On the Optimal Number of Representatives*

Emmanuelle Auriol† and Robert J. Gary-Bobo ‡

November 2007.

Abstract

We propose a normative theory of the number of representatives based on a stylized model of a representative democracy. We derive a simple formula which gives the number of representatives in parliament as proportional to the square root of total population. Econometric tests of the formula on a sample of a 100 countries yield surprisingly good results. These results suggest that the United States have too few representatives, while France and Italy have too many. The excess number of representatives is positively correlated with indicators of red tape, barriers to entrepreneurship and perceived corruption.

Keywords: Representative Democracy, Number of Representatives, Constitution Design, Incentives.

JEL N°: D7, H11, H40.

*We thank 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 and extended version of a manuscript circulated under the same title in the year 2000 (DP no 1286 of the Center for Mathematical Studies, in Economics and Management Sciences, Northwestern University).

†Toulouse School of Economics, ARQADE and IDEI, email: eauriol@cict.fr

‡Université Paris 1 Panthéon-Sorbonne, Paris School of Economics, email: garybobo@univ–paris1.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 even 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 representatives. As a counterpart, protection against the opportunistic behavior of these representatives becomes a major justification of 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 focus on a theory of the optimal number of representatives, that we also test with the help of political data.

We adopt a two-stage approach to constitutional design,[1] 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

[1] 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. The novelty of this article is that we do not assume that this distribution is common knowledge. A robust mechanism is therefore required, in the following particular sense: among nonmanipulable mechanisms, the Founding Fathers pick a decision rule that maximizes expected utility against a vague (or noninformative) prior relative to citizens' preferences. 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.[2]

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. It must work under very different distributions of preferences. 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, which trades off the direct and opportunity costs of representatives with 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 the help of a sample of more than 100 countries, and we find that our square-root theory is almost true and reasonably robust. World data is 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 US, France and Italy as outliers. The former lie below the regression line; the latter two much above. The same model does not fit the data on the 50 US State Legislatures very well. We conclude that the historical rigidity of political US institutions is likely to be the cause of this lower quality of fit. Indeed constitutional History shows that the representation ratio has constantly decreased for more than 200 years in the United States. 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

[2] See Auriol and Gary-Bobo (2007).

in the recent years. Furthermore, the number of seats in the House of Representatives has reached a ceiling of 435 in 1910.[3] According to our results, the US Lower and Upper Houses should have a total of 807 members.

We finally check for correlation of the number of representatives with some indices measuring openness to trade, the costs of setting up a new firm (i.e. "red tape"), the degree of state interference in markets, and perceived corruption.[4] The results are clearly that the number of representatives matters: it is positively and significantly correlated with state interference, red tape, and corruption. 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.

The question of the appropriate number of seats in US Parliament has been posed a long time ago by the founding fathers and opponents of the American Constitution. James Madison addressed the question in a famous passage of Federalist n°10:

> In the first place, it is to be remarked that however small the Republic may be, the Representatives must be raised to a certain number, in order to guard against the cabals of a few; and however large it may be, they must be divided to certain number, in order to guard against the confusion of a multitude.
>
> Madison, *Federalist 10* (in Pole (1987), p 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 possess their sentiments and feelings, and 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

[3]This number has been fixed by statute in 1929. See O'Connor ans Sabato (1993), p 191.

[4]We use indices constructed by Barro and Lee (1994), Djankov *et al.* (2002), Treisman (2000), and Transparency International, respectively.

> 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), p 123)

Some essential ideas are condensed in the above quotations. In the paper we interpret them as stating that there exists a tradeoff between the need to protect citizens against the dictatorship of a minority and that of reducing the costs of public decision-making. As far as we know the problem of the optimal number of legislators has been studied by a handful of economists only.[5] In contemporary writings, Buchanan and Tullock (1962) are clearly the forerunners of the approach followed here. Thinking about constitutional design, they developed a theory of the optimal constitution based on 4 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 4 variables in order to minimize the expected sum of decision-making and *external* costs of institutions. Another forerunner is Stigler (1976), who sketched a theory of the degree of representation and proposed 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 a long time ago: following Dahl (1970), Mueller *et al.* (1972) discuss random representation. Tullock (1977) went as far as to ponder over the practical possibility of using pivotal mechanisms in the US Congress to make public decisions. In the present paper, our intention is not to advocate the 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.

We are not the first to adopt a "reduced-form approach" to models politics. For instance, in Becker (1983), political parties and voting receive little attention because "they

[5]This problem is essentially distinct from that of fair representation or apportionment, that has been studied quite extensively, *e.g.* Balinski and Young (2001). Our theory is not related to L. S. Penrose's (1946) square-root formula. Penrose's formula determines the size of a country's delegation in supra-national institutions like the UN or EU, not the number of representatives itself.

are assumed mainly to transmit the pressure of active groups". Becker (1983) defines political equilibrium as a Nash equilibrium among pressure groups using expenditures on influence as strategic variables. Such an approach has both limitations and advantages. 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)).

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), Erlenmaier and Gersbach (2001). 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). An extension of optimal taxation theory to a dynamic setting in which citizens must rely on a non-benevolent politician to implement redistribution policies has been proposed by Acemoglu *et al.* (2007). Political Economy considerations are now more and more introduced in normative analysis and give rise to new constraints, in addition to revelation or incentive constraints. Our underlying philosophy has much in common with that of these recent contributions.

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, state interference and corruption indices. 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 $i = 0, 1, ..., N$. A public decision, denoted q, must be chosen in 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.

These valuation functions can be viewed as random draws in some unknown probability distribution P on a set of admissible valuation functions V.

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 executing the collective decision q. After some relabelling if necessary, agents $i = 1, ..., n$ are representatives; and agents $i = n+1, ..., N$ are passive citizens. The task of representatives is to transmit information on preferences. We assume the following.

Assumption 3. (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 to become aware of their own preferences. The two interpretations are compatible.

A representation is basically a random sample of $n \leq N$ agents (or, equivalently, a random sample of preferences $v = (v_1, ..., v_n)$).

Assumption 4. (Perfect Representation) 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 4 might seem a rather naive idealization, 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. Representation by lot existed in some societies of the past (see Hansen (1991), Manin (1997));[6] it has been discussed by political scientists (Dahl (1990)) and is still used to select juries in some countries. However, the representation biases induced by voting systems cannot be studied with the simplest form of this model. We will nevertheless continue to work with this convenient idealization.

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 $\widehat{v} = (\widehat{v}_1, ..., \widehat{v}_n)$ into Q, i.e., $q = f(\widehat{v})$, and a list of tax functions denoted $t = (t_0, t_1, ...t_N)$, 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 maximizes the function

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

[6]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)).

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:

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

where $E_P(v(.))$ is the average utility function in the population. To understand how this first-best optimum looks like, assume for example that preferences are quadratic, with a single-dimensional parameter θ, 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 specifications, (2) becomes

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

This immediately yields the optimal decision

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

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

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

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

Proposition 1. *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 the individual cost of participating in the collective decision process 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 μ 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 always be the recommended solution. 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, μ 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 benevolent planner, endowed with prior knowledge of the distribution of preferences (i.e., (μ, σ) in the quadratic example), but in a large economy with quadratic preferences, knowing μ means knowing almost everything that is useful: Democracy is 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

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 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 imposes 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 $i = n+1, ..., N$ 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 $i = 1, ...n$, 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 $v_{-i} = (v_1, ..., v_{i-1}, v_{i+1}, ..., v_n)$ 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 separation-of-powers, subsidiarity and anonymity principles, non-manipulable mechanisms must assume the following form: the decision rule $f(.)$ must maximize an objective which is the sum of an arbitrary function k and of the utilities reported by representatives, i.e.,

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

And for all $i = 1, ..., n$, representatives must be subjected to the following tranfer schedules:

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

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 $f(.)$ 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.[7]

Note that these mechanisms are budget-balanced by construction, because there is at least one citizen which 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. 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.

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

[7]On 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).

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 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 decison-maker uses a set of probability 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.[8] 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 The Founding Fathers' Objective

We first assume that the FF constrain themselves to choose a decision rule that satisfies "Weak Utilitarianism".[9]

Definition 4 (Weak Utilitarianism). The decision rule f should maximize the expected utility $E_{P_0}(v(q))$ with respect to q in Q for *some* probability distribution P_0 on 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 assumed to be Utilitarians, this requirement is very weak, because P_0 can be chosen arbitrarily.

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. If the set

[8] Bayesian statisticians have developed the theory of improper priors. See e.g. Bernardo and Smith (1994).

[9] 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.

of possible utilities V is convex, the weak utilitarianism requirement imposes to choose k of the form $k(q) = bv_0(q)$, where $b \geq 0$ is a scalar and v_0 is in V, for otherwise, the maximand $k(q)+ \sum_{i=1}^{n} \widehat{v}_i(q)$ could not be expressed as the expected social utility for some probability distribution.[10]

Formally, the social surplus function is defined as

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

The FF would like to maximize the expected value of this social welfare with respect to decision rule $f(.)$, 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).

4.2 The Founding Fathers' Beliefs

The only problem is now to give formal content to the idea that the FF will choose a nonmanipulable $f(.)$ so as to maximize $E_B E_P(W)$ under a vague (or non-informative) probability B. Such a decision rule will simply 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 which 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

[10] In other words, $bv_0(q)+ \sum_{i=1}^{n} \widehat{v}_i(q)$ is proportional to a weighted average of utilities for all b and v_0.

normality requirement, but will obtain a tractable model only if utility is assumed 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,

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

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

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

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

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

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 $\widehat{z}_k^2$ goes to $+\infty$, while $\widehat{\sigma}_k^2/\widehat{z}_k^2$ goes to zero.

To understand this definition, assume that all possible P distributions have the same variance $\sigma_P^2 = \widehat{\sigma}^2$, but that their mean μ_P is unknown to the FF. To approach complete ignorance, we let the variance of the possible means, i.e. $\widehat{z}^2$, 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 of the form $v_0(q) = \alpha q - q^2/2$ with a weight $\beta \geq 0$ and a decision rule $f^{**}(.)$, such that

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

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

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

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

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

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

Lemma 1.

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

For proof, see the appendix.

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

Lemma 2. *For given B, the optimal values of α and β are $\alpha^{**} = \widehat{\mu}$, and*

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

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 β^{**} as $\zeta \to 0$,

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

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 quadratic, we get $q^{**} = f^{**}(\widehat{v}_1, ..., \widehat{v}_n) = (1/n)\sum_{i=1}^{n} \widehat{\theta}_i$. 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).[11]

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 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 $\lim_{\zeta \to \infty} \beta^{**} = N + 1 - n$. This means that the FF know the distribution of preferences in society for sure. In this case, the recommended solution is the

[11] Again, see Auriol and Gary-Bobo (2007).

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 $E[W \mid q, \widehat{v}_1, ..., \widehat{v}_n]$ (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

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

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

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

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

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

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

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

It is then easy to check that

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

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 $\widehat{z}^2$ (which can thus be arbitrarily large). It follows that if the FF had prior information on the variance of preferences $\widehat{\sigma}^2$, 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 $\widehat{\sigma}$, but have been 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 seats in parliament should not be fixed by the constitution.[12]

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

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

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

If n is smaller than 1, we choose $n^{**} = 1$. This appears when F is very large, or $\widehat{\sigma}$ very small. In this case, a single person (a "technocrat") will make the public decision.[13] If, on the contrary, F is small, or $\widehat{\sigma}$ 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

[12] This does not mean that it should be free. In our stylized model, the rule to change 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.

[13] But the technocrat is not a dictator, because, when $\widehat{\sigma}$ is small, preferences tend to be very close, and there is a consensus about the optimal decision.

almost implemented.[14]

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

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

where ϵ 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 national representation sizes. 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 (1995). To fix ideas, the USA are in the sample with $n = 535$ and $N = 260.341$. France has 898 representatives (*députés* plus *sénateurs*).[15] We have estimated the same model separately with data relative to the 50 parliaments of the US States.

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

[14] In the first best, strictly speaking, we have $n^* = N + 1$ (see sub-section 2.2).

[15] In the case of the UK, adding some 1221 peers to 651 MPs (in 1995) would have created an outlier: so we decided not to add the Peers.

significant estimates of a and b, but with a poor R^2 (the adjusted R^2 is .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,

$$\log(n) = \underset{(75.26)}{4.324} + \underset{(17.63)}{0.41} \log(N) \tag{24}$$

In the above regression, t-statistics are between brackets. The adjusted R^2 is .74, and the global F-statistic is a highly significant 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. The estimated constant captures some of the effect of the omitted variables. But the result is surprisingly good for such a crude regression. See Fig.1, for a plot of n against N in the studied sample.

Insert Figure 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 σ^2 and the per capita (opportunity) cost of representation F in the regression. We can only hope to find proxies for σ and F. We were not able to find a database, or even international comparison studies on the social cost to maintain a representative assembly. We have checked some national accounts in order to get a sense of the costs involved. They are quite large. For instance, in the US, the legislative branch funding rose from USD 2.8 billions in 2001 to USD 4.3 billions, requested in 2007 (a 57% growth). The average annual cost of maintaining one representative can hence be estimated in 2006 around USD 8 millions, 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 millions in 2004. This puts the average annual cost of maintaining one representative around AD 2 millions (i.e., USD 2.6 millions), more than 100 times the Australian GNP per capita. In Canada, the total cost was CD 468 millions in 2004-2005. The average annual cost per representative is then CD 5.5 millions (i.e., USD 4.95 millions), more than 200 times the Canadian GNP per capita. None of these amounts include the costs of running an election (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.[16] According to 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 the legislative bodies, we rely on several proxies. We add the logarithm of the 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 a large assembly. 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 wages (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 the wage data for 62 countries only. We first add the 3 variables sequentially to the $\log(n)$ regression. The results are presented in Table 1 below.

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 the government wages is negative, as predicted by the theory. We next run a regression with the 3 variables simultaneously, reported in column (2) of Table 1. There are only 62 countries because of the 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. Coefficients are fairly robust, and

[16] This is presumably due to the fact that, contrary to their US and Canadian counterparts, Australian representatives do not set their own wages and benefits (they are fixed by an independent court).

particularly the coefficient on log(N): its values are closer to the theoretical prediction in columns (2) and (5), with a value of .44, possibly because the simplest log-linear model has an omitted-variable bias.

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

	(1)	(2)	(3)	(4)	(5)
Constant	4.32	2.98	5.46	4.28	3.98
	(75.26)***	(7.9)***	(27.16)***	(10.0)***	(6.55)***
log(N)	0.41	0.44	0.4	0.41	0.44
	(17.63)***	(14.71)***	(16.12)***	(16.88)***	(12.9)***
log(GNP)		0.04		0.04	0.02
		(1.21)		(1.53)	(0.55)
log(TAXREV)		0.34		0.17	0.24
		(2.98)***		(1.96)*	(1.76)*
log(GOVWAGE)		-0.12			-0.18
		(-2.08)**			(-1.82)*
DENSITY			-0.0001	-0.0001	$- 8.5\times10^{-5}$
			(-2.48)**	(-2.96)***	(2.26)**
GINI			-2.53	-1.79	-1.1
			(-5.48)***	(-3.79)***	(-1.73)*
ELF			-1.74	-1.25	-1.37
			(-3.85)***	(-3.07)***	(-2.41)**
GINI×ELF			3.54	2.69	2.95
			(3.49)***	(3.06)***	(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%).

In order to measure population heterogeneity σ, we have tried different variables. We first added population density (inhabitants per square kilometer, divided by 10,000) in the basic regression. The intuition is that people who leave far apart do not interact much, and may differ more. They have different ways of life, different jobs, face different climates, etc. High density countries (e.g., Japan) should present less heterogeneity than low density countries (e.g., the US). 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 (to fix ideas, a decrease of the density variable by 1 unit in the US would result in one additional representative only).

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.[17] This index varies between 0 and 1 and gives the probability that two randomly chosen individuals do not speak the same language. Higher ethno-linguistic fractionalization indices may signal more heterogeneous populations and the sign of the ELF coefficient should then be positive. This index is known for 93 of the 111 countries considered.

Another variable that might reflect population heterogeneity is the Gini coefficient of each country. Gini coefficients provide a measure of heterogeneity in the sense that 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 representatives assemblies. The coefficient of the *Gini* variable should then be positive. For the sake of comparison with the ELF coefficient, the data, which are from the World Bank (World Development Indicators, 1998), are divided by 100 to be normalized between 0 and 1. 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 by columns (3)-(5) in Table 1. This seemingly contradicts the theoretical result that more heterogeneity in the population should be associated with a larger representation. However, when we consider the joint effect of wealth inequality and ethno-linguistic fragmentation (i.e., the product $ELF \times Gini$), the coefficient of this interaction 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 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 really good proxies for the

[17]Source: Easterly and Levine, World Bank, 2003

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 capture the effect of preference heterogeneity on the size of representations. 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.

Finally, we have added several other variables, capturing some aspect of countries' history and legal institutions, as robustness checks.[18] The only additional variables that turned out to be significant are dummies, denoted $DEM46$ and $OCDE$ indicating when the country experienced 46 years of continuous democracy and a member of $OCDE$, respectively (see Table A1, in the Appendix). $DEM46$ has a negative coefficient: everything else equal, old democracies have less representatives than young democracies. This suggests that the number of seats in legislatures is characterized by institutional rigidity. $OCDE$ has a positive coefficient and seems to capture the same effects as $TAXREV$ and GNP.

To sum up, results indicate that the number of representatives n is not determined by 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 for a 100 millions of 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.

Apart from assessing whether the normative principles underlying the analysis are at work in actual political institutions, our empirical exercise can also be viewed as a first step

[18] $PERCENTPROT$ is the percentage of protestants in the country; $OECD$ indicates an $OECD$ country, TRANS a former socialist (i.e;, transition) country; $FEDERAL$ a country with a federal structure (e.g., the US, Germany); $AFRICA$ indicates a country form the African continent; $FORMBRITCOL$ indicates a former British colony and the UK; $COMLAW$ indicates a country with a Common Law system; $DEM46$ equals 1 if the country has been democratic in all 46 years (1950-1995) (see Treisman (2000)).

to compare the political systems of different countries, in terms of representation size. Table A2, in the appendix, gives the list of countries, the observed value of n and its fitted value $\widehat{n}$, computed with the help of Regression 4 in Table 1. For the missing data the simple regression 1 in Table 1 is used. These figures are reported in italic. Figure 2 plots the actual number n of representatives (denoted $REPRE$) against the predicted number of representatives $\widehat{n}$ (denoted $REPREF$). This plot has been drawn with the results of Regression 4 in Table 1.

Insert Fig. 2 here

We find that France and Italy have "too many" representatives, whereas the United States do not have "enough" of them (i.e., they lie below the regression line). In fact, both France and Italy have more representatives than the US. According to our results, the US Lower and Upper House should have 807 members instead of 535. The political US institutions seem characterized by historical rigidity. We confirm this result in the next sub-section.

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 with the 50 US States, adding the state senators and representatives together to form the n variable. The state population (in million) is for 2005. We find that the crude log-linear regression yields

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

the adjusted R^2 is equal to .21, and the global $F = 14.16$, with exactly 49 observations (t-statistics are in parentheses). Adjustment quality is mediocre as compared to the results obtained with world data. Among the US states, New Hampshire, Nebraska and Nevada are outliers. New Hampshire has a plethoric $n = 400$ representatives and we have removed this state from the sample. 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 $R^2 = .12$. Fig. 3, which includes New Hampshire, shows that the quality of the adjustment is bad. Removing another outlier will not change results much (this will not increase the coefficient on $\log(N)$ substantially).

Insert Fig.3 here

We then added the population density in 2000 and representatives' salary, 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:

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

The adjusted $R^2 = .35$, the global $F = 4.78$ (significant at 3%) and t-statistics are in parentheses. We already observed that the US are an outlier among nations in the world. A strong dependence on history, and a very slow historic speed of adjustment of 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 D.C. has itself been fixed by statute in 1929 (see O'Connor and Sabato (1993)). The US seem 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 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 the corruption level could be higher in countries characterized by an "excessive" number of representatives. We indeed find a result of that sort below.

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 lobby activity and the number of representatives (see Mueller (2003)): the influence of each representative should decrease with their number; lobbies would be ready to pay more to buy a vote when the number of seats in parliament is low. Becker's (1983) approach is also compatible with the existence of an impact of the number of representatives, if n affects the pressure groups' *influence functions* through changes in the "pressure technology".

To check for the presence of a possible influence of the number of representatives on variables related to lobbying activity, we consider in turn 3 indices: (i) a measure of trade openness, denoted $FREEOP$ (and taken from Barro and Lee (1994)); (ii) 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)); (iii) and a measure of whether state interference hinders business development, denoted $STATEINTERF$ (due to Treisman (2000)). We regress the three variables on $\log(n)$, $\log(N)$, $\log(GNP)$, and many controls. The only significant variables are $LAND$, which measures the country's land surface in millions of square kilometers, $DEM46$ and $TRANS$ dummies (defined above). The regressions, presented in Tables 2 and 3, 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 a high rank in the hierarchy of norms, that cannot be changed easily. It is of course endogeneous in the long run. Hence, the number of representatives has good chances of being "exogenous" compared to $FREEOP$, $SUNKCOST$ and $STATEINTERF$, that can be changed easily. Table 2 and 3 presents the results obtained with OLS.

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, $\log(n) - \log(\widehat{n}) = \log(n) - (0.4)\log(N) - 4.32$. This hypothesis can be tested by means of the standard F-test of a single linear restiction, because using $\log(n) - \log(\widehat{n})$ 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 $\log(n)$.

According to some theories, we should observe more restrictions to trade, and thus less openness, in countries with a smaller legislature: protectionist policies would be easier to implement with a relatively smaller parliament.[19] But the impact of $\log(n)$ on trade

[19] On the other hand, since there exists a positive correlation between an economy's exposure to international trade and the size of the government (*e.g.* Rodrik (1998), Alesina and Wacziarg (1998)), FREEOP could presumably be positively related to the number of representatives.

openness is not significant. Table 2 presents a study of the possible impact of $\log(n)$ or $(\log(n) - \log(\widehat{n}))$ on $FREEOP$. Column (1) in Table 2 is the unconstrained version of the regression, while the constrained version is given by column (3). However $\log(n) - \log(\widehat{n})$ is not significant in regression (3). Column (2) yields good results, but the significant coefficient on $\log(n)$ is likely to indirectly capture the effect of $\log(N)$, which has been removed from the regression. Total population and land surface are significant with the expected negative sign, as shown by column (4). This result is robust to the addition of many controls.[20] We find more interesting results with the other indices.

TABLE 2

Dependent Variable:	FREEOP			
	(1)	(2)	(3)	(4)
Constant	0.09	0.22	0.11	0.14
	(1.3)	(3.8)***	(3.33)***	(5.39)***
log(n)	0.01	-0.024		
	(0.83)	(-2.39)**		
log(N)	-0.02			-0.02
	(-3.09)***			(-3.09)***
log(GNP)	0.015	0.017	0.02	0.02
	(3.76)***	(4.65)***	(3.05)***	4.48***
LAND	-1.6×10^{-5}	-1.89×10^{-5}	-2.19×10^{-5}	-1.65×10^{-5}
	(-5.3)***	(-5.31)***	(-5.5)***	(-5.37)***
DEM46	0.05	0.054	0.048	0.05
	(2.51)**	(2.53)***	(2.53)**	(2.69)***
log(n)-log($\widehat{n}$)			0.01	
			(0.49)	
No. Obs.	67	67	67	67
R^2	0.65	0.61	0.57	0.64
Adjusted R^2	0.62	0.58	0.54	0.62
Sum squared Resid	0.13	0.15	0.16	0.13

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%).

The first dependent variable studied in Table 3 , $STATEINTERF$, provides a measure of barriers to business for existing firms (i.e., whether State interference hinders the development of business). According to some theories, we should observe larger barriers

[20] The result holds when we sequentially add: AFRICA, COMMONLAW, ELF, FEDERAL, FORMBRITCOL, PERCENTPROT, GOVWAGE. None of these variables have a coefficient significantly different from zero.

to business in countries with a smaller legislature. The rent-seeking strategies of lobbies would be easier to carry out with a smaller number of seats in parliament. Yet, in Table 3, column (1a) exhibits a positive and significant coefficient on $\log(n)$. This result is robust to the addition of many controls.[21] 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 $(\log(n) - \log(\widehat{n}))$. The F-test testing the model of column (2b) against that of column (2a) yields $(45 - 5)(13.32 - 12.32)/12.32 = 3.246$. The critical value being $F_{95}(1, 40) = 4.08$ at the 5% level, we cannot reject the assumption that it is the excess number of representatives, $\log(n) - \log(\widehat{n})$, 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 3 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.[22] It is also robust when the regression is run without France and Italy, which have been identified as outliers above.[23] The unconstrained regression is given by column (2a), while the constrained regression, with $\log(n) - \log(\widehat{n})$ as a regressor, is given by column (2b). Note that $\log(n) - \log(\widehat{n})$ 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 $SUNKCOST$.

[21] The result holds when we sequentially add, AFRICA, ELF, COMMONLAW, LAND, FEDERAL, FORMBRITCOL, PERCENTPROT, GINI, GOVWAGE, OECD, TRANS, to the regression.None of theses variables have a significant coefficient.

[22] The result holds when we sequentially add to the regression:AFRICA, COMMONLAW, DEM46, LAND, FEDERAL, FORMBRITCOL, GINI, GOVWAGE. None of these variables have a coefficient which is significantly different from zero.

[23] The coefficient of the log-number of representatives is then positive and significant at the 1% level.

TABLE 3

Dependent Variable:	STATEINTERF	STATEINTERF	SUNKCOST	SUNKCOST
	(1a)	(1b)	(2a)	(2b)
Constant	1.58	4.65	0.56	2.32
	(1.45)	(6.22)***	(1.18)	(3.38)***
log(n)	0.47		0.45	
	(2.2)**		(2.26)**	
log(N)	-0.08		-0.25	
	(-0.63)		(-2.29)**	
log(GNP)	-0.19	-0.26	-0.23	-0.23
	(-2.31)**	(-3.0)***	(-3.05)***	(-3.07)***
DEM46	-0.4	-0.38		
	(-2.27)**	(-1.81)*		
TRANS			-0.34	-0.3
			(-2.16)**	(-2.07)**
log(n)-log($\widehat{n}$)		0.51		0.42
		(2.08)**		(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%).

The results of Table 3 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 causing a higher level of 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. We are not aware of a theory explaining these facts, but they suggest a possible straightforward "quantity theory" of the legislators' activity and meddling in the operation 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. A related idea is provided by Becker (1983, p 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, the low 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. If these ideas are true, there are some chances that we will find a positive correlation between the number of representatives and corruption.

5.4 The Number of Representatives, and Corruption

We finally study the possible impact of the number of representatives on the level of perceived corruption in a country. Treisman (2000) has analyzed the causes of perceived corruption in a cross-national study, using $TISCORE$, a well-known corruption index computed by *Transparency International*. Treisman (2000) finds that countries with more developed economies, protestant traditions, histories of British rule, long exposure to democracy and higher imports are less "corrupt". Federal states are more "corrupt". We have reproduced Treisman's regressions with additional controls and chiefly, we added the number of representatives in the regressions.

The results are presented in Table 4; they strongly suggest that the number of representatives has a positive impact on corruption. Columns (1), (2), and (3) of Table 4 reproduce columns (2), (3), and (4), respectively, in Treisman (2000, table 3, p 58), for the year 1996, with the addition of the number of representatives n.[24] With these changes, contrary to Treisman's results, $FORMBRITCOL$ is not significant, while $DEM46$ becomes significant.[25]

[24] The dataset is not exactly the same: more countries are included in our results. Some variables are different: we use $FREEOP$ instead of the imports/GNP ratio in Treisman's work. $RAWMAT$ is the percentage of raw materials in the country's exports.

[25] This is true, with or without the number of representatives.

TABLE 4

Dependent Variable: TISCORE				
	(1)	(2)	(3)	(4)
Constant	14.94	13.69	13.82	12.17
	(13.28)***	(11.31)***	(10.14)***	(10.21)***
n	0.0018	0.0017	0.002	0.0023
	(2.61)**	(2.07)**	(1.93)*	(3.51)***
COMLAW	-0.18	-0.17	-0.12	
	(-0.45)	(-0.46)	(-0.32)	
FORMBRITCOL	-0.72	-0.56	-0.64	
	(-1.61)	(-1.28)	(-1.46)	
PERCENTPROT	-0.026	-0.014	-0.01	-0.01
	(-3.85)***	(-2.15)**	(-1.77)*	(-3.22)***
ELF	-0.24	-0.2	-0.42	
	(-0.33)	(-0.31)	(-0.57)	
RAWMAT	-0.008	-0.013	-0.01	
	(-0.69)	(-1.19)	(-1.16)	
log(GNP)	-1.29	-1.02	-1.03	-0.81
	(-9.24)***	(-6.8)***	(-5.85)***	(-5.08)***
FEDERAL		0.95	0.93	
		(2.35)**	(2.06)**	
DEM46		-1.53	-1.72	-1.63
		(-3.34)***	(-3.57)***	(-4.73)***
FREEOP			0.3	
			(0.13)	
AFRICA				-1.12
				(-2.81)***
OECD				-0.76
				(-1.95)*
DENSITY				-0.0005
				(-4.73)***
No. Obs.	69	69	56	79
R^2	0.70	0.75	0.76	0.77
Adjusted R^2	0.67	0.71	0.71	0.75
Sum squared Resid	115.03	96.71	80.65	101.74

Columns (1) to (4) 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%).

But the novelty is of course that the coefficient on n is positive and significant in the 4 variants presented in Table 4. The result is thus fairly robust.[26] We are tempted to interpret

[26] The significant coefficient on n in column (4) is robust to the inclusion of many controls, *i.e.*, ELF, COMLAW, LAND, FEDERAL, FORMBRITCOL, GINI, GOVWAGE, TRANS. None of these variables have significant coefficients.

this fact as a confirmation of the “quantity theory” sketched above. More representatives produce more red tape and induce more corruption.

To sum up, putting together the results of Tables 3 and 4, suggests 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 subject 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 constitution 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 almost true. We then find that the USA 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 is 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 itself has a significant, positive impact on the degree of perceived corruption. The number of representatives thus matters and we suggest that a “quantity theory” of representatives hold: more seats in parliament are associated with more red tape, more state interference in business, and more corruption.

7 References

Acemoglu, Daron, Golosov, Michael, and Aleh Tsyvinski (2007), "Political Economy of Mechanisms", *manuscript, MIT*, Cambridge, Massachusetts.

Aghion, Philippe, and Patrick Bolton (2003), "Incomplete Social Contracts", *Journal of the European Economic Association*, 1, 38-67.

Alesina, Alberto, and Romain Wacziarg (1998), "Openness, Country Size, and Government", *Journal of Public Economics*, 69, 305-321.

Auriol, Emmanuelle, and Robert J. Gary-Bobo (2007), "On Robust Constitution Design", *Theory and Decision*, 62, 241-279.

Auriol, Emmanuelle, and Michael Warlters (2005), "Taxation Base in Developing Countries", *Journal of Public Economics*, 89, 625-646.

Balinski, Michel L., and H. Peyton Young (2001), *Fair Representation: Meeting the Ideal of One Man One Vote*, Second edition, Brookings Institution Press, Washngton, D.C.

Barbera, Salvador, and Matthew Jackson (2004), "Choosing How to Choose: Self-Stable Majority Rules and Constitutions", *Quarterly Journal of Economics*, 119, 1011-1048.

Barro, Robert, and J-W Lee (1994), "Data Set for a Panel of 138 Countries", available at *www.nber.org/ftp/barro.lee/*.

Becker, Gary S. (1983), "A Theory of Competition Among Pressure Groups for Political Influence", *Quarterly Journal of Economics*, 98, 371-400.

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

Besley, Timothy, and Stephen Coate, (1998), "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis", *American Economic Review*, 88, 139-156.

Buchanan, James M., and Gordon Tullock (1962), *The Calculus of Consent. Logical Foundations of Constitutional Democracy*, The University of Michigan Press, Ann Arbor, Michigan.

Casella, Alessandra (2005), "Storable Votes", *Games and Economic Behavior*, 51, 391-419.

Ciccone, Antonio, and Elias Papaioannou (2007), "Red Tape and Delayed Entry", *Journal of the European Economic Association*, 5, 459-469.

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

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

Dixit, Avinash, Grossman, Gene M., and Elhanan Helpman (1997), "Common Agency and Coordination: General Theory and Application to Government Policy Making", *Journal of Political Economy*, 105, 752-769.

Djankov, Simeon, La Porta, Rafael, Lopez-de-Silvanes, Florencio, and Andrei Shleifer (2002), "The Regulation of Entry", *Quarterly Journal of Economics*, 117, 1-37.

Erlenmaier, Ulrich, and Hans Gersbach (2001), "Flexible Majority Rules", *CES Ifo Working Paper*, n° 464, CES Ifo, Munich, Germany.

The Europa World Year Book (1995), Europa Publications Ltd, London, U.K.

Garelli, Stephane (2001), *World Competitiveness Yearbook*, International Institute for Management Development, Lausanne, Switzerland.

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

Green, Jerry, and Jean-Jacques Laffont (1977), "Imperfect Personal Information and the Demand-Revealing Process: A Sampling Approach", *Public Choice*, 29, 79-94.

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

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

Groves, Theodore. and Martin Loeb (1975), "Incentives and Public Inputs", *Journal of Public Economics*, 10, 187-217.

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

Harsanyi, John C. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, 63, 309-321, reprinted in Harsanyi (1976), *Essays in Ethics, Social Behaviour, and Scientific Explanation*, Reidel, Dordrecht.

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

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

Manin, Bernard (1997), *The Principles of Representative Government*, Cambridge University Press, Cambridge.

McCormick, Robert E., and Chad S. Turner (2001), "On Legislatures and Legislative Efficiency Wages", in William F. Shughart II and Laura Razzolini, *The Elgar Companion to Public Choice*, Edward Elgar, London, UK.

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

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

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

Mueller, Dennis C. (2003), *Public Choice III*, Cambridge University Press, Cambridge, UK.

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

Osborne, Martin J. and Al Slivinski (1996), "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics,* 111, 65-96.

Penrose, Lionel. S. (1946), "The Elementary Statistics of Majority Voting", *Journal of the Royal Statistical Society*, 109, 53-57.

Pole, J.R. (1987), *The American Constitution, For and Against, the Federalist and Anti-Federalist Papers*, Hill and Wang, New York.

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

Rodrik, Dani (1998), "Why Do More Open Economies Have Bigger Government?", *Journal of Political Economy*, 106, 997-1032.

Stigler, George (1976), "The Size of Legislatures", *Journal of Legal Studies*, 5, 17-34.

Storing, Herbert 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, Alexis de (1835), *De la démocratie en Amérique, vol. 1.*, reprinted 1981, Garnier-Flammarion, Paris.

Treisman, Daniel (2000), "The Causes of Corruption: A Cross-National Study", *Journal of Public Econmics*, 76, 399-457.

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

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.

8 Appendix

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

Proposition A1. *Assume that there are at least three possible decisions, that the separation-of-powers, subsidiarity and anonymity principles hold. Assume that any utility function v is possible (V is a universal domain). Then, (f,t) is nonmanipulable if and only if the following three conditions hold:*

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

where k is an arbitrary fixed function, and for all $i = 1, ..., n$;

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

where m is an arbitrary fixed function that doesn't depend on v_i; and finally,

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

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 which is not a representative (*i.e.*, at least agent 0 does not report about his (her) preferences). We further assume that the FF are constrained to choose $f(.)$ in the set defined by Proposition A1 above, even if the domain V is restricted to a particular class of utility functions (to keep matters simple.[27]) We now provide a short proof of the Lemmas.

[27]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.

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

$$W_P(a,b) = E_P\left\{(r+s)\left(\frac{r+ab}{n+b}\right) - \frac{(N+1)(r+ab)^2}{2(n+b)^2}\right\} - nF.$$

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

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

We then take the expectation of $W_P(a,b)$ with respect the the prior distribution B. This yields the stated result.

Q.E.D.

Proof of Lemma 2. We first maximize the expression of $E_B\left[W_P(\alpha,\beta)\right]$ given by Lemma 1 with respect to a. This is equivalent to maximizing $(\alpha\widehat{\mu}-\alpha^2/2)$. Hence, $\alpha^*=\widehat{\mu}$. Substitute next $\alpha^*=\widehat{\mu}$ in $\overline{W}(\alpha,\beta)\equiv E_B\left[W_P(\alpha,\beta)\right]$. We then easily obtain

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

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

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

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

Q.E.D.

TABLE A1. Dependent Variable: $log(n)$

	(1)	(2)	(3)	(4)	(5)
Constant	3.4***	4.21***	4.03***	4.22***	4.02***
	(7.16)	(9.97)	(9.34)	(10.21)	(11.67)
log(N)	0.45***	0.41***	0.4***	0.4***	0.39***
	(15.27)	(16.79)	(16.66)	(16.16)	(18.32)
log(GNP)	0.08**	0.07**	0.06*	0.04	0.02
	(2.11)	(2.35)	(1.69)	(1.09)	(0.76)
log(TAXREV)	0.3***	0.16*	0.12	0.16*	0.15*
	(3.19)	(1.94)	(1.37)	(1.94)	(1.91)
DENSITY	-0.0001**	-0.0001***	-0.0001***	- 0.0001***	- 0.0001***
	(-2.27)	(-3.44)	(-2.82)	(-2.67)	(-2.9)
GINI	-1.63***	-2.1***	-1.37**	-1.57***	-0.71**
	(-3.49)	(-4.32)	(-2.24)	(-2.72)	(-2.21)
ELF	-1.22***	-1.31***	-0.83*	-0.98**	
	(-2.67)	(-3.15)	(-1.74)	(-2.07)	
GINI×ELF	2.7***	2.89***	1.69	2.18**	
	(2.74)	(3.26)	(1.6)	(2.16)	
DEM46	-0.19	-0.21*	-0.22*	-0.24**	-0.2*
	(-1.27)	(-1.92)	(-1.82)	(-2.31)	(-1.96)
PERCENTPROT	-0.0006				
	(-0.34)				
FEDERAL	0.009				
	(0.085)				
COMLAW	-0.13				
	(-1.1)				
FORMBRITCOL	0.06				
	(0.54)				
OECD			0.24**	0.26**	0.31***
			(2.08)	(2.13)	(3.11)
TRANS			0.10		
			(0.85)		
AFRICA			0.14		
			(1.28)		
No. Obs.	71	93	93	93	111
R^2	0.88	0.86	0.87	0.86	0.83
Adjusted R^2	0.86	0.84	0.85	0.85	0.82
Sum squared Resid	4.51	7.4	6.9	7.06	9.5

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%). PERCENTPROT is the percentage of protestants in the country; OECD indicates an OECD country, TRANS a former socialist (i.e, transition) country; FEDERAL a country with a federal structure (e.g., the US, Germany); AFRICA indicates a country form the African continent; FORMBRITCOL indicates a former British colony and the UK; COMLAW indicates a country with a Common Law system; DEM46 equals 1 if the country has been democratic in all 46 years (1950-1995) (source Treisman (2000)).

Table A2: Actual and fitted number of representatives
Value $\widehat{n}$ has been computed based on regression 4 in Table 1. For missing data, reported in italic, regression 1 in Table 1 has been used.

COUNTRY	n	$\widehat{n}$
Albania	140	113.21610
Angola	220	187.31671
Argentina	329	285.41909
Armenia	190	*129.68857*
Australia	219	308.46221
Austria	247	293.03681
Azerbaijan	350	*170.93651*
Bangladesh	300	460.99839
Belgium	221	262.90633
Benin	83	121.36892
Bolivia	157	140.49486
Bosnie-Herz	240	*128.93847*
Brazil	594	463.96416
Bulgaria	240	213.43878
Burkina-Faso	227	155.45687
Cambodia	120	144.75544
Cameroon	180	161.83940
Canada	399	336.39994
Cl Africa Rep.	85	79.448245
Chile	167	178.12343
Colombia	267	233.24201
Costa Rica	57	114.81261
Coast Ivory	175	188.57065
Croatia	201	*143.04017*
Czech Rep.	281	251.82664
Denmark	175	243.65874
Dominican Rep.	150	138.41592
Egypt	664	460.33356
El Salvador	84	113.00308
Equ. Guinea	80	39.270870
Estonia	101	*89.256839*
Fiji	104	63.739114
Finland	200	218.41965
France	898	545.84014
Gabon	120	80.252787
Germany	740	661.84103
Ghana	200	182.14392
Greece	300	231.76985
Grenada	28	24.882613

COUNTRY	n	$\widehat{n}$
Guatemala	116	142.01617
Guyana	65	69.559220
Honduras	128	117.29668
Hungary	386	277.03339
Iceland	63	54.259966
India	790	860.18700
Indonesia	500	542.36577
Ireland	226	165.18585
Israel	120	175.02468
Italy	945	570.24767
Jamaica	81	103.79495
Japan	763	704.17869
Jordan	120	141.44161
Kazakhstan	177	*238.86700*
Kenya	188	239.66050
Rep. of Korea	299	415.62166
Kyrgyr. Rep.	105	*139.26255*
Latvia	100	*110.94319*
Lebanon	128	84.179405
Lesotho	65	82.219469
Lithuania	141	*129.08048*
Macedonia	120	*98.900478*
Madagascar	138	157.81119
Malawi	177	155.33208
Malaysia	212	239.34138
Mali	129	145.56781
Mauritania	135	90.158273
Mauritus	62	67.472499
Mexico	628	389.86490
Moldava	104	*137.72470*
Mongolia	76	112.42436
Mozambiq.	250	165.65331
Nanibia	72	91.836596
Nepal	255	198.00898
Netherlands	225	325.95926
New Zealand	99	152.13716
Nicaragua	92	116.27401
Niger	83	127.39640
Norway	165	219.39632
Pakistan	304	448.11270
Panama	72	89.919477
Papua Guin.	109	121.11107
Paraguay	125	105.57714

COUNTRY	n	$\widehat{n}$
Peru	120	239.17824
Phlippines	250	351.30956
Poland	560	476.16203
Portugal	230	247.83124
Romania	484	349.69833
Russian Fed.	628	*581.97390*
Senegal	120	153.68839
Singapore	81	70.236058
Slovakia	150	*149.63106*
Slovenia	88	100.84635
South Africa	490	356.47748
Spain	605	414.31214
Sweden	349	296.31429
Switzerland	200	*166.87908*
Tajikistan	181	*154.28113*
Tanzania	291	*284.14307*
Thailand	391	361.26580
Trinidad Tob.	67	84.508863
Tunisia	163	191.99075
Turkey	450	371.82386
Ukraine	338	*379.52819*
United King.	651	527.30617
USA	535	807.49016
Uruguay	130	139.27179
Usbekistan	250	*265.43888*
Venezuela	248	230.87274
Yemen	301	182.78597
Zambia	150	148.53332
Zimbabwe	150	178.27591

Figure 1

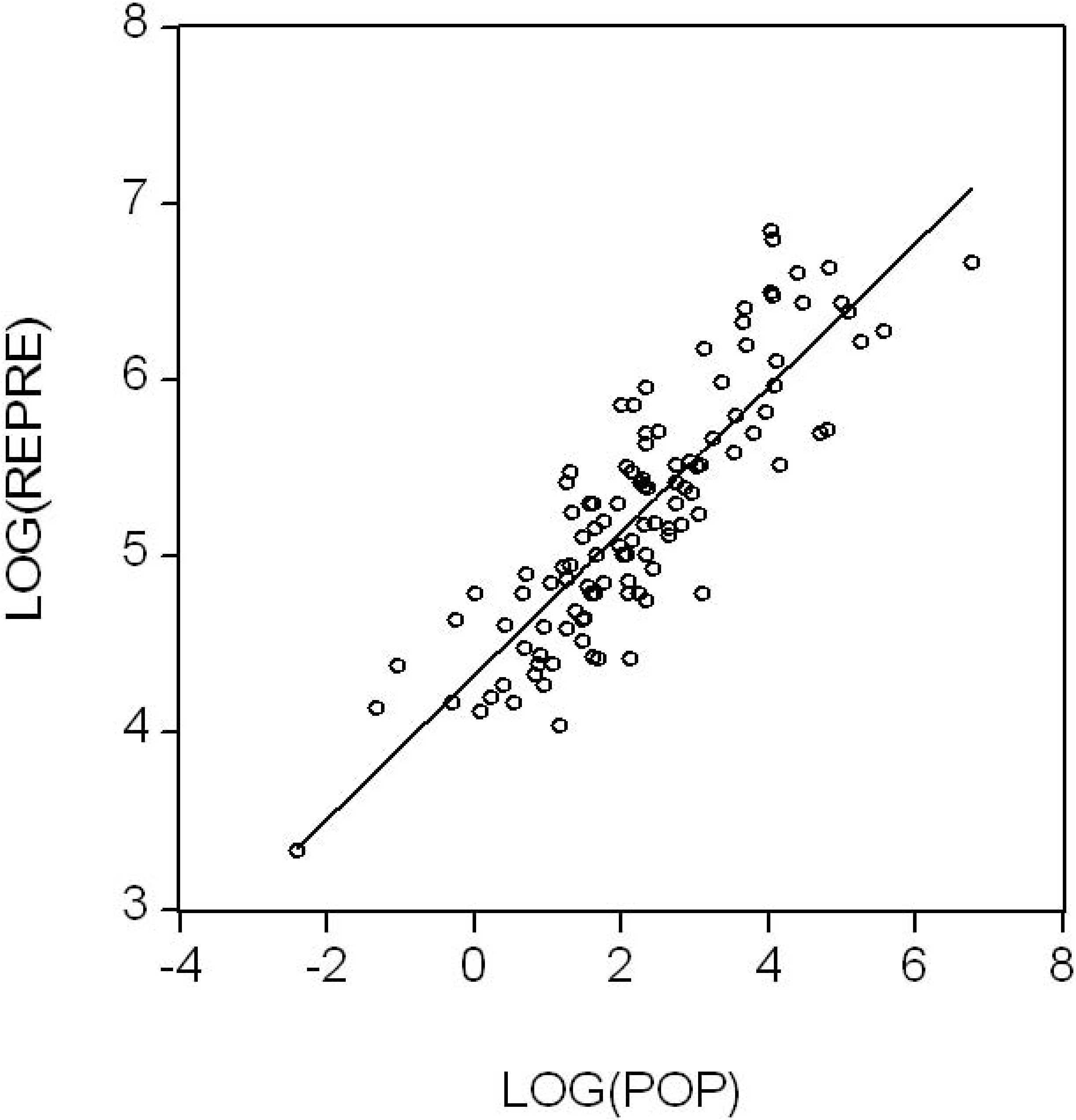
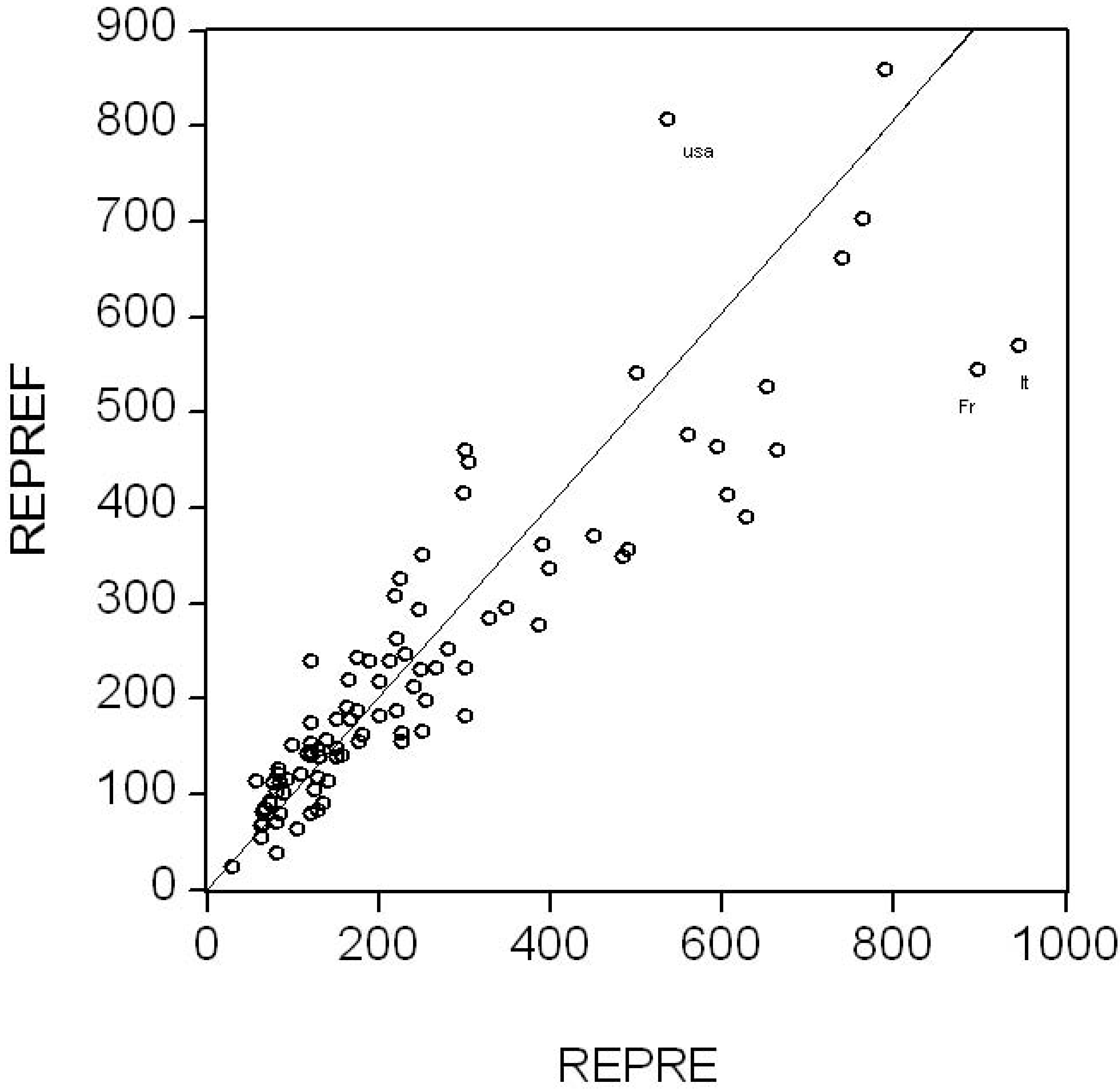
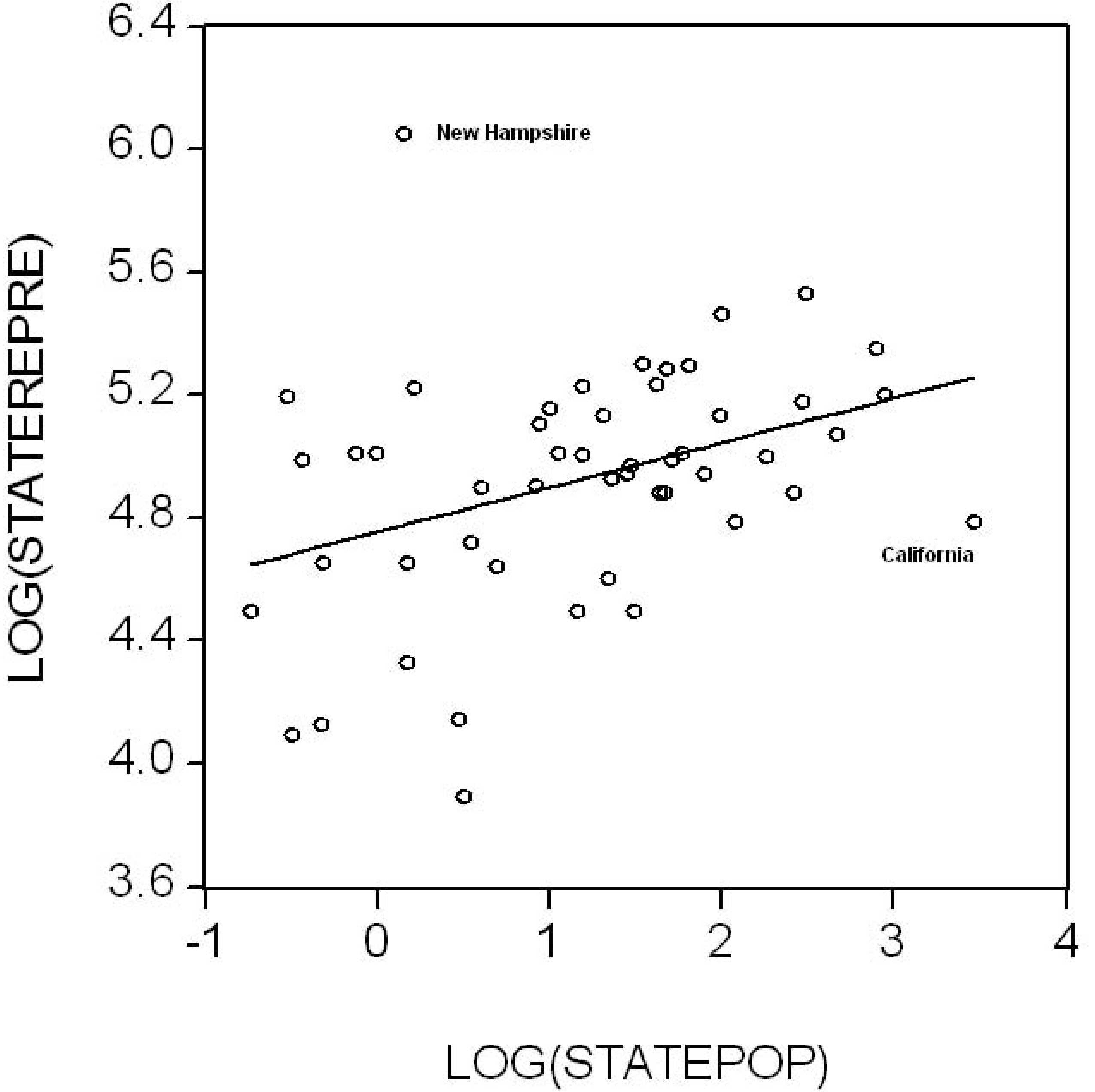


Figure 2



900
800
700
600
500
400
300
200
100
0
REPREF
usa
Fr
It
0
200
400
600
800
1000
REPRE

Figure 3



6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
LOG(STATEREPRE)
New Hampshire
California
-1
0
1
2
3
4
LOG(STATEPOP)