2 = 0.35$ , the global  $F = 4.78$  (significant at 3%) and  $t$ -statistics are in parentheses. We already observed that the United States is an outlier among nations in the world. A strong dependence on history, and a very slow historic speed of adjustment in the number of representatives seem to be the main explanations for the low quality of the adjustment. Population has increased enormously in some US states during the 20th century, without much change in the number of state representatives. The number of federal representatives in Washington has itself been fixed by statute in 1929 (see O'Connor and Sabato 1993). The United States seems to be characterized by an extreme form of rigidity in these matters.

These results and those obtained with world data suggest that some countries have an excessive number of representatives while others have too few. It seems important to analyze the impact of having too few or too many representatives on the performance of political institutions. With too few representatives, public decisions could well be biased in favor of active minorities, to the detriment of under-represented (or less organized) groups. Casual observations also suggest that corruption could be higher in countries characterized by an “excessive” number of representatives.

### **5.3 The number of representatives and red tape**

We now examine the link between the number of representatives and barriers to business entry, entrepreneurship and trade. The public choice school offers a theory relating lobbying activity to the number of representatives (see Mueller 2003, McCormick and Tollison 1981): the influence of each representative should decrease with their numbers; lobbies would be ready to pay more to buy a vote when the number of seats in parliament is small.

To check for the presence of a possible influence of the number of representatives on variables related to lobbying activity, we consider two indices: (i) a measure of the direct cost of meeting government requirements to open a new business, expressed as a fraction of 1999 GDP per capita, denoted *SUNKCOST* (due to Djankov *et al.* 2002); (ii) and a measure of whether state interference hinders business development, denoted *STATEINTERF*

(due to Treisman 2000).<sup>32</sup> We regress the two variables on  $\log(n)$ ,  $\log(N)$ ,  $\log(GNP)$ , and many controls. The only additional variables that are significant are *DEM46*, a dummy which equals 1 if the country has been democratic in all years from 1950 until 1995, and *TRANS* indicating a former socialist country. The OLS regressions, presented in Table 2, check for correlation, not necessarily for causality. Yet, as explained above, the number of representatives is largely predetermined and rigid in most countries: it is either fixed by the constitution, or by statutes with high ranks in the hierarchy of norms, that cannot be changed easily. It is of course endogenous in the long run. Hence, the number of representatives has good chances of being "exogenous" compared to *SUNKCOST* and *STATEINTERF*, that can be changed more easily.

We ask the following question: is it true that the variables under study are in fact influenced, not by  $\log(n)$  itself, but by the *excess number of representatives*, defined as the residual of the crude log-linear regression,  $\log(n) - \log(\hat{n}) = \log(n) - (0.4) \log(N) - 4.32$  ? This hypothesis can be tested by means of the standard *F*-test of a single linear restriction, because using  $\log(n) - \log(\hat{n})$  as a control instead of  $\log(n)$  and  $\log(N)$  is tantamount to assuming that the coefficient on  $\log(N)$  is equal to  $-0.4$  times the coefficient on  $\log(n)$ .

The first dependent variable studied in Table 2, *STATEINTERF*, provides a measure of barriers to the expansion of existing firms (i.e., whether state interference hinders the development of business). According to some theories, we should observe that the barriers to business expansion are higher in countries with smaller legislatures. The rent-seeking strategies of lobbies would be easier to carry out with fewer parliamentary seats. Yet, in Table 2, column (1a) exhibits a positive and significant coefficient on  $\log(n)$ . This result is robust to the addition of many controls. It seems that the residual of the regression of  $\log(n)$  on  $\log(N)$  is in fact the appropriate explanatory variable. Column (1a) is the unconstrained version of the regression, and column (1b) is the constrained version, exhibiting a positive and significant coefficient on  $(\log(n) - \log(\hat{n}))$ . The *F*-test for the model of column (2b) against that of column (2a) yields  $(45 - 5)(13.32 - 12.32)/12.32 = 3.246$ . The critical value

---

<sup>32</sup>We also studied a measure of trade openness, denoted *FREEOP* (and taken from Barro and Lee 1994); but the results obtained are disappointing. For details, see, Auriol and Gary-Bobo (2007b).

being  $F_{95}(1, 40) = 4.08$  at the 5% level, we cannot reject the assumption that it is the excess number of representatives,  $\log(n) - \log(\hat{n})$ , which has an impact on *STATEINTERF*.

The second variable, *SUNKCOST*, measures barriers to entry for entrepreneurs (*i.e.*, barriers to the creation of new firms). This variable has been shown to be a major determinant of the size of a country’s informal sector, and also to contribute to the level of rents in the legal sector (see Auriol and Warlters 2005, Ciccone and Papaioannou 2007). Again, we have reasons to expect higher barriers to entry in countries with relatively smaller legislatures. Yet, column (2a) in Table 2 shows the opposite result: the coefficient on  $\log(n)$  is positive and significant at the 5% level. The result is again robust to the addition of many controls. It is also robust when the regression is run without France and Italy, which have been identified as outliers above.<sup>33</sup> The unconstrained regression is given by column (2a), while the constrained regression, with  $\log(n) - \log(\hat{n})$  as a regressor, is given by column (2b). Note that  $\log(n) - \log(\hat{n})$  has a significant coefficient in column (2b). The *F*-test comparing columns (2a) and (2b) is  $(71 - 5)(20.46 - 20.05)/20.05 = 1.35$ , and the critical value is  $F_{95}(1, 66) \approx 4$ , so we cannot reject the assumption that it is the residual of the crude log-linear regression which has an impact on *SUNKCOST*.

INSERT TABLE 2 HERE

The results of Table 2 suggest that it is the part of  $\log(n)$  which is unexplained by total population (*i.e.*, the excess number of representatives) which is in fact erecting higher barriers to entry. These results show a negative correlation between the number of representatives and the degree of laissez-faire (or free-market orientation) of a country. These facts suggest a possible straightforward “quantity theory” of legislative activity and meddling in the functioning of markets: the more representatives, the more people work on law and regulation, the higher their “output”, and the more they meddle in business activity. We are not aware of many theories explaining these facts. Myerson (1993) shows that electoral systems vary in their ability to reduce government corruption. The recent literature on vote-buying provides interesting theoretical insights: when the number of representatives

---

<sup>33</sup>The coefficient of the log-number of representatives is then positive and significant at the 1% level.

increases, it seems at first glance more costly to buy votes, because a larger number of bribes must be paid to secure an outcome; but if pressure groups compete to buy votes, because of strategic effects, large assemblies may in fact be more easily “buyable” than smaller ones (on this point, see Groseclose and Snyder 1996, Morgan and Vardy 2007). Related ideas are provided by Becker (1983: 388):

Cooperation among pressure groups is necessary to prevent the wasteful expenditures on political pressure that result from the competition for influence. Various laws and political rules may well be the result of cooperation to reduce political expenditures, including restrictions on campaign contributions and the outside earnings of Congressmen, the regulation of and monitoring of lobbying organizations, and legislative and executive rules of thumb that anticipate (and thereby reduce) the production of pressure by various groups.

To this list, we add the number of representatives itself. It might well be that in some countries, a small number of representatives is a long-established, endogenous response of the political system, reflecting cooperation among various forces to reduce lobbying and inefficient state interventions.

To sum up, the results in Table 2 suggest that the number of representatives really matters. Political regimes in which the rate of representation is low, the influence and “value” of each representative are correlatively high, could paradoxically be regimes in which the supply of intervention is less elastic, and the occasions for corruption more limited. These results are of course just an indication that the topic deserves more attention. Additional work is needed to check for robustness and causality.

## 6 Conclusion

We have proposed a model of a representative democracy, based on a two-stage model of constitutional design, with a constitutional and a legislative stage. This model embodies a notion of political stability of the constitution, called robustness, which emphasizes the idea that the founding fathers do not know the distribution of citizens’ preferences. From

this model, we derived a “square-root formula” for the number of representatives, stating that the optimal number should be proportional to the square root of total population. Regression work on a sample of more than a 100 countries shows that the number of national representatives is proportional to total population to the power of 0.4 : the square-root theory is verified approximately. We then find that the United States is an outlier with too few representatives, while France and Italy, for instance, have too many. The quality of fit is lower when data on the 50 US state legislatures are used. We finally cannot reject the assumption that the excess number of representatives has an impact on the degree of state interference and on an index of barriers to entry of new firms (i.e., red tape). The number of representatives thus matters and we suggest that a ‘quantity theory’ of representatives holds: more seats in parliament are associated with more red tape and more state interference in business.

ACKNOWLEDGEMENTS. We thank William Shugart II and two anonymous referees of Public Choice, Nabil Al Najjar, David Austen-Smith, Tim Besley, Helge Berger, Jacques Drèze, Gabrielle Demange, Jean-Pierre Florens, Françoise Forges, Roger Guesnerie, Arye Hillman, Mamoru Kaneko, Jean-Jacques Laffont, Jean-François Laslier, John Ledyard, Philippe Mongin and Thomas Palfrey for their help and comments, as well as seminar audiences at the Free University of Berlin, UC Dublin, Universities of Tilburg, 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. The present paper is a deeply revised version of a manuscript circulated under the same title in the year 1999 (DP no 1286 of the Center for Mathematical Studies, in Economics and Management Sciences, Northwestern University).

## 7 Appendix

The formal statement of the result used in the derivation of the theorem is as follows.

**Proposition A1.** *Assume that the set of possible public decisions  $Q$  contains at least three elements, that the separation-of-powers, subsidiarity and anonymity principles hold.*

Assume that any valuation function  $v$  is possible ( $V$  is a universal domain). Then,  $(f, t)$  is nonmanipulable if and only if the following three conditions hold:

$$f(\widehat{v}) \in \arg \max_{q \in Q} \left\{ k(q) + \sum_{i=1}^n \widehat{v}_i(q) \right\} \quad (27)$$

where  $k$  is an arbitrarily fixed function, and for all  $i = 1, \dots, n$ ;

$$t_i(\widehat{v}) = - \sum_{j \neq i} \widehat{v}_j(f(\widehat{v})) - k(f(\widehat{v})) + m(\widehat{v}_{-i}), \quad (28)$$

where  $m$  is an arbitrarily fixed function that does not depend on  $v_i$ ; and finally,

$$k, \text{ and } m \text{ are fixed in the constitution.} \quad (29)$$

See Auriol and Gary-Bobo (2007) for a proof of this result, which is an adaptation of the classic characterization of dominant strategy mechanisms, under the assumption of quasi-linear preferences. The hard part in the proof of this proposition is the “only if” part; it heavily relies on K. Roberts’ (1979) characterization theorem. It is intuitive that the requirement of dominant strategies restricts the set of admissible mechanisms in such a way, even if it is difficult to prove that these mechanisms are the only nonmanipulable ones. Note that these mechanisms are also budget-balanced by construction, because there is at least one citizen who is not a representative (*i.e.*, at least agent 0 does not report about his (her) preferences). For the sake of simplification, we further assume that the FF are constrained to choose  $f(\cdot)$  in the set defined by Proposition A1 above, even if the domain  $V$  is restricted to a particular class of utility functions.<sup>34</sup>

We now provide a short proof of the three Lemmas.

*Proof of Lemma 1.*

If  $f$  is nonmanipulable, then  $f$  maximizes  $(k(q) + \sum_{i=1}^n \widehat{v}_i(q))$  with respect to  $q$ . So, equivalently,  $f$  maximizes

$$\frac{b}{b+n} v_0(q) + \frac{n}{b+n} \bar{v}(q),$$

---

<sup>34</sup>In the case of quadratic preferences, it is well-known that there exists a fully optimal, budget-balanced, Groves mechanism: but it is a member of the same family (see Moulin (1988), chapter 8, Groves and Loeb (1975)). In the quadratic case, we can design the transfers so as to “isolate” the representatives: they can self-finance their revelation incentives.

where, by definition,  $\bar{v} = (1/n)\sum_{i=1}^n \hat{v}_i$ ,  $v_0 = k/b$ ,  $b > 0$  is a fixed parameter and  $v_0$  is an arbitrary function of  $q$ . Suppose first that  $k \neq 0$ . Since  $V$  is convex,  $\bar{v}$  belongs to  $V$ , and more generally, a given function  $w$  belongs to  $V$  if and only if there exists a probability  $P_0$  with support included in  $V$ , such that  $E_{P_0}(v) = w$ .

Next, if  $v_0 \notin V$ , then, for some  $\bar{v} \in V$ , we have

$$\left( \frac{b}{b+n}v_0 + \frac{n}{b+n}\bar{v} \right) \notin V.$$

Since  $V$  is closed, the above result is obtained if, for instance,  $\bar{v}$  is the projection of  $v_0$  on  $V$ . Then, the intersection of the segment joining  $v_0$  and  $\bar{v}$  with  $V$  itself contains only  $\bar{v}$ , and any convex combination of  $v_0$  and  $\bar{v}$  that puts a positive weight on  $v_0$  is outside of  $V$ . It follows in this case that for every probability  $P_0$  with support included in  $V$ , we would have

$$E_{P_0}(v) \neq \frac{b}{b+n}v_0 + \frac{n}{b+n}\bar{v}.$$

This contradicts the weak utilitarianism requirement. If  $k = 0$ , then we set  $b = 0$ , and  $\bar{v}$  itself trivially satisfies the latter requirement.

*Q.E.D.*

*Proof of Lemma 2.* Let  $\sum_{i=0}^N \theta_i = x + y$ , where  $x = \sum_{i=1}^n \theta_i$  and  $y = \theta_0 + \sum_{i=n+1}^N \theta_i$ . Then,

$$W_P(\alpha, \beta) = E_P \left\{ (x + y) \left( \frac{x + \alpha\beta}{n + \beta} \right) - \frac{(N + 1)(x + \alpha\beta)^2}{2(n + \beta)^2} \right\} - nF.$$

Using the fact that  $E_P(x) = n\mu_P$ ,  $E_P(y) = (N + 1 - n)\mu_P$ ,  $E_P(xy) = E_P(x)E_P(y)$  (because of independence),  $E_P(x^2) = n\sigma_P^2 + n^2\mu_P^2$  and after some elementary computations, we find

$$\begin{aligned} W_P(\alpha, \beta) &= \left( 1 - \frac{N + 1}{2(n + \beta)} \right) \frac{n\sigma_P^2}{n + \beta} + \frac{n(N + 1)}{(n + \beta)^2} \left( \frac{n}{2} + \beta \right) \mu_P^2 \\ &\quad + \frac{\alpha\beta^2(N + 1)}{(n + \beta)^2} \mu_P - \frac{\alpha^2\beta^2(N + 1)}{2(n + \beta)^2} - nF. \end{aligned} \tag{30}$$

We then take the expectation of  $W_P(\alpha, \beta)$  with respect to the prior distribution  $B$ . This yields the stated result.

*Q.E.D.*

*Proof of Lemma 3.* We first maximize the expression of  $E_B [W_P(\alpha, \beta)]$  given by Lemma 2 with respect to  $a$ . This is equivalent to maximizing  $(\alpha\hat{\mu} - \alpha^2/2)$ . Hence,  $\alpha^* = \hat{\mu}$ . Substitute next  $\alpha^* = \hat{\mu}$  in  $\bar{W}(\alpha, \beta) \equiv E_B [W_P(\alpha, \beta)]$ . We then easily obtain

$$E_B [W_P(\hat{\mu}, \beta)] = \left( n + \beta - \frac{N+1}{2} \right) \frac{n\hat{\sigma}^2}{(n+\beta)^2} + \frac{n(N+1)}{(n+\beta)^2} \left( \frac{n}{2} + \beta \right) (\hat{\mu}^2 + \hat{z}^2) + \frac{\hat{\mu}^2\beta^2(N+1)}{2(n+\beta)^2} - nF. \quad (31)$$

We finally maximize  $\bar{W}(\hat{\mu}, \beta)$  with respect to  $\beta$ . After some simplifications, we find the first-order condition

$$\frac{\partial \bar{W}(\hat{\mu}, \beta)}{\partial \beta} = \frac{n}{(n+\beta)^3} (\hat{\sigma}^2(N+1-n-\beta) - \hat{z}^2\beta(N+1)) = 0.$$

We then solve this equation for  $\beta^*$ . It is easy to check that  $\bar{W}$  is strictly quasi-concave and it follows that  $\beta^*$  is the unique global maximizer of  $\bar{W}(\hat{\mu}, \beta)$ .

*Q.E.D.*

## 8 References

- Aghion, P., and Bolton, P.(2003), Incomplete social contracts, *Journal of the European Economic Association*, 1, 38-67.
- Auriol, E., and Gary-Bobo, R. J. (1999), On the optimal number of representatives, Discussion paper no 1286, *Center for Mathematical Studies in Economics and Management Science*, Northwestern University, Evanston, Illinois.
- Auriol, E., and Gary-Bobo, R. J. (2007), On robust constitution design, *Theory and Decision*, 62, 241-279.
- Auriol, E., and Gary-Bobo, R. J. (2007b), On the optimal number of representatives, *manuscript*, <http://ces.univ-paris1.fr/membre/Gary-Bobo/>.
- Auriol, E., and Warlters, M. (2005), Taxation base in developing countries, *Journal of Public Economics*, 89, 625-646.

- Austen-Smith, D. and Banks, J. (1996), Information aggregation, rationality and the Condorcet jury theorem, *American Political Science Review*, 90, 34-45.
- Balinski, M. L., and Young H. P. (2001), *Fair representation: meeting the ideal of one man one vote*, Second edition, Brookings Institution Press, Washington, D.C.
- Barbera, S., and Jackson, M. (2004), Choosing how to choose: self-stable majority rules and constitutions, *Quarterly Journal of Economics*, 119, 1011-1048.
- Barro, R., and Lee, J-W. (1994), "Data set for a panel of 138 countries, available at [www.nber.org/ftp/barro.lee/](http://www.nber.org/ftp/barro.lee/).
- Becker, G. S. (1983), A theory of competition among pressure groups for political influence, *Quarterly Journal of Economics*, 98, 371-400.
- Bergemann, D. and Välimäki, J. (2002), Information acquisition and efficient mechanism design, *Econometrica*, 70, 1007-1033.
- Bernardo, J., and Smith, A. F. M. (1994), *Bayesian theory*, Wiley and Sons, New York.
- Besley, T., and Coate, S. (1998), Sources of inefficiency in a representative democracy: a dynamic analysis, *American Economic Review*, 88, 139-156.
- Börgers, T. (2004), Costly voting, *American Economic Review*, 94, 57-66.
- Bradbury, J. C., and Crain, W. M. (2001), Legislative organization and government spending: cross-country evidence", *Journal of Public Economics*, 82, 309-325.
- Bradbury, J. C., and Crain, W. M. (2002), Bicameral legislatures and fiscal policy, *Southern Economic Journal*, 68, 646-659.
- Buchanan, J. M., and Tullock G. (1962), *The calculus of consent. Logical foundations of constitutional democracy*, The University of Michigan Press, Ann Arbor, Michigan.
- Casella, A. (2005), Storable votes, *Games and Economic Behavior*, 51, 391-419.
- Ciccone, A, and Papaioannou, E. (2007), Red tape and delayed entry", *Journal of the European Economic Association*, 5, 459-469.

- Clarke, E. H. (1971), Multipart pricing of public goods, *Public Choice*, 8, 19-33.
- Compte, O. and Jehiel, P. (2007), Auctions and information acquisition: sealed-bid or dynamic formats?, *RAND Journal of Economics*, 38, 355–372.
- Dahl, R. A. (1970), *After the revolution: authority in a good society*, Yale University Press, New Haven.
- Dixit, A., Grossman, G. M., and Helpman, E. (1997), Common agency and coordination: general theory and application to government policy making, *Journal of Political Economy*, 105, 752-769.
- Djankov, S., La Porta, R., Lopez-de-Silvanes, F., and Shleifer, A. (2002), The regulation of entry, *Quarterly Journal of Economics*, 117, 1-37.
- The Europa world year book* (1995), Europa Publications Ltd, London, U.K.
- Feddersen, T. and Pesendorfer, W. (1998), Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting, *American Political Science Review*, 92, 23-35.
- Garelli, S. (2001), *World competitiveness yearbook*, International Institute for Management Development, Lausanne, Switzerland.
- Gary-Bobo, R.J., and Jaaidane, T. (2000), Polling mechanisms and the demand revelation problem”, *Journal of Public Economics*, 76, 203-238.
- Gerardi, D., and Yaari, L. (2008), Information acquisition in committees, *Games and Economic Behavior*, 62, 436-459.
- Gersbach, H. (1995), Information efficiency and majority decisions, *Social Choice and Welfare*, 12, 363-370.
- Gersbach, H. (2000), Public information and social choice, *Social Choice and Welfare*, 17, 25-31.
- Gersbach, H. (2009), Democratic mechanisms, *Journal of the European Economic Association*, 7, 1436-1469.

- Green, J., and Laffont, J.J. (1977), Imperfect personal information and the demand-revealing process: a sampling approach”, *Public Choice*, 29, 79-94.
- Green, J., and Laffont, J.J. (1979), *Incentives in public decision-making*, North Holland, Amsterdam.
- Groseclose, T. J., and Snyder, J. M. (1996), Buying supermajorities, *American Political Science Review*, 90, 303-315.
- Groves, T. (1973), “Incentives in teams”, *Econometrica*, 41, 617-631.
- Groves, T., and Loeb M. (1975), Incentives and public inputs, *Journal of Public Economics*, 10, 187-217.
- Hansen, M. H. (1991), *The Athenian democracy in the age of Demosthenes. Structures, principles and ideology*, Blackwell, Oxford.
- Harsanyi, J. C. (1955), Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy*, 63, 309-321, reprinted in Harsanyi J. C. (1976), *Essays in ethics, social behaviour, and scientific explanation*, Reidel, Dordrecht.
- Holmström, B. (1979), Groves schemes on restricted domains, *Econometrica*, 47, 1137-1147.
- Ledyard, J. O. (1978), Incentive compatibility and incomplete information, *Journal of Economic Theory*, 18, 171-189.
- Li, H.. (2001), A theory of conservatism, *Journal of Political Economy*, 109, 617–636.
- Manin, B. (1997), *The principles of representative government*, Cambridge University Press, Cambridge.
- Matthews, S. A. (1984), Information acquisition in discriminatory auctions”, in M. Boyer and R. Kihlstrom eds., *Bayesian models in economic theory*, North-Holland, Amsterdam, 181-207.
- McCormick, R. E., and Tollison, R. D. (1978), Legislatures as unions, *Journal of Political Economy*, 86, 63-78.

- McCormick, R. E., and Tollison, R. D. (1981), *Politicians, legislation and the economy*, Martinus Nijhoff, Boston.
- McCormick, R. E., and Turner, C. S. (2001), On legislatures and legislative efficiency wages, in W. F. Shughart II and Laura Razzolini, *The Elgar companion to public choice*, Edward Elgar, London, UK.
- Morgan, J. and P. C. Stocken (2008), Information aggregation in polls, *American Economic Review*, 98, 846-896.
- Morgan, J., and F. Vardy (2007), On the buyability of voting bodies, *manuscript*, UC Berkeley.
- 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 Willett, T.D. (1972), Representative democracy via random selection, *Public Choice*, 12, 57-68.
- Mueller, D. C. (2003), *Public choice III*, Cambridge University Press, Cambridge, UK.
- Myerson, R. B. (1993), Effectiveness of electoral systems for reducing government corruption: a game-theoretic analysis", *Games and Economic Behavior*, 5, 118-132.
- O'Connor, K. and Sabato, L. (1993), *American government, roots and reform*, McMillan, New York.
- Osborne, M. J. and Slivinski, A. (1996), A model of political competition with citizen-candidates," *Quarterly Journal of Economics*, 111, 65-96.
- Osborne, M. J., Rosenthal, J. S. and Turner, M. A. (2000), Meetings with costly participation, *American Economic Review*, 90, 927-943.
- Palfrey, T. R. and Rosenthal, H. (1985), Voter participation and strategic uncertainty, *American Political Science Review*, 79, 62-78.

- Penrose, L. S. (1946), The elementary statistics of majority voting, *Journal of the Royal Statistical Society*, 109, 53-57.
- Persico, N. (2004), Committee design with endogenous information, *Review of Economic Studies*, 71, 165-191.
- Pole, J. R. (1987), *The American constitution, for and against, the federalist and anti-federalist papers*, Hill and Wang, New York.
- Roberts, K. (1979), The characterization of implementable choice rules, chap. 18 in J-J. Laffont ed., *Agregation and revelation of preferences*, North Holland, Amsterdam.
- Stigler, G. (1976), The size of legislatures, *Journal of Legal Studies*, 5, 17-34.
- Storing, H. J. (1981), *The anti-federalist: writings by the opponents of the constitution*, selected by M. Dry from: *The complete anti-federalist*, Chicago University Press, Chicago.
- Tocqueville, A. de, (1835), *De la démocratie en Amérique, vol. 1.*, reprinted 1981, Garnier-Flammarion, Paris.
- Treisman, D. (2000), The causes of corruption: a cross-national study, *Journal of Public Economics*, 76, 399-457.
- Tullock, G. (1977), Practical problems and practical solutions, *Public Choice*, 39, 27-35.
- Weingast, B., R., Shepsle, K. A., and Johnsen, C. (1981), The political economy of benefits and costs: a neoclassical approach to distributive politics", *Journal of Political Economy*, 89, 642-664.
- World Bank (1996), *World development indicators, 1996*, Washington, D.C.
- World Bank (1998), *World development indicators, 1998*, Washington, D.C.
- World Bank (2001), *World development indicators, 2001*, CD-ROM, Washington, D.C.

TABLE 1. Dependent Variable:  $\log(n)$ 

|                   | (1)                | (2)                | (3)                  | (4)                   | (5)                               |
|-------------------|--------------------|--------------------|----------------------|-----------------------|-----------------------------------|
| Constant          | 4.32<br>(75.26)*** | 2.98<br>(7.9)***   | 5.46<br>(27.16)***   | 4.28<br>(10.0)***     | 3.98<br>(6.55)***                 |
| $\log(N)$         | 0.41<br>(17.63)*** | 0.44<br>(14.71)*** | 0.4<br>(16.12)***    | 0.41<br>(16.88)***    | 0.44<br>(12.9)***                 |
| $\log(GNP)$       |                    | 0.04<br>(1.21)     |                      | 0.04<br>(1.53)        | 0.02<br>(0.55)                    |
| $\log(TAXREV)$    |                    | 0.34<br>(2.98)***  |                      | 0.17<br>(1.96)*       | 0.24<br>(1.76)*                   |
| $\log(GOVWAGE)$   |                    | -0.12<br>(-2.08)** |                      |                       | -0.18<br>(-1.82)*                 |
| DENSITY           |                    |                    | -0.0001<br>(-2.48)** | -0.0001<br>(-2.96)*** | $-8.5 \times 10^{-5}$<br>(2.26)** |
| GINI              |                    |                    | -2.53<br>(-5.48)***  | -1.79<br>(-3.79)***   | -1.1<br>(-1.73)*                  |
| ELF               |                    |                    | -1.74<br>(-3.85)***  | -1.25<br>(-3.07)***   | -1.37<br>(-2.41)**                |
| GINI $\times$ ELF |                    |                    | 3.54<br>(3.49)***    | 2.69<br>(3.06)***     | 2.95<br>(2.43)**                  |
| No. Obs.          | 111                | 62                 | 93                   | 93                    | 55                                |
| $R^2$             | 0.74               | 0.83               | 0.83                 | 0.85                  | 0.86                              |
| Adjusted $R^2$    | 0.74               | 0.82               | 0.82                 | 0.84                  | 0.83                              |
| Sum squared Resid | 14.5               | 4.25               | 8.89                 | 7.74                  | 3.26                              |

Columns (1)–(5) were estimated by ordinary least squares. White heteroskedastic-consistent standard errors are used to calculate t-statistics, which are reported in parentheses. Significance is denoted by \*\*\* (1%); \*\* (5%); \* (10%).

TABLE 2

| Dependent Variable:     | STATEINTERF        | STATEINTERF        | SUNKCOST            | SUNKCOST            |
|-------------------------|--------------------|--------------------|---------------------|---------------------|
|                         | (1a)               | (1b)               | (2a)                | (2b)                |
| Constant                | 1.58<br>(1.45)     | 4.65<br>(6.22)***  | 0.56<br>(1.18)      | 2.32<br>(3.38)***   |
| log(n)                  | 0.47<br>(2.2)**    |                    | 0.45<br>(2.26)**    | .                   |
| log(N)                  | -0.08<br>(-0.63)   |                    | -0.25<br>(-2.29)**  | .                   |
| log(GNP)                | -0.19<br>(-2.31)** | -0.26<br>(-3.0)*** | -0.23<br>(-3.05)*** | -0.23<br>(-3.07)*** |
| DEM46                   | -0.4<br>(-2.27)**  | -0.38<br>(-1.81)*  |                     | .                   |
| TRANS                   |                    |                    | -0.34<br>(-2.16)**  | -0.3<br>(-2.07)**   |
| log(n)-log( $\hat{n}$ ) |                    | 0.51<br>(2.08)**   |                     | 0.42<br>(2.22)**    |
| No. Obs.                | 45                 | 45                 | 71                  | 71                  |
| $R^2$                   | 0.43               | 0.38               | 0.3                 | 0.28                |
| Adjusted $R^2$          | 0.37               | 0.34               | 0.25                | 0.25                |
| Sum squared Resid       | 12.32              | 13.32              | 20.05               | 20.46               |

All columns were estimated by ordinary least squares. White heteroskedastic-consistent standard errors are used to calculate t-statistics, which are reported in parentheses. Significance is denoted by \*\*\* (1%); \*\* (5%); \* (10%).

Figure 1

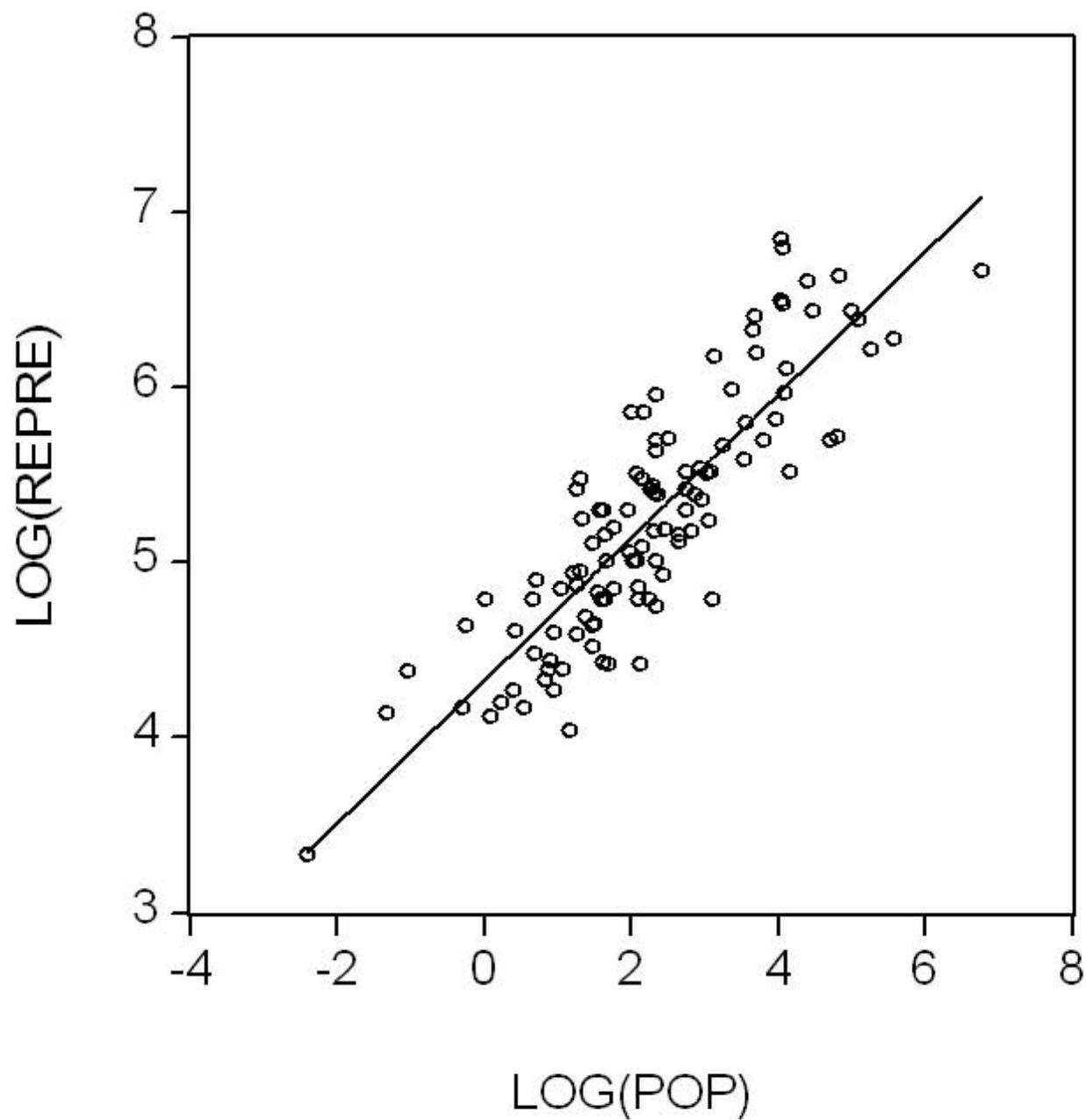


Figure 2

