Capitalism and Society

Volume 8, Issue 1 2013 *Article 1*

Excessive Ambitions (II)

Jon Elster
Columbia University

Recommended Citation:
Elster, Jon (2013) "Excessive Ambitions (II)," *Capitalism and Society*: Vol. 8 : Iss. 1, Article 1.

Excessive Ambitions (II)

Jon Elster

Abstract

Following an earlier article criticizing excessive uses of rational-choice modeling and statistical analyses in the social sciences, the present article argues that much of normative political theory, notably many theories of optimal institutional design, also suffer from various forms of overreaching. It is argued that both attempts to design democratic institutions that will track independently defined good outcomes and attempts to choose good democratic decision-makers are bound to fail. The article also presents a positive alternative, inspired by Jeremy Bentham's Political Tactics: institutional designers should reduce as much as possible the impact of self-interest, emotion, prejudice and cognitive bias on the decision-makers, and then let the chips fall where they may.

1. Introduction

In an earlier article, "Excessive ambitions", I argued that large bodies of social science are permeated by explanatory hubris. Economists and political scientists, in particular, rely on deductive models and statistical tools that are vastly less robust and reliable than their practitioners claim. In this article, I argue that large bodies of normative political theory also suffer from excessive ambitions.[1*] In part, the problem arises because of reliance on inadequate causal theories. In part, however, it is due to inadequate normative theories.

I shall not discuss normative political philosophy at the general level, to assess the value of the approaches proposed by John Rawls, Ronald Dworkin, John Harsanyi, Robert Nozick, Thomas Scanlon, Amartya Sen, and others. My aim is more limited: to assess the capacity of a normative theory to identify good outcomes and to design institutions that can track these outcomes. I shall mainly consider three democratic institutions: juries, elected assemblies, and electoral bodies. Although some of my arguments carry over to non-democratic institutions such as constitutional courts, central bank committees, and the FDA advisory committees, I shall touch on these bodies only in passing.

My argument is inspired by the political writings of Jeremy Bentham and by John Hart Ely's *Democracy and Distrust*. While I do not embrace – and in fact shall argue against – Bentham's utilitarianism, I do adopt, with some qualifications, his argument that when designing institutions, "the end is so to speak of a *negative character*. […] The art of the legislator is limited to the prevention of everything that might prevent the development of their liberty and their intelligence."[2] From Ely's work I can cite the following statement: "The approach to constitutional adjudication recommended here is akin to what might be called an 'antitrust' as opposed to a 'regulatory' orientation to economic affairs – rather than to dictate substantive results it intervenes only when the 'market', in our case the political market, is systematically malfunctioning".[3]

The natural standard for assessing the institutions I discuss - jury trials, political assemblies and electoral systems - might seem to be whether they tend to *produce good outcomes*. One difficulty is that of defining what counts as good outcomes. Political and legal philosophers have proposed a number of answers, and it is not clear how we can choose among them. Assuming that we opt for one of them, we then have to assess institutions in light of

[*] Just as my earlier article (Elster 2009) provided the analytical skeleton of Elster (2007), the present one offers a précis of Elster (2013). This book provides numerous examples and illustrations as well as a more nuanced theoretical argument than what can be presented in an article.

[2] Bentham (1999), p.15. Bentham's pedantic and impractical mind combines with his barbaric writing style, especially in his later writings, to make reading him an arduous task. Without having become any kind of Bentham scholar, I have found that the investment is worth the effort.

[3] Ely (1980), p. 102-3.

their tendency to promote the chosen conception of goodness. To do so, we need a causal theory. Social scientists have produced a number of such theories, and it is not clear how we can choose among them. Over the last few years the *Financial Times* has probably published dozens of different theories of what caused the current financial crisis, together with equally many remedy proposals. The number of theories and proposals is matched only by the certainty with which each of them is propounded. The large number of competing theories should, however, undermine our confidence that *any* of them is right.[4]

I believe that this *double indeterminacy* – of plausible-sounding but unprovable normative views, and of plausible-sounding but unprovable causal theories – has led to a deep disillusionment in public debates. References to the common good or the general interest are routinely dismissed as cant. There is, in fact, hardly any policy proposal – however partial in its origins – that cannot be justified on impartial grounds. Blue-collar and white collar workers tend to invoke different norms of equity, the former arguing that work should be rewarded according to the burdens imposed on the workers and the latter that wages should reflect skills and benefit to society.[5] A study of the appeal to principles of equity in allocating the burdens of climate change abatement found that "the economic costs implied by the respective equity rules explain the perceived support by EU, Russia, and the USA."[6] These examples could be multiplied indefinitely.

Tocqueville's observation still rings true:

> [A] politician first tries to identify his own interests and finds out what similar interests might be joined with his. He then casts about to discover whether there might not by chance exist some doctrine or principle around which this new association might be organized, so that it may present itself to the world and gain ready acceptance.[7]

He might also have accepted the following variation on his theme:

> A politician first tries to identify his own interests and finds out what similar interests might be joined with his. He then casts about to discover whether there might not by chance exist some causal theory or statistical model according to which the promotion of this interest coincides with the general interest, so that it may present itself to the world and gain ready acceptance.

[4] Føllesdal (1979), p. 405-6.

[5] Hyman and Brough (1975), p. 49.

[6] Lange et al. (2010), p. 367.

[7] Tocqueville (2004), p. 202.

Thus Chicago-style economists find that for each murderer who is executed, up to 18 murders are not committed, and also that the right to carry concealed handguns saves lives. To buttress these claims, they engage in fragile statistical analyses that amount to little more than data-mining or curve-fitting.[8] Although in their case it is not a matter of self-interest, but of ideology, the general point is the same. To justify a policy to which one is attached on self-interested or ideological grounds, one can shop around for a causal or statistical model just as one can shop around for a principle. Once it has been found, one can reverse the sequence and present the policy as the conclusion. This process can occur anywhere on the continuum between deception and self-deception (or wishful thinking), usually no doubt closer to the latter.

I shall *argue against positive institutional design*, aimed at creating institutions that will (or tend to) produce good decisions, select good decision-makers, or create good decision-making bodies. The criteria for what counts as a good decision by juries, elected assemblies and electoral bodies are often indeterminate. When they are not, we cannot verify whether they are satisfied. Although some decision-makers are no doubt better than others, we cannot reliably select them through institutional design. Although one can make plausible qualitative arguments for how to design good decision-making bodies, more ambitious quantitative models fail because one cannot verify empirically whether the relevant conditions obtain.

I shall also *argue for negative institutional design*, as a more modest and more robust way of reaching good collective decisions. Although the reference to good decisions might seem to contradict the claims of the previous paragraph, properly understood it does not. I shall argue for various reforms on the grounds that they *can't hurt and might help*. Specifically, one should insulate decision-makers as much as possible from the influences of self-interest, passion (emotion or intoxication), prejudice and cognitive bias. Once that has been done, one should let the chips fall where they may. Extending the typology of John Rawls[9] to include a category that he does not mention, this program would constitute a form of *impure procedural justice*. It is impure since it has to be supplemented in two respects: to include the positive desideratum of diversity and the need, in some cases, for an override mechanism.

To illustrate the argument, consider jury decision-making. One natural positive proposal would be to design juries to produce good decisions or, in a more sophisticated approach, juries that minimize an appropriately weighted sum of false convictions and false acquittals. If we limit ourselves to the verdict, there is, in standard cases, a fact of the matter that the jury is supposed to track. The goodness of the verdict consists in its truth. We can never tell,

[8] For criticisms of these analyses see Donohue and Wolfers (2005) and Ayres and Donohue (2003).

[9] Rawls (1971), § 14. He discusses perfect procedural justice (illustrated by the principle "I divide, you choose" for dividing a cake), imperfect procedural justice (illustrated by the criminal trial), and pure procedural justice (illustrated by his own theory of justice).

however, whether juries are better than judges, or juries with 12 members better than juries with 10 members, at tracking the truth or at minimizing the objective function. If we had a procedure for deciding the correctness of a verdict, we would obviously use that instead of going through the trial. [10] Also, what would justify the choice of the weights in the function to be minimized?[11] Finally, when we turn from verdicts to discretionary sentencing or awards, there is usually no fact of the matter for the jury to track.

Suppose, *per impossibile*, that all these problems could be overcome. We would still be confronted with insuperable causal problems in trying to decide which particular jury arrangement is likely to achieve the aims of truth-tracking and error-avoidance. On this point I refer to my earlier article. Briefly restated in the present context, the argument is that neither deductive models based on the assumption of rational jurors nor statistical models based on observation or experiments will tell us whether 12-member juries perform better than 10-member juries, whether juries deciding by unanimity perform better than juries deciding by simple or qualified majority, or whether judges or juries perform best in civil trials.

An alternative positive approach would be to try to select good jurors. Historically, jurors have been subject to many of the economic qualifications imposed on voters. As I argue more fully below in the context of voting, there is no reason to believe that income and property are useful proxies for the qualities we might wish decision-makers to have. Literacy requirements may be somewhat more justified, although jurors are typically exposed to oral arguments only. In American jury trials, the prosecutor and the lawyer for the defense routinely use peremptory and for-cause challenges to eliminate jurors who might be prejudiced for or against a defendant of a given race. The pathologies and absurdities of this procedure are well-known.[12]

The negative approach is more promising. The need to shield jurors from bribes and threats has been recognized for centuries. Whereas bribes always appeal to interest, threats can appeal either to the prudential (interest-based) fear of the jurors or to their visceral (emotion-based) fear. One might also, as is done in some jurisdictions, prevent the jurors from drinking alcohol during or between their deliberations. Photographs of victims of violence should be presented in black and white, not in color. In high-profile cases, one can use delays and changes of venue to reduce undue influence of the media on the jurors. To prevent jurors from acting on the prejudiced beliefs that "who steals an egg steals a cow" or "false in one, false in all", they should not be told about previous convictions of the

[10] Although DNA testing has made it possible to produce estimates of the frequency of wrongful convictions (Risinger 2007, Garrett 2011), nobody to my knowledge has tried to use this frequency as a dependent variable with alternative jury procedures (or jury-versus-judge) as independent variables. Nor do the data lend themselves to this kind of analysis.

[11] On the range of possible weights, see Vologh (1997).

[12] See Amar (1995) for a devastating indictment.

accused. The accused should appear in court in normal dress, not in prison attire. (It has also been argued, although implausibly, that minority defendants should be allowed to be physically absent from the courtroom throughout the trial to avoid triggering racial prejudice.[13]) To reduce cognitive bias, statistical information, for instance concerning DNA evidence in a murder trial, should be communicated in terms of natural frequencies rather than probabilities.[14] Once one has gone as far as possible in shielding jurors from influences that can distort their judgment, their decision has to be accepted, with the exception that judges should be allowed to override what they see as a wrongful conviction.

I shall now proceed as follows. In Section **II** I consider the internal logic of welfare-based and preference-based conceptions of the public interest. Section **III** considers what one might call the external adequacy of these conceptions. In Section **IV** I discuss the properties we would wish decision-makers to have, and whether one can design institutions to select individuals with those qualities. In Section **V** I turn to the details of the negative approach that emphasizes the screening of decision-makers from distorting influences. I consider some problems to which the proposal may give rise, and discuss how they might, at least in part, be attenuated.

II. Internal problems of aggregation

I shall consider collective decision-making procedures that are defined by the aggregation of subjective mental states: beliefs, preferences, or welfare levels.[15] Here I assume that these are the mental states of the decision-makers. (In the next Section I allow for a larger set of mental states.) The main argument is that in many situations the outcome of the aggregation is indeterminate. There is an obvious parallel, although not a very instructive one, between this indeterminacy and the indeterminacy of rational-choice theory that I discussed in my earlier article.

Some decisions aim at cognitive goodness only. Examples include criminal verdicts and many expert-based decisions. Other decisions aim at non-cognitive goodness only. Examples include debates over abortion, torture, amnesty for wrongdoers, and euthanasia. Although defenders of these practices often invoke consequentialist arguments that can be assessed by cognitive standards, their opponents typically refuse the relevance of factual or causal claims. Some opponents also refuse the appropriateness of procedures such as majority voting in settling these issues. Still other decisions aim at both cognitive and non-

[13] Note (2005)

[14] Hoffrage et al. (2000).

[15] I understand preferences in the wide sense of Hausman (2011).

cognitive goodness: choosing appropriate ends and identifying factual constraints as well as effective causal means. Examples include redistributive policies and measures for climate change abatement.

When all decision-makers *agree* on a cognitive or non-cognitive aspect of the decision, the collective conclusion follows immediately. From the internalist perspective that I adopt in this Section, the only important issues arise when there is disagreement. In jury decisions, where only cognitive goodness is relevant, it might seem obvious that majority voting is appropriate. Although most jury systems that do not require unanimity for verdict use qualified majorities, the 15-member Scottish juries decide by simple majority. The choice of the majority size might be guided by the desire to maximize cognitive goodness or by some non-cognitive end such as minimizing the cost of decision-making. This choice belongs, however, to the legislature, not to the jury itself. The task of the jury is to deliberate and then to aggregate the beliefs of its members by voting.

This procedure can, however, be indeterminate, a fact first recognized by the French mathematician Poisson in 1837:

> Two individuals, whom I shall call Pierre and Paul, are accused of theft; to the question whether Pierre is guilty, four jurors say *yes*, three others *yes*, and the five remaining *no*: the defendant is declared guilty by a majority of seven votes to five; to the question whether Paul is guilty, the first four jurors say *yes*, the three others who had said *yes* against Pierre say *no* against Paul, and the five remaining say *yes*: Pierre is therefore declared guilty by a majority of nine votes to three. Next one asks whether the theft has been committed by *several* individuals, which in case of an affirmative answer entails a more serious punishment. Following their previous votes, the first four jurors say *yes* and the remaining eight who had declared either Paul or Pierre to be innocent, say *no*. Hence even though there is no contradiction in the votes of the jurors, the decision of the jury is that both are guilty of theft and that the theft has not been committed by several individuals. [16]

The jury could, in other words, use either of two procedures to decide whether the theft was committed by several individuals. On the one hand it could vote directly on this issue and reach a negative answer. This is usually referred to as a *conclusion-based procedure*. On the other hand, it could first vote separately on the guilt of the two individuals and then, from the finding that they are both guilty, infer a positive answer to the issue of their joint guilt. This is usually referred to as a *premise-based procedure*. Neither procedure is unambiguously optimal on normative grounds. This paradox was rediscovered in 1921 by the Italian legal philosopher Vacca, and then re-rediscovered in 1986 by Lewis Kornhauser and Larry

[16] Poisson (1837), p. 21 n.

Sager[17] It is deeply unsettling. In many cases, there is no determinate answer to the question "What is the opinion of the group?" In a trial, a defendant might be guilty according to the views of the *jury* in a one-step aggregation but not according to the views of the *jurors* in a two-step aggregation.

The Poisson paradox can also arise in joint aggregation of beliefs and fundamental preferences. Table 1 illustrates in stylized form the debates over unicameralism versus bicameralism in the French Assemblée Constituante of 1789.[18] Broadly speaking, the assembly contained three roughly equal-sized groups. The reactionary right wanted to set the clock back to absolute monarchy, the moderate center wanted a constitutional monarchy with strong checks on parliament, and the left wanted a constitutional monarchy with weak checks on parliament. On the issue of bicameralism, the constellations were, highly simplified, the following:

	Fundamental preferences	Beliefs	Policy preserences
Reactionaries	Destabilize regime	Bicameralism is stabilizing	Unicameralism
Moderates	Stabilize regime	Bicameralism is stabilizing	Bicameralism
Radicals	Stabilize regime	Bicameralism is destabilizing	Unicameralism
Majority	Stabilize regime	Bicameralism	Unicameralism

Table 1

In this stylized rendering, a majority made up of reactionaries and moderates *believed* that bicameralism would stabilize the regime, while a majority made up of moderates and radicals *desired* to stabilize the regime. The reactionaries desired, of course, to destabilize it. The actual decision, to adopt *unicameralism*, was taken by voting over policy preferences. If instead the decision had been taken by first aggregating beliefs by (sincere) majority voting, next aggregating fundamental preferences by (sincere) majority voting, and finally taking the

[17] Vacca (1921); Kornhauser and Sager (1986).

[18] Egret (1950).

action which according to the aggregate belief would best realize the aggregate preference, *bicameralism* would have been the choice.

This is a hypothetical example. In reality, large assemblies always vote over policy options in a single step rather than using a two-step aggregation. The latter procedure would be paralyzing. Smaller groups may, though, face a real dilemma. When deciding on the rate of interest, members of the board of a central bank may have to resolve the normative issue of the trade-off between inflation and unemployment as well as the causal issues involved in a choice of an economic model. In general, there is no way of assessing whether the public interest is best served by each member voting twice (on the normative and the causal premises) or only once (on the interest rate).[19] Although the boards of central banks are not democratic institutions, one can certainly imagine similar paradoxes arising in small elected assemblies such as municipal councils or the governing bodies of academic institutions.

Prior to the 1986 re-rediscovery of Poisson's paradox, the Condorcet paradox offered the best-known example of procedural indeterminacy. Although the empirical frequency of cyclical majorities remains a contested issue, their possibility and occasional reality are uncontroversial. In a Poisson paradox, the group cannot state what it *believes*; in a Condorcet paradox it cannot state what it *wants*.

This seductive parallel, though, is misleading. In voting, the informational input to the aggregation process – the ordinal preferences of the voters – is, as it were, *doubly impoverished*. It neglects both the cardinal dimension of individual preferences and interpersonal differences. The second neglect is the decisive one. The Condorcet paradox would disappear if one could attach interpersonally comparable utilities to the options and choose the one that realizes the largest sum of individual utilities. In belief aggregation, the informational input is *impoverished in one respect only*, by neglecting the cardinal aspect of the beliefs (their strength). A juror has the choice only between "guilty" and "not guilty" (and, in Scotland, "not proven"). The Poisson paradox could still arise, however, if individual beliefs were to be expressed in the form of quantitative subjective probabilities.

Utilitarianism does not really, of course, offer a way out Condorcet's paradox. The conceptual and practical obstacles to interpersonal comparisons of utility are known to be formidable and perhaps insurmountable.[20] To be sure, in practice we sometimes do use this utilitarian approach in a rough and ready manner. On the reasonable assumptions that individuals are pretty much similar in their capacity for transforming money into subjective welfare and that the marginal utility of money is decreasing, we feel justified in transferring some income from the rich to the poor. Often, however, the utilitarian approach provides no guidance for action. If some members of the municipal council want to build a

[19] Claussen and Røisland (2010 a, b).

[20] Elster and Roemer, eds. (1991).

swimming pool, others a school library and still others wish to fund an orchestra, interpersonal comparisons of the utility the citizens would derive from the various options are likely to unfeasible.

Michel Balinski and Rida Laraki have recently offered a more promising way out of the paradox, by substituting the information-richer input of qualitative *judgments* for the mere expression of preferences. They argue, on the basis of extensive empirical investigations of grading in wine tasting, ice skating and other domains, that "inputs [to an aggregation mechanism] given in the six-word language of grades *Excellent*, *Very Good*, *Good*, *Acceptable*, *Poor*, and *To Reject* convey much more precise common meanings within a culture, than the input of rank-order, the name of one candidate, or the names of several candidates".[21] Crucially, politics also seems to have a common language. In a survey where 1752 persons were asked to grade twelve French politicians for the office of President of the Republic, "[m] ore than one of every three participants gave their highest grade to two or more candidates. Only half of the voters used the grade of *Excellent*. […] *This proves that voters do not have in mind rank-orderings of the candidates*".[22] Balinski and Laraki show that the choice of the candidate with the highest median grade is invulnerable to the Condorcet paradox and to strategic manipulation as well. If there there is no common language of qualitative grades, one can use the Borda count and choose the option with the highest median score rather than, as in the usual procedure, the highest mean score. Although vulnerable to some forms of manipulation, this procedure is not vulnerable to the Condorcet paradox.

If adopted, the Balinski-Laraki proposal would attenuate some of the indeterminacy of majority decision-making. It would not, however, offer a way out of the Poisson paradox. More important, for the reasons set out in the next Section, it would fail as a full-blown normative theory.

III. External problems of aggregation

In this Section I shall bracket the problems explained in the previous Section, by assuming away the paradoxes of collective belief formation and collective preference formation. Although I shall conduct the discussion in the framework of classical social-choice theory, the arguments also apply to the Balinski-Laraki proposal. My claim is that even an internally coherent procedure may be externally incomplete, in the sense I shall now explain.

I shall distinguish between the agents of a collective decision (the *deciders*) and the *targets* of the decision. The latter are the individual or individuals on whom the decision has *an intended or expected effect*. The *primary targets* are those whose behavior or situation the deciders

[21] Balinski and Laraki (2010), p. 389.

[22] *Ibid*., p. 13; italics in original.

intend to affect, whereas *collateral targets* are those whom the deciders expect to be affected by the decision. These targets often form a much smaller group than the set of individuals actually affected by the decision. In Boolean terms, the relation between deciders and targets can take five different forms: (i) the agents and the targets coincide; (ii) the deciders are a proper subgroup of the targets; (iii) the targets are a proper subgroup of the deciders; (iv) the deciders and the target overlap, but neither group is included in the other; and (v) the deciders and the targets are disjoint groups.

The framework of social-choice theory is limited to the deciders. Although these may include the welfare of the targets in their own preferences, nothing in the theory requires them to do so. Adults may legislate for children in the best interest of children, but also in the best interest of adults. An analogy with rational-choice theory may be instructive. That theory tells agents how best to realize their desires, *whatever these might be*. If Hitler had listened to such advice, he might have won the war. Instead, his emotions caused him to act irrationally: his hubris triggered by the victory against all odds over France led to his disastrous conduct of the war in the East, his anger at the British bombing of Hamburg caused him to direct air attacks against London instead of targeting airfields, his contempt for the Slavs made him blind to the benefits of cooperating with the Western Ukrainians, and his hatred of Jews made him divert scarce resources to the Holocaust. In all these cases, a less emotional response would have served his ends, *such as they were*, better. Similarly, the normative question asked by social choice theory is how to aggregate individual preferences, *whatever these might be*. If all deciders prefer exterminating all Jews to eliminating some Jews or none, the Holocaust will be the socially preferred outcome. Like rational-choice theory, social-choice theory and its cognates are in an important sense normatively neutral. Although a good collective decision ought to promote the welfare of the targets, social-choice theory does not have the conceptual wherewithal to achieve this end.

One might try address this problem in various ways, either by (1) "input filters" or (2) by "output filters".[23] Input filters could take the form of (1a) "laundering preferences"[24] or of (1b) selecting good decision-makers. Although (1a) bears some resemblance to the Benthamite proposal that I discuss in Section **V**, the proposal is too devoid of institutional detail to be helpful. I consider (1b) in the next Section. Output filters might take the form of constraining the social-choice framework by individual *rights*. Although any version of this proposal would certainly block genocidal decisions, different versions will have widely different implications and depend heavily on often-arbitrary legal interpretations. As Montaigne observed, "[j]ust as no event and no form completely resembles another, neither does any completely differ. [...] All things are connected by some similarity;

[23] Goodin (1986).

[24] *Ibid.*

yet every example limps and any correspondence which we draw from experience is always feeble and imperfect; we can nevertheless find some corner or other by which to link our comparisons. And that is how laws serve us: they can be adapted to each one of our concerns by means of some twisted, forced or oblique interpretation."[25] For this reason, I share Ely's skepticism towards substantive judicial review. In the Conclusion I espouse his theory of procedural judicial review, in an expanded version that addresses Benthamite concerns.

Although important, these issues are secondary compared to a more fundamental problem. Even if we suppose that deciders are concerned only with promoting the welfare of the targets, that *objective function is not well-defined* if we include, as we should, future generations among the targets. In addition to the thorny and seemingly irresolvable normative problem of how to discount the welfare of future generations, there is basic uncertainty about factual as well as causal matters. Concerning the facts, we cannot know or estimate how many individuals there will be in the distant future. Concerning causality, the debate over climate change suggests that the parameter linking increases in CO_2 concentration in the atmosphere to temperature change is subject to deep structural uncertainty.[26] It is arguable, and in my view indisputable, that these issues dwarf those that arise among the individuals who are currently living. While we may get some second-decimal matters right, we shall remain ignorant about the first decimal.

IV. Choosing good decision-makers and good decision-making bodies

I have argued that the idea of good collective decisions is, in general, indeterminate. A natural response might be to lower one's sight and trust in the selection of good decision-makers or the creation of good decision-making bodies. "Anyone who could discover the means by which men could be justly judged and reasonably chosen would, at a stroke, establish a perfect form of commonwealth."[27] I shall argue that, with one exception, even this idea is too ambitious.

To my knowledge, Bentham provided the best analysis of the qualities or "aptitudes", as he called them, that we would want good decision-makers to possess. Although he was mainly concerned with *removing obstacles* to the deployment of these aptitudes, I shall here consider – and reject - the idea of *favoring* them by institutional design.

It is a commonplace to wish for decision-makers to possess both virtue and ability or, in Bentham's language, both moral and intellectual aptitude. To this traditional idea he added two new ones. First, he wanted decision-makers to possess *active* aptitude - dedication,

[25] Montaigne (1991), p. 1213.

[26] Weitzman (2009).

[27] Montaigne (1991), p. 1057. He believes, of course, that it can't be done.

application, energy, exertion, or industry. He may have read and applied to legislators Gibbon's observation that "the time of a prince is the property of his people".[28] Second, he emphasized the (multiplicative) *interaction* among the three kinds of aptitude, arguing for instance that Napoleon's combination of low moral aptitude and high intellectual and active aptitude rendered him especially dangerous. Conversely, he would have subscribed to the common characterization of tsarism or of the Habsburg monarchy as "tyranny tempered by incompetence". Although I cannot cite instances of despotism tempered by laziness, Louis XVI offered, in Bentham's view, an example of (relative) moral and intellectual aptitude rendered ineffective by laziness.[29] Gibbon cites the poet Claudian (4th century A. D.) as comparing "in a lively epigram, the opposite characters of two præfects of Italy; he contrasts the innocent repose of a philosopher, who sometimes resigned the hours of business to slumber, perhaps to study, with the interested diligence of a rapacious minister, indefatigable in the pursuit of unjust, or sacrilegious gain. 'How happy […] might it be for the people of Italy if Mallius could be constantly awake, and if Hadrian would always sleep!' "[30]

To my knowledge, there have been no attempts to select decision-makers for their active aptitude. The practice of choosing jurors from the voter lists has been defended as a selection mechanism, on the grounds that it "automatically eliminates those individuals not interested enough in their government to vote",[31] but was hardly established for that purpose. By contrast, over the centuries there have been many attempts to choose voters, deputies and jurors for their moral or intellectual aptitude. Often, these have been based on implicit or explicit assumptions that income or wealth were good proxies for these aptitudes. The acquisition of wealth was supposed to signal general ability. The possession of landed property was supposed to create a virtuous attachment to the long-term interests of the nation. These assumptions have been so thoroughly discredited that there is little point in discussing them. Let me note, however, that even if one may to some extent select for intellectual aptitude by requiring literacy tests, there is no reason whatsoever to believe in a positive correlation between intellectual and moral aptitude. A negative correlation is at least as plausible. Distrust of elites has often been rooted in the perception that "[great] abilities have generally […] been employed to mislead the honest, unwary multitude".[32]

[28] Gibbon (1995), vol. 2, p. 52, commenting on the fact that "the virtuous mind of Theodosius was often relaxed by indolence".

[29] Bentham (1989), pp. 103, 139, 178.

[30] Gibbon (1995), vol. 2, p. 162.

[31] Judge Irving Kaufman testifying in the Senate, cited in Van Dyke (1977), p. 90. For a somewhat different argument along similar lines, see Amar (1995), p. 1179-80.

[32] *Cato's Letters*, cited in Storing (1981), vol. 4, p. 244.

More sophisticated proposals rely on cleverly designed electoral systems. It has notably been argued that the centrifugal or extremist mechanisms in divided societies can be overcome by centripetal mechanisms designed to produce moderate candidates.[33] The "alternative vote" system has been advocated on these grounds.[34] Another proposal with roots in classical antiquity is that of "cross-voting". Members of group A may delegate to another group B the power or some of the power to choose the representatives of group A. This system will allegedly produce candidates for group A that can appeal to a broad electorate. The process can also involve more than two groups. There are many variations on this general theme. The main variation is quantitative, in the following sense: in the election of representatives for group A, the proportion of voters who belong to group A may vary from 0 (Brittany 1576, 1614) through 1/6 (Maryland 1789), 1/3 (Georgia 1789, France 1484, 1789) to some number between ½ and 1 (proposed scheme for Belgium).[35]

Broadly speaking, these schemes have proved quite fragile, in part because, they can set up incentives for undesirable as well as for desirable behavior, and because they rely on an assumption of rational voters that is not always be satisfied. Voters may game the system, if members of group B vote for incompetent candidates for group A rather than for moderate candidates. In the related system of cross-over voting in presidential primaries, Democrats may vote for the Republican candidate who is least likely to be elected rather than for the candidate who, if elected, would be closest to Democratic values. This is a pervasive problem in institutional design: "It may be impossible to anticipate the many ways in which a particular incentive scheme may be gamed […]. There are a virtually infinite number of dimensions along which strategic […] personnel can act to game the current system".[36]

Voters may also, irrationally, attach greater importance to what they lose than to what they gain. In cross-voting, a group gives up the exclusive right to select its own representative in exchange for a right to contribute to the choice of representatives for other groups. What is lost in this exchange is the option of choosing a representative who would promote its interest vigorously and aggressively. What is gained is the elimination of candidates who would oppose that interest with equal vigor and aggression. Even if an observer judges that the gains outweigh the losses, the theory of loss aversion predicts that actors will judge differently.

Yet even if attempts to select individual decision-makers for desirable aptitudes are likely to fail, could one design decision-making *bodies* to improve the quality of their

---

[33] Reilly (2011).

[34] The sharply opposed views on the efficacy of this system in Fraenkel and Grofman (2006) and in Horowitz (2007) suggest that the problem of indeterminacy arises here as well.

[35] I refer to Elster (2013) for details.

[36] Jakob and Levitt (2003), pp. 843, 871.

decisions? The two main instruments that have been proposed are *size* and *diversity*. Formal theorems show that under certain conditions, larger bodies outperform smaller ones, and that diverse bodies outperform bodies that have higher individual expertise but less diversity.

Condorcet's Jury Theorem purports to show the cognitive benefits of size. It is an indisputable mathematical fact that if all deciders have a greater than 50 % chance of being right, form their opinions independently of each other, and vote sincerely on the basis of these opinions, the efficacy of majority voting in tracking the truth converges to 100 % as the number of voters grows indefinitely large. Numerous variations on this theorem have also been proved. Their practical relevance is highly disputable, however. We cannot determine, in practice or even in theory, whether the conditions for the theorem to hold - individual competence, independence, and sincerity – obtain. This is not simply an empirical difficulty, but a logical one. Although the competence and the independence conditions may be simultaneously true, they cannot simultaneously be *shown* to be satisfied.[37] In addition, the independence condition in itself seems questionable. While it prevents bad interactions effects that arise through conformism and social pressure, it also excludes good interaction effects arising from information-pooling.

Bentham pointed to another weakness in the theorem: "[Claim:] *With the number of members increases the chance of wisdom*. So many members, so many sources of light. Answer : the reduction which that same cause operates in the strength of the motive to bring out this light […] offsets this advantage."[38] The tendency towards informational free riding increases with the size of the body.[39] In Bentham's language, active aptitude decreases with size. Aanund Hylland has shown (personal communication) that if p_N is the probability that each of N deputies will "get it right", the probability of a majority vote "getting it right" converges to 100 % only if (p_N -1/2) goes to 0 more slowly than the square root of N goes to infinity.[40] If that is not the case, the dilution of active aptitude may offset the increase in "light".[41] Although Condorcet also recognized that individual competence would decrease in large groups, his argument was based on the need to select less "naturally competent" decision-makers (scraping the bottom of the barrel). By contrast, Bentham's argument relies on an endogenous rather than an exogenous source of diminishing competence in ever-larger groups.

---

[37] Dietrich (2008).

[38] Bentham (2002), p. 122.

[39] Karotkin and Paroush (2003), Mukhopadhaya (2003). Grossman and Stiglitz (1980) show that *markets*, as well as *assemblies*, can induce suboptimal investment in information.

[40] The same result obtains if the competence of the deputies diminishes for exogenous rather than endogenous reasons, that is, if an expanding group has to enroll less competent members (Berend and Paroush 1998).

[41] Grofman and Feld (1989), p. 1338 also noted this weakness in Condorcet's argument.

One might look, therefore, for the optimal size of a decision-making body, at which the costs and benefits from adding an extra member are equalized. This idea is chimerical, however. It presupposes the kind of science-fiction modeling that was the target of my earlier article.

The argument from (cognitive) diversity goes back to Aristotle: "[…] the many, who are not as individuals excellent men, nevertheless can, when they have come together, be better than the few best people, not individually but collectively, just as feasts to which many contribute are better than feasts provided at one person's expense" (*Politics* 1281a). Intuitively, the idea makes good sense. In an assembly, diversity of membership may help it both to identify problems (where the shoe pinches) and to propose solutions (how to make better shoes). The experiences from twelve very different state constitutions were invaluable in helping the framers in Philadelphia to discard unworkable arrangements and adopt feasible ones.[42] When Louis XVI imposed a double representation for the third estate at the Estates-General in 1789, it was to represent their information rather than their interests. The Estates should "gather together all knowledge that might be useful for the good of the State, and one cannot contest that this diversity of knowledge belongs above all to the third estate, since there are numerous public matters that only this order is informed about".[43]

In a jury, diversity of experience among jurors may help them catch inconsistencies in testimonies and in narratives proposed by the prosecution or by the defense. In addition to the poignant episodes in *Twelve Angry Men*, an example from a real jury trial illustrates the point. In a civil case " a woman claimed that as a result of an automobile accident she continued to have severe back pain. During discussions, [a female juror] observed that she was wearing high heels, and that when she stepped off the raised witness stand after her testimony, she didn't even wince".[44] A man might not have noticed this discrepancy. In Norway, juries are required by law to have roughly 50% female jurors. It is at least arguable that this feature of juries makes them take rape cases more seriously than all-male juries would do.

Scott Page has recently formalized the conditions under which "diversity trumps ability".[45] As is also the case for Condorcet's Jury Theorem, it seems hard and perhaps

[42] See Adams (2001), Ch. XIV for a survey of the impact of the state constitutions on the Federal constitution. Bryce (1995), p. 31 asserts that "It has been truly said that nearly every provision of the federal Constitution that has worked well is one borrowed from or suggested by some state constitution; nearly every provision that has worked badly is one which the Convention, for want of a precedent, was obliged to devise for itself". By contrast, the assembly's capacity for detecting shoe-pinching was undermined by the massive underrepresentation of the backcountry at the Federal Convention (Elster 2012).

[43] AP 1, p. 492. By contrast, the assembly's capacity for improving shoe-making was undermined by the lack of political experience of the deputies.

[44] Vidmar and Hans (2007), p. 272.

[45] Page (2007).

impossible to verify empirically when these conditions are satisfied. The stylized examples Page offers are compelling. It is quite possible that actual cases in which diversity proves beneficial follow the same logic – but we don't know. There is a simpler argument for the value of diversity, however. Since, as I have argued, we cannot reliably choose democratic decision-makers for their ability, the question of the diversity-ability trade-off does not arise. Since we *can* choose decision-makers for their diversity – selecting jurors by lottery or deputies by proportional voting – we should do so.

V. Towards a more modest and more robust political theory

The Benthamite program I sketch in this Section has three interlocking parts and two supplements. First, institutions should be designed to protect the *active, intellectual and moral aptitude* of the decision-makers, in the sense of removing or minimizing negative influences on these aptitudes. The main focus will be on moral and active aptitude. Second, these influences can be spelled out as *self-interest, passion, prejudice, and bias*. Third, the most important means for neutralizing them are *secrecy, publicity, and ignorance*. In some cases, *rotation* of decision-makers to prevent capture by interest groups can also be a valuable tool. In a compelling metaphor, John Lilburne told the private soldiers in the New Model Army, "Suffer not one sort of men too long to remaine adjutators [delegates], least they be corrupted by bribes of offices, or places of preferment, for *standing water though never so pure at first, in time putrifies*".[46] To this negative program we should add the positive value of *diversity*. Juries and assemblies should also be subject to narrowly defined *overriding mechanisms*.

I shall consider this program with respect to three institutions: juries, assemblies, and elections.

Jury trials are public, to prevent the operation of what Bentham called "sinister interest", that is, political justice. Jury deliberations, by contrast, are secret in the sense that non-jurors have no access to them during the trial. In some countries, jurors are also forbidden to report from their discussions after the trial. These bans create a serious obstacle to understanding how juries work, and perhaps also to jury reform. With a few minor exceptions, jury votes are also secret in the same sense. The rationale for secrecy is the need to protect jurors from threats and to make it impossible to bribe them, by preventing them from making credible promises to vote one way or another. Jury votes may also be secret in a different sense, if the jury uses secret ballot. In that case, the jurors themselves have no access to how other jurors voted. The rationale for this "ultra-secret" voting might be the avoidance of conformism or group pressure on holdouts. At the same time, however, this practice would

[46] Woolrych (1986), p. 192; my italics.

allow jurors to vote according to idiosyncratic prejudices, without having to argue for their vote.

This dilemma can be solved by adopting a remarkable proposal that Bentham made for voting in assemblies and that also applies to juries. He argued that voting should be secret ex ante, but public ex post. At the time of voting, no one would know how others are voting; after the vote, everybody will learn how others have voted. While the secrecy will reduce conformism and eliminate group pressure, the publicity will induce responsibility by the fact that jurors will know that a frivolous vote will trigger blaming and shaming by their peers. In my view, this procedure should be adopted in all bodies that decide by voting preceded by deliberation. Although one might object that seriatim voting in public can provide useful information for those who vote late in the sequence, that information can equally well be provided in the deliberation phase.

In some circumstances, legislators or judges may wish jurors to be ignorant of certain facts about the law or about the defendant so that they will decide "behind a veil of ignorance", to use a phrase coined for a different context. The rationale may either be to induce other agents to behave in socially desirable ways, even at the expense of a good decision in the given case, or to increase the likelihood of a good decision in the given case. Concerning the first reason, the best-known example is perhaps the ban on the fruit of the poisoned tree. In a given case, illegally obtained evidence might provide useful information to the jury, but over time banning such evidence will dissuade law enforcement agents from seeking to obtain it. An illustration of the second reason is provided by American criminal law, where "the criminal record of the defendant who does not testify generally cannot be made available to juries as evidence of his or her propensity to commit such acts. While a defendant with such a record may be more likely to have committed the offense currently charged, the probative value of this information is seen as outweighed by its potential prejudicial effect".[47] Other biases, such as the hindsight bias in tort cases, are probably ineradicable.

In contemporary democracies, parliamentary debates and votes are public. The principal (the electorate) has a right to know what their agents (the deputies) say and do. The publicity may either take the form of printed parliamentary records or of access of the public to the gallery. As Bentham argued, the latter is essential for the credibility of the former. As in the case of the jury, we may distinguish between revelation of votes to other deputies and revelation to the public at large. Bentham argued that the worst system would be one that provided the first but not the latter form of publicity. In this regime, deputies would be held accountable to each other and be able to make credible logrolling promises, but not be accountable to the public for the deals they make. The Federal Convention illustrates this case. It is not clear, however, that the system was as bad as Bentham argued. Later Madison said,

[47] Diamond, Casper and Ostergren (1989), p. 249.

in conversation, that "had the members committed themselves publicly at first, they would have afterwards supposed consistency required them to maintain their ground, whereas by secret discussion no man felt himself obliged to retain his opinions any longer than he was satisfied of their propriety and truth, and was open to the force of argument".[48]

The Federal Convention illustrates a general problem to which I shall return: a device that shields the decision-makers from one distorting factor may leave them vulnerable to another. In an assembly that debates and votes before an audience, deputies may be deterred from acting according to naked self-interest, e.g. voting in favor of large salary increases to themselves, but they may also be subject to vanity, glory-seeking and fear. In ordinary parliamentary assemblies, the need to counteract self-dealing is probably more important than the need to keep the proceedings in a dispassionate register. In constituent assemblies, where the interests of the delegates and the groups they represent tend to have less purchase on the issues (the Federal Convention being an obvious exception), debates behind closed doors may be the more desirable mode. Also, when the constituent assembly is a one-off event (the Federal Convention being an obvious example), the rationale for publicity that is linked to the possibility of not reelecting unfaithful agents does not exist.

Voting in assemblies can take several forms. In one extreme case, represented by France between 1798 and 1843 and the Italian parliament until 1988, votes were secret to other deputies and *a fortiori* to the public at large. The effect of – and probably the rationale for – this practice has been to shield deputies from party discipline and from pressure by the executive. A probably-unintended consequence of secret voting has been to prevent logrolling. From a normative perspective, this effect may nor may not be desirable.[49] Even assuming it is desirable, in modern democracies it is almost certainly swamped by the negative effects stemming from lack of accountability.

At the opposite extreme, votes are cast in public in full view of all deputies and of the nation. Earlier, I have mentioned two intermediate or hybrid schemes, one defined by the nature of the public and the other by the timing of the publication. Votes might be known to other deputies, but not to the public at large. They might be secret ex ante but public ex post. In ordinary parliamentary assemblies, the first hybrid is clearly undesirable, except when secret sessions are required by a military emergency. The second seems desirable under all circumstances, even in secret sessions.

As one indication of the opprobrium attached to self-serving assembly behavior, one may cite the fact that when the members of the French *Constituante* enacted emoluments for

[48] Farrand (1966), vol.III, p.479. The debates were in fact leaked, but only to a French diplomat (*ibid*., p. 61-63).

[49] "Today, no consensus exists in the normative public choice literature as to whether logrolling is on net welfare enhancing or welfare reducing" (Stratmann 1997, p.322); see also Mueller (2003), pp. 104-12. Riker and Brahms (1973) claim that in the long run *everybody* loses by vote trading.

themselves on September 1 1789, they were so afraid of being seen as self-interested that they did not insert the decision in the published record.[50] In the United States, there was such a high "degree of citizen indignation when legislators voted themselves a pay increase in 1816 that almost two thirds of them failed to return to Capital Hill after the next election, even though they had hastily repealed the compensation law in the meantime".[51] *Ignorance* can achieve the same aim as publicity, if politicians are required to put their assets in a blind trust. In the United States, there are no strict requirements to this effect, but some members of Congress choose to adopt this practice. The Congressional Blind Trust Act introduced in 2011 would, if adopted, make it mandatory.

Publicity and ignorance are tools to promote *moral aptitude*, by making it costly, risky or impossible for politicians to act in their own interest. For Bentham, morality equaled the impotence to do harm – a far cry from the idea of republican virtue. He claimed that in "the United States, Bonaparte might have been a Washington: in France, Washington might have been no more than a Bonaparte".[52] He was also concerned with the *active aptitude* of deputies. He discussed the problem of abstentionism in parliament and proposed to require "of each member a deposit, at the commencement of each quarter, of a certain sum for each day of sitting in the quarter; this deposit to be returned to him at the end of the term, deduction being made of the amount deposited for each day he was absent".[53] Since wealthy deputies would not be affected by this scheme, he proposed to supplement it by coercive measures: "one day of arrest for each contravention"[54]. Finally, he argued for a register of non-attendance, to be published at the end of each session.[55] As he observed with characteristic astuteness, these measures should be mechanical and automatic. The English practice of requiring in each case a vote of the House to punish an absentee member is unlikely to be efficient, "when all the judges are interested in the contravention of the laws".[56]

There are several other measures one can take to promote the active aptitude of deputies. In France, one could abolish the "cumul des mandats". There are far more politicians holding more than one elected office in France than in any other European country. In the European parliament, French members are much more likely than others to hold a

[50] Pierre (1893), p. 1154.

[51] Young (1986), p. 59.

[52] Bentham (1989), p. 213.

[53] Bentham (1999), p. 58.

[54] *Ibid.*, p. 59.

[55] *Ibid.*, p. 60.

[56] *Ibid.*, p. 61.

concurrent national office.[57] Although there is little doubt that the regime generates negative externalities, it is unlikely to be abolished anytime soon given that members of parliament have an interest in retaining it. The same comments apply to the common, although unconstitutional practice of voting by proxy in the French National Assembly. In the United States, members of Congress spend an inordinate time raising funds for reelection, at the expense of what they are elected to do. To counter this tendency, one might either impose a ceiling on the aggregate campaign contributions any deputy might receive (or have campaigns fully funded by the state). A ceiling on *individual* contributions does not solve this problem, since this cap would only induce officeholders to spend more time seeking out contributors. The individual ceiling might nevertheless be useful to promote *moral* aptitude, by making deputies less beholden to important donors.

As noted earlier, active aptitude can also be undermined by informational free riding. An assembly can overcome this collective action problem by setting up a separate informational-gathering structure staffed by officials who are paid to determine the facts. An example is provided by the Unfunded Mandates Reform Act of 1995, which obligates Congress to determine the obligations that a federal law would impose on states, municipalities and tribes, and indicate which of them would not be funded by the federal government.[58] The decision by the Belgian parliament in 1875 that no proposition of law could be signed by *more* than six members[59] can be interpreted in the same spirit.

A minor, amusing and instructive example of institutional – in fact physical – promotion of active aptitude is provided by the nineteenth-century claim that "the desk on the floor [in the American Congress] encouraged members to sit there attending to personal business instead of listening: the House, it was said, ought just to have benches, like the House of Commons".[60] From classical Antiquity, one can cite Plutarch's Comparison of Solon and Publicola, where he asserts that Publicola's reason for appointing "quaestors over the public moneys […] was that the consul, if a worthy officer, might not be without leisure for his more important duties and, if unworthy, might not have greater opportunities for injustice by having both the administration and the treasury in his hands" – thus ensuring active aptitude in the best-case scenario and moral aptitude in the worst case.

The constructive vote of no confidence adopted in the German constitution of 1949 and later imitated in many countries can also be seen in this perspective. Although the mechanism owed its origin to the experience of the coalitions of extremes that brought down

[57] For data, see the special issue on « cumul des mandats » of *French Politics* 2007 (No. 3), as well as the very full analysis in Bach (2009).

[58] Vermeule (2007), p. 228-31.

[59] Pierre (1893), p. 724. The four other European countries he cites required a *minimum* number of signatories.

[60] Miller (1996), p. 46; see also Goodsell (1988).

governments of the Weimar Republic, it would have been a stabilizing factor in fragmented multi-party regimes such as the Second Polish Republic or the Third and Fourth French Republics. When deputies have to *compare* candidates for the office of prime minister rather than simply assess the incumbent, they will be forced to look more carefully. "To throw off the burden of a present evil is no cure unless the general condition is improved".[61]

Bentham affirmed that "[o]n the part of the people appropriate aptitude in the shape of moral aptitude is at all times at a maximum".[62] His argument was that "nothing more is required than the placing of [the right to vote] in the hands of individuals disposed each of them to take that course which in his judgment is most conducive to his own individual interest: so disposed, he will be disposed to take that course which is most conducive to the universal interest, for the universal interest is nothing else but the aggregate of all individual interests".[63] This utilitarian argument must be rejected. It is true, however, as Tocqueville says, that the majority of citizens "may be mistaken but cannot be in conflict with themselves"[64] and, as James Harrington says, that the people would not "cast themselves into the sea" as a mad prince might do.[65] A tacit presupposition of these statements is, of course, an effective secret ballot to prevent candidates from influencing the vote by bribes and threats. It is also important, although more difficult, to prevent the use of bribes and threats to induce voting abstentions.

To enhance both the moral and the intellectual aptitude of the voters, they should also be shielded from passion and bias. Many countries and several American states ban the sale of liquor on election day, partly no doubt because of the historical practice of candidates plying voters with liquor[66] and partly because voters are believed to be more open to reason when not "under the influence". The common practice of requiring televised debates and the publication of opinion polls to cease some time before election day may be motivated in part by the cooling-down effect of delay and in part by the desire to minimize bandwagon effects and the recency bias.

Along more utopian lines one might require that debates be broadcast on radio rather than televised, to shield the voters from potentially distorting visual impressions. The dialogue *Hermotimus* by Lucian of Samosata (2nd century AD) provides a precedent: "Reason insists that […] it must further be allowed ample time; he will collect the rival candidates together, and make his choice with long, lingering, repeated deliberation; he will give no

61 Montaigne (1991), p. 1085.

62 I*bid*., p. 143.

63 *Ibid*., p. 133; see also Bentham (1990), p. 56 along similar lines.

64 Tocqueville (2004), p. 265.

65 *Ibid*., p. 429.

66 Pierre (1893) p.266-68.

heed to the candidate's age, appearance, or repute for wisdom, but perform his functions like the [Athenian court of the] Areopagites, who judge in the darkness of night, *so that they must regard not the pleaders*, but the pleadings" (my italics). On the other hand, this advantage might be offset or more than offset by the fact that fewer people would tune in to the radio than would turn on their TV.

Bentham also asserted that "[in] the case of the people in their quality of Electors, no demand for active aptitude has place"[67]. On this point I believe he was wrong. Voting induces a collective action problem that, if not addressed, could lead to unacceptably low voter turnout. Relevant measures to increase voter participation include small electoral districts, compulsory voting, and publication of the names of abstainers on the Internet. (Note, however, that although small electoral districts would reduce free riding among voters, they would also lead to large assemblies that would increase free riding among deputies.) One could hold elections on market day or after church services, to enlist the vote of those who wouldn't otherwise bother to go to the polling station. More radically, one could reduce the number of voters through an "enfranchisement lottery", which would both enhance the motivation of voters to inform themselves and make it possible to provide them with information at low cost.[68]. Although these measures cannot substitute for civic spirit among the citizens, they can amplify its effects.

The Benthamite program needs to be supplemented on two counts. I shall not repeat what I said about diversity, but focus on the need for institutional overrides of decisions by jurors and assemblies.

In the chapter on "Cross-overs" in their classic study *The American Jury*, Harry Kalven and Hans Zeisel argue that in the United States there is (and should be) an asymmetry in the role of judges in acquittals and convictions by juries. The cross-over cases are those in which the judges presiding over the cases said that they would have been more lenient than the jury was. One case concerned "an auto accident where defendant's auto knocks down a female pedestrian. The jury finds him guilty of reckless driving, the judge only of the lesser crime of leaving the scene. […] The unusual feature of this case is that the jury is so incensed at the defendant that it is not satisfied to find him guilty of the very crime that angers them but most go on to find him guilty of reckless driving".[69] In such cases, the judge has the power to set aside the conviction or to impose the minimum penalty. There is a sharp asymmetry between acquittal and conviction: "As a matter of law [the judge] has no power over jury verdicts which, in his view, are too favorable to the defendant. In the cross-over case, however, the judge may have legal power to intervene […] In the end *the institutional arrangement*

[67] Bentham (1989), p. 142 n.

[68] López-Guerra (2011).

[69] Kalven and Zeisel (1966), pp. 298, 299.

is impressive. It gives the jury autonomy to do equity on behalf of the criminal defendant. Where the jury's freedom leads to 'illegally' harsh results, the judge is at hand, ready to erase them".[70] Kalven and Zeisel report that in 56% of the 103 cross-over cases in their sample, the judge let stand a verdict he would not have reached himself. In trials for a serious crime, the percentage falls to 39%[71] The override mechanism, although not perfect, had a considerable effect.

In Scotland, where juries can convict by a bare majority of eight to seven, the right to appeal against convictions was introduced in 1926 and further streamlined in 1980.[72] In Belgium, judges can overturn a conviction if the jury was divided 7 to 5. In Norway, judges can overturn jury acquittals as well as convictions, but the requirements for overrides are more stringent for acquittals. There have notably been several overrides of acquittals for rape. The diversity requirement that half the members of the jury be women is likely to lead to fewer acquittals in rape cases in any case.

Bentham argued that parliament should be omnipotent. He discussed and rejected indirect elections, bicameralism, supermajorities, entrenchment, judicial review, and prorogation as remedies for inaptitude. On this point he was manifestly misguided. If Parliament was not subject to constitutional constraint, it could postpone elections at will. As Tom Paine wrote, "Were a Bill to be brought into any of the American legislatures similar to that which was passed into an act by the English parliament, at the commencement of George the First, to extend the duration of the assemblies to a longer period than they now sit, the check is in the constitution, which in effect says, Thus far shalt thou go and no further."[73] Publicity cannot prevent self-dealing by the legislators if the voters are prevented from punishing them.

More generally, constitutional constraints are needed to *prevent those in power from using their power to stay in power*. The independence of state-owned media, of the central bank or of the National Bureau of Statistics is another implication of this idea. If measures constraining parliament are to have any effect, however, the constitution has to be relatively rigid, that is, difficult to amend; otherwise a simple majority in parliament will simply do in two steps what it is forbidden to do in one. In addition to this a democracy-based argument for rigid constitutions, one can cite an efficiency-based one: "If the vote of a simple majority could change the basic form of the government or expropriate the wealth of a minority, enormous resources might be devoted to seeking and resisting such legislation. In a sense, a supermajoritarian constitutional provision confines legislative discretion to matters that do not

[70] *Ibid.*, pp. 411, 413; my italics.

[71] *Ibid.*, pp. 412, 413.

[72] Duff (2000), pp. 277-79.

[73] Paine (1791), II. 4.

matter all that much; the stakes are not large enough to evoke a disproportionate expenditure of resources on redistributing wealth or utility". [74]

We must ask, however, how a written constitution can have causal efficacy. In France between 1789 and 1958 (or 1974), it was essentially an ineffective parchment barrier that had at most moral authority. In a speech justifying what is arguably the first written constitution in the modern sense of the term, the 1653 *Instrument of Government*, Cromwell argued that a single person, such as the Lord Protector (himself), was needed. "Of what assurance is a Law to prevent so great an evil, if it lie in the same Legislature to unlaw it again? Is such a law like to be lasting? It will be a rope of sand; will give it no security; for the same men may unbuild what they have built".[75]

Cromwell's solution to this classical problem was to enforce the constitution, and limit legislative omnipotence, by an executive veto. In the *Assemblée Constituante* of 1789, Mounier made a similar proposal.[76] Sieyes proposed to give *both* the legislature and the executive the right to call for an assembly to revise the constitution if they believed the other party was encroaching on its constitutional rights.[77] In 1795, he proposed "a *jurie constitutionnaire* that would ensure that the constitution was obeyed by annulling acts of the legislature and the executive that were contrary to it".[78] In the 19th century, the French and Swedish presidents of the national assemblies arrogated to themselves the right to declare a proposed bill unconstitutional.[79] Contemporary democracies rely on judicial review, a mechanism already implicit in *Federalist* No. 78.

John Hart Ely is the foremost exponent of the procedural theory of judicial review. According to him, the sole role of the Supreme Court is to ensure fair representation in Congress. Once that is done, no substantive override is justified. I believe this conception is too narrow. Because of its focus on elections, his theory does not exhaust the set of procedural remedies one might want to apply before letting the chips fall where they may. Nothing in Ely's argument would exclude, for instance, the *cumul des mandats* or require that deputies put their assets in a blind trust. A Benthamite extension of Ely's argument dictates that measures to promote the active, moral and intellectual aptitude of deputies should also be put into the constitution and be enforced by judicial review.

---

[74] Posner (1988), p. 9.

[75] Cromwell (1845), p. 120-21; my italics.

[76] AP 8, p. 585-86.

[77] AP 9, p. 219.

[78] Bell (1994), p. 21. See Troper (2006), Ch. VIII for a full discussion.

[79] Pierre (1893), p. 65. He also reports that in Sweden, the constitutional committee in parliament decided when a doubt arises.

Overrides of assemblies by courts and of juries by judges are widely practiced. It is harder, perhaps impossible, to find examples of institutional mechanisms for overriding the outcome of *elections* in cases where the impact of distorting factors has been reduced as much as possible. Overrides happen, to be sure. In Algeria, the Islamic Salvation Front won 188 out of 232 seats in the first electoral round in 1991, and was confidently expected to win sufficiently many of the remaining 198 seats to form the 75% majority in parliament required to change the constitution and create an Islamic state. The government cancelled the second round, "saving democracy" by violating it.[80] Since the ISF owed its victory in part to massive electoral fraud, the assumption that distorting factors had been minimized does not hold. Moreover, the override was purely ad-hoc, not based on an institutional mechanism. Whether it is because of limitations on my imagination or because of the nature of the issue, I do not believe that one could find an ex post control mechanism on elections for which all appropriate ex ante precautions are in place.

I conclude by considering two possible – and partly valid - objections to the line of argument I have been pursuing in this Section. First, does not the Benthamite approach also lend itself to the charge of indeterminacy that I have leveled against the traditional theories? It might seem as if I have assumed, naively, that all good things go together. This assumption is obviously vulnerable to the objection that one cannot expect that several functions will be minimized or maximized by the same value of the instrumental variable. I have observed, in fact, that a measure to reduce the impact of self-interest on decision-makers can enhance the impact of passion, and that measures to promote the active aptitude of voters may reduce that of deputies. Whereas the selection of jurors who are ignorant about a high-profile case they are to decide will eliminate bias caused by pre-trial publicity, jurors who do not read newspapers or watch TV are likely to be lacking in intellectual or active aptitude. There may also be a tradeoff between moral and active aptitude: "Delay replaces self-interested motivation with impartial reason, but the latter motivation is frequently too feeble to produce action".[81] Whereas rotation of officials and term limits of deputies may reduce the risk of capture by interest groups, these measures will have a negative impact on their knowledge of the issues. To these five examples one could add many others.

One response to this predicament would be to look for an optimal trade-off, such as, for instance, the size of electoral districts that maximizes some appropriate function of the active aptitude of voters and of deputies, or the optimal length of tenure for elected or appointed officials. For the reasons I set out both in my earlier article and in the first Sections of the present one, I do not believe that such attempts at fine-tuning are likely to succeed. They presuppose a knowledge of social causality as well as a determinate conception of the

[80] Bouandel (2005).

[81] Vermeule (2007), p. 62.

public interest that, in most cases, we simply do not possess. This being said, it would be absurd to be dogmatic on this issue. In some cases where a proposed reform would entail costs as well as benefits, one can confidently claim that the net effect will be positive. Examples include the abolition of life tenure for judges (an American practice created when life expectancies were less than half of what they are today) or of annual elections to parliament (also a practice that may have made sense in the 18th century but not in the 21st). By contrast, the merits of triennial versus quadrennial elections or of six-year versus eight-year tenure for judges on constitutional courts seem entirely indeterminate.

When a departure from the status quo can be predicted to have significant positive as well as negative effects, institutional designers have to tread carefully. Often, the combination of *uncertainty* concerning the long-term net effect of reform and the *transaction costs* of reform will provide a compelling argument for retaining the status quo. *The burden of proof is on the reformer*, simply because reform always has short-term costs. This emphasis on indeterminacy confers an undeniable conservative slant to my argument. It does not, however, amount to a "cult of complexity, with its inevitable strong suggestion that any but the most piecemeal and modest tinkering with the social mechanism [is] ill-fated".[82] The conservatism is offset by the radical character of many Benthamite proposals, such as the promotion of the active aptitude of deputies and the combination of secrecy ex ante and publicity ex post in voting systems. If carried out systematically, the program would imply a very thoroughgoing purge of many institutions.

This last remark triggers a second objection, however: who would be motivated to carry out this purge? Where would the *political will* come from? Once again we can take our cue from Bentham, who proposed elaborate devices for making the ruler of Tripoli adopt measures that would deprive him of most of his powers.[83] Although his proposals were utterly impractical, the question to which they were offered as an answer is a real one.

It seems safe to say that the question does not arise in jury reform. It is hard to see why legislators should be opposed to measures that reduce the impact of self-interest, passion, prejudice and cognitive bias on jurors.

It is also hard, although not quite as hard, to see why anyone would be opposed to measures, in any arena, that would reduce the impact of passion, prejudice and bias. On any given occasion, a group or an individual might wish these influences to operate, but in the long run there is little benefit to be had from irrational decision-makers. The main exception concerns prejudice: racist legislators have often, for example, adopted mechanisms favoring the choice of racist jurors.

[82] Novick (1988), p. 324.

[83] Bentham (1990).

The hardest question arises with regard to measures to curb the impact of self-interest in elections and, especially, in assemblies. Why would a parliament adopt legislation to prevent vote-buying if all parties believe that it might one day be in their interest to use this stratagem? Why would framers motivated by partisan interests create a central bank that would render monetary policy independent of partisan interests?[84] Why would a parliament abolish the *cumul des mandats* or adopt strict rules to punish members for being absent if all deputies would like to retain these options?

These problems might be overcome at the constitutional stage, but only if the framers follow the example of Solon and leave political life. At the post-constitutional stage, much depends on the nature of the issue and on the motivation of the deciders. Let me illustrate with two reforms to promote active aptitude in the American Congress, past and present. In the 19th century, some representatives may well have understood that public business would be carried out more efficiently if they sat on benches rather than at desks, yet opposed the reform because they cared more about carrying on their private business. In the 21st century, reforms limiting the time representatives spend on financing their reelection would enable them to carry out their public business more efficiently, with no sacrifice of their private interest.

I wish I could say more about the conditions for political will formation. If it is true, as Kant said, that nothing great is ever done without passion, and that passions are unpredictable and ungovernable, this limit to my analysis is perhaps inevitable.

References

Adams, W. P. (2001), *The First American Constitutions*, New York: Rowman and Littlefield

Amar, A. (1995), "Reinventing juries: Ten suggested reforms", *University of California Davis Law Review* 28, 1169-94.

AP = *Archives Parlementaires,* Serie I: 1787-1799, Paris 1875-1888

Ayres, I. and Donahue, J. (2003), "Shooting down the 'more guns, less crime' hypothesis", *Stanford Law Review* 55, 1193-1312

Bach, L. (2009), "Faut-il interdire le cumul des mandats?", Working Paper, Ecole d'Economie de Paris. (http://www.jourdan.ens.fr/~lbach/documents/Cumul_Bach.pdf)

[84] Przeworski and Limongi (1993).

Balinski, M. and Laraki, R. (2010), *Majority Judgment*, Cambridge MA: MIT Press.

Bentham 1989, *First Principles Preparatory to a Constitutional Code,* Oxford University Press

Bentham, J. (1990), *Securities against Misrule and other Constitutional Writings for Tripoli and Greece*, Oxford University Press

Bentham, J. (1999), *Political Tactics,* Oxford University Press

Bentham, J. (2002*), Rights, Representation, and Reform,* Oxford University Press

Berend, D. and Paroush, J. (1998), "When is Condorcet's jury theorem valid?", *Social Choice and Welfare* 15, 481-88.

Bouandel, Y. (2005), "Reforming the Algerian electoral system", *Journal of Modern African Studies* 43, 393-415.

Bryce, J. (1995*), The American Commonwealth*, Indianapolis IN: Liberty Fund.

Claussen, C. and Røisland, Ø. (2010 a), "A quantitative discursive dilemma", *Social Choice and Welfare* 35, 49-64.

Claussen, C. and Røisland, Ø. (2010 b), "The discursive dilemma in monetary policy", Bank of Norway Working Paper no. 5.

Cromwell, O. (1945), *Letters and Speeches*, Vol. 2, ed. T. Carlyle, New York: Wiley and Putnam.

Diamond, S., Casper, J. and Ostergren, L. (1989), "Blindfolding the jury", *Law and Contemporary Problems* 52, 247-68.

Dietrich, F. (2008), "The premises of Condorcet's Jury Theorem are not simultaneously satisfied", *Episteme* 5, 56-73.

Donahue, J. and Wolfers, J. (2005), "Uses and abuses of empirical evidence in the death penalty debate", *Stanford Law Review* 58, 791-46.

Donovan, T. (2008), "The Limbaugh effect", *The Forum* 6(2), 1-8.

Duff, P. (2000), "The Scottish criminal jury", in N. Vidmar (ed.), *World Jury Systems,* Oxford University Press, pp.249-82.

Elster, J. (2007 a), *Explaining Social Behavior*, Cambridge University Press.

Elster, J. (2009), "Excessive ambitions", *Capitalism and Society* 4 (2), 1-30.

Elster, J. (2012), "Constitution-making and violence", *Journal of Legal Analysis* 4, 000-000.

Elster, J. (2013), *Securities against Misrule*, Cambridge University Press.

Elster, J. and Roemer, J., eds. (1988), *Interpersonal Comparisons of Utility*, Cambridge University Press.

Ely, J, (1980), *Democracy and Distrust*, Cambridge MA: Harvard University Press.

Farrand, M. (1966), *Records of the Federal Convention*, New Haven CO: Yale University Press.

Fraenkel, J. and Grofman, B. (2006), "Does the alternative vote foster moderation in ethnically divided societies?", *Comparative Political Studies* 39, 623-51.

Føllesdal, D. (1979), "Some ethical aspects of recombinant DAN research", *Social Science Information* 18, 401-19.

Garrett, B. (2011*), Convicting the Innocent*, Cambridge MA: Harvard University Press.

Gibbon, E. (2005), *The History of the Decline and Fall of the Roman Empire*, London: Penguin.

Goodin, R. (1986), "Laundering preferences", in J. Elster and A. Hylland (eds.), *Foundations of Social Choice Theory*, Cambridge University Press, pp. 75-102.

Goodsell, C. (1988), "The architecture of parliaments", *British Journal of Political Science* 18, 287-302.

Grofman, B. and Feld, S. (1989), in D. Estlund et al., "Democratic theory and the public interest: Condorcet and Rousseau revisited", *American Political Science Review* 83, 1328-40.

Grossman, S. and Stiglitz, J. (1980), "On the impossibility of informationally efficient markets", *American Economic Review* 70, 393-408.

Hausman, D. (2011), *Preference, Value, Choice, and Welfare*, Cambridge University Press.

Hoffrage, U. et al. (2000), "Representation facilitates reasoning", *Cognition* 84, 343-52.

Horowitz, D. (2007-2008), "Conciliatory institutions and constitutional process sin post-conflict states", *William and Mary Law Quarterly* 49, 1213-48.

Hyman, R. and Brough, I. (1975), *Social Values and Industrial Relations*, Oxford: Blackwell.

Quarterly Journal of Economics 118, 843–77.

Kalven, H. and Zeisel, H. (1966), *The American Jury*, University of Chicago Press.

Karotkin, D. and Paroush, J. (2003), "Optimal committee size", *Social Choice and Welfare*, 20, 429-41.

Kornhauser, L. and Sager, L. (1986), "Unpacking the Court", *Yale Law Journal* 96, 82–117

Lange, A. et al. (2010), "On the self-interested use of equity in international climate negotiations", *European Economic Review* 54, 359-75.

López-Guerra, C. (2011), "The enfranchisement lottery", *Politics, Philosophy, & Economics 10,* 211-33.

Miller, W. (1996), *Arguing about Slavery*, New York: Knopf.

Montaigne, M. de, (1991), *Essays*, London: Allen Lane.

Mueller, D. (2003), *Public Choice III*, Cambridge University Press

Mukhopadhaya, K. (2003), "Jury size and the free rider problem", *Journal of Law, Economics & Organization* 19, 24-44.

Note (2005), "Trumping the race card: Permitting criminal defendants to remain anonymous and absent from trials to eliminate racial jury bias", *Georgetown Journal of Legal Ethics 18* (2005), 1150-60.

Novick, P. (1988), *That Noble Dream: The 'Objectivity' Question and the American Historical Profession*, Cambridge University Press.

Page. S. (2007), *The Difference*, Princeton University Press

Paine, T. (1791), *The Rights of Man*.

Pierre, E. (1893), *Traité de droit politique electoral et parlementaire*, Paris.

Poisson, S.-D. (1837), *Recherches sur la probabilité des jugements en matières*

Posner, R. (1988), "The Constitution as an economic document", *George Washington Law Review* 56, 4-49.

Przeworski. A. and Limongi, F. (1993), "Political regimes and economic growth", *Journal of Economic Perspectives* 7, 51-69.

Rawls, J. (1971), *A Theory of Justice*, Cambridge MA: Harvard University Press.

Reilly, B. (2011), "Centripetalism" in S. Wolff and C. Yakinthou (eds), *Conflict Management in Divided Societies*, London: Routledge, pp. 57-65.

Riker, W. and Brams, S. (1973), " "The paradox of vote trading", *American Political Science Review* 67, 1235-1247.

Risinger, D. (2007), "Innocents convicted", *Journal of Criminal Law and Criminology* 97, 761-806.

Tocqueville, A. de (2004), *Democracy in America*, New York: Library of America.

Troper, M. (2006), *Terminer la Révolution: La Constitution de 1795*, Paris: Fayard.

Vacca, R. (1921), « Opinioni individuali e deliberazione collettive », *Rivista Internazionale di Filosofia del Diritto* 52, 52-59.

Vermeule, A. (2007), *Mechanisms of Democracy*, Oxford University Press.

Vidmar, N, and Hans, V. (2007), *American Juries: the Verdict, A*mherst, NY: Prometheus Books.

Vologh, A. (1997), "n guilty men", *University of Pennsylvania Law Review* 146, 173-218.

Weitzman, M. (2009), "On modeling and interpreting the economics of catastrophic climate change", *Review of Economics and Statistics* 91, 1-19.

Woolrych, A. (1986), *Soldiers and Statesmen*, Oxford University Press.

Young, J. (1986), *The Washington Community 1800-1828*, New York: Columbia University Press.