Introduction

Securities Against Misrule was the title of a work by Bentham from 1822–23, one of several writings on a proposed constitution for Tripoli. The main idea was already stated in Political Tactics (1789):

The tactics of political assemblies form the science . . . which teaches how to guide them to the end of their institution, by means of the order to be observed in their proceedings.

In this branch of government, as in many others, the end is so to speak of a negative character. The object is to avoid the inconveniences, to prevent the difficulties, which must result from a large assembly of men being called to deliberate in common. The art of the legislator is limited to the prevention of everything which might prevent the development of their liberty and their intelligence.¹

The aim of the present book is to develop, generalize, and, to some degree, modify the claim made in the sentence I have italicized.

Before deciding on the title, I thought of calling the book Bentham against Condorcet. In Chapter 1 I argue, in fact, that the social-choice tradition stemming from Condorcet has little relevance for the normative theory of collective decision making. In Chapter 3 I make a similar claim with regard to the Condorcet Jury Theorem. By contrast, Bentham’s relentless realism and his insight into the operations and the failings of actual institutions provide, or so I shall argue throughout the book, a superior guide to institutional reform. To get a feeling for the drift of my argument, readers are encouraged to consult the list—which is far from exhaustive—of Benthamite schemes that I offer in the Conclusion (p. 272).

In Bentham’s view, the object of institutional design is security against misrule, or the prevention of mischief—the removal of obstacles that will thwart the realization of the greatest good for the greatest number. In a sense, I go further than he did, since I do not stipulate any particular goal to which the removal of these obstacles is conducive. I shall argue that once collective decisions have been

¹ Bentham (1999), p. 15. For Bentham, the subject matter of “tactics” was not strategic behavior (which I discuss in Chapter 1), but the design of political institutions that promote the “greatest happiness of the greatest number.”
sheltered as much as possible from distortions, we have to let the chips fall where they may. In this respect I follow the memorable Conclusion to John Hart Ely’s *Democracy and Distrust*, which I cite and discuss in the Conclusion. The distortions stem from *self-interest* (”sinister interest” in Bentham’s phrase) and from *irrationality*: passion, prejudice, and bias.

Ely and, above all, Bentham are the direct sources of inspiration for this book. Other partial ancestors or precursors can also be cited. Given that action depends on opportunities as well as on motives, institutional designers can try to prevent undesirable actions either by acting on the motives of the agents (the incentive approach) or by limiting their opportunities to do harm (the Benthamite approach). The American framers used both strategies. In Chapter 5 I provide an example of how Madison relied on incentives in designing the mechanism for electing the President. According to an authoritative interpretation of *The Federalist*, he also proposed to act on opportunities:

> Madison’s idea that a republican government will deprive corrupt representatives of the opportunity to carry out schemes of oppression is evident in Number 63, when he says that even if indirectly elected senators were to be impelled by corrupt motives, the system of government provided for in the Constitution would deny them the opportunity to act corruptly. It is hard to exaggerate the extent to which Madison and his fellow authors employed the concepts of motive and opportunity.

In a very different way, the Benthamite approach may also be seen as a generalization of Habermas’s idea of “the uncoerced force of the better argument.” It is a generalization of his idea, because it takes account of other distorting factors than coercion. Coercion – the threat of violence – may induce either prudential, interest-based fear or visceral, emotion-based fear (see Chapter 4). My conjecture – based on the lack of any sustained discussion by Habermas of the role of emotions in public life – is that he refers to prudential fear. In that case, his theory needs to be supplemented by an account of emotion. Also, bribes can be as powerful as threats. Plutarch tells us that Solon “with good reason [thought] that being seduced into wrong was as bad as being forced . . . , since both may equally suspend the exercise of reason” (Solon XXI.4). Institutional designers should insulate the decision makers from both, as well as from the other distortions I mentioned.

The idea of “preventing the prevention of intelligence” has a strong verbal resemblance to the claim of the British idealists that the main task of the state is to “hinder hindrances” or “remove obstacles” to the free development of the individual. Yet these writers did not focus, as I shall, on removing obstacles to collective decision making. They did, for instance, endorse restrictions on the sale and purchase of alcohol so that individuals “may become more free to exercise the faculties and improve the talents which God has given them.”

---

3 Green (1986), p. 212. I am grateful to Claus Offe for directing my attention to Green’s work.
Introduction

From the Benthamite perspective, the relevant task would be to limit the sale of alcohol to jurors, deputies, and voters. Let me give some examples. A handbook for jurors in Minnesota tells them, “While serving as a juror, do not drink alcoholic beverages during trials breaks.” In *Tanner v. United States*, the Supreme Court found that copious consumption of alcohol by jurors during a trial did not constitute grounds for overturning the verdict: “However severe their effect and improper their use, drugs or alcohol voluntarily ingested by a juror seems no more an ‘outside influence’ than a virus, poorly prepared food, or a lack of sleep.” (I return to this amazing piece of bad reasoning in Chapter 1.) From the Benthamite perspective, the Minnesota ex ante ban solves the problem. When the French constituants abolished feudalism on the night of August 4, 1789, some of them may have been, in the words applied to a later self-denying ordinance, “drunk with disinterestedness.” In addition, it is likely that a good portion were drunk in the literal sense of the term. One historian claimed that a few days later, “to avoid nocturnal and intemperate deliberations, and the reproaches that good Frenchmen who are not deputies could make to good patriots for deliberating upon coming from the dinner table, it was decided that in future votes on important matters the assembly should be fasting (*à jeun*).” Many countries and American states, finally, ban the sale of liquor on election day, partly no doubt because of the historical practice of candidates plying voters with liquor and partly, I assume, because voters are believed to be more open to reason when not “under the influence.”

The natural standard for assessing the institutions I discuss in this book – jury trials, political assemblies, and electoral systems – might seem to be whether they tend to produce good outcomes. In Chapter 1 I discuss some of the problems raised by that approach. 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 their tendency to promote the chosen conception of goodness. To do so, we need a causal theory (a purely non-consequentialist theory will never be sufficient in institutional design). 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

---

4 Lebègue (1910), p. 261, referring to the decision by the Assembly on May 16, 1791, to render its members ineligible to the first ordinary legislature (see Chapter 4).
7 This is not a problem if the good outcome is defined as finding a factual truth, as in jury verdicts. The Condorcet Jury Theorem is often, in fact, held up as an example of how institutional design can promote good outcomes. I believe the relevance of the theorem to actual decisions is minimal. Also, as I explain in Chapter 3, Bentham spotted a possibly fatal weakness in Condorcet’s argument.
is propounded. The large number of competing theories should, however, undermine our confidence that any of them is right.8

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.9 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.”10 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.11

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 these interests coincides with the general interest, so that it may present itself to the world and gain ready acceptance.

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.12 Although in

8 “Where we have several competing theories, which give different predictions, all these theories should be regarded with suspicion, and we should be prepared for a risk that is higher than what is predicted by any of the theories” (Føllesdal 1979, pp. 405–6).
9 Hyman and Brough (1975), p. 49.
11 Tocqueville (2004a), p. 202. Anticipating and influencing a more famous statement by Pascal, Montaigne (1991), p. 615, asked, “What kind of truth can be limited by a range of mountains, becoming a lie for the world on the other side?” Tocqueville might have observed that this relativism creates an incentive to cross the mountain. Under the entry “Morals,” Bierce (2002), p. 166, cites the following passage from the mythical Gooke’s Meditations: “It is sayd there be a raunge of mountaynes in the Easte, on one syde of the which certayn conductes are immorall, yet on the other syde they are holden in good esteeme; whereby the mountaineer is much conveynyenced, for it is given to him to goe downe eyther way and act as it shall suite his moode, withouten offence.”
12 For a general criticism of many uses of statistical models (admitting my lack of firsthand expertise), see Elster (2009d). Concerning the death penalty and handguns see, respectively, Donohue and Wolters (2005) and Ayres and Donohue (2003).
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.

Being aware of these temptations and suspecting oneself of being subject to their influence, what should one do? Mme de Staël wrote that in “this world, there is no greater trial for morality than political employment, for the arguments one can invoke on this subject, to reconcile one’s conscience with one’s interest, are without number. Yet the principle from which one ought not to deviate is to support the weak; one rarely goes wrong in orienting oneself by this compass.” The Norwegian writer Helge Krog told his readers that when in doubt, they should adopt the policy that would harm themselves the most. Along the same lines, Proust refers to “the soldier who chooses the post not where he can be of most use but where he is most exposed.” While these reactions may seem morally admirable, they can lead one astray. Pascal wrote, “The most equitable man in the world is not permitted to be judge in his own cause: I know some who, in order not to be entrapped by this amour-propre, have been as unjust as possible by a counter-bias; the sure way to lose a perfectly just cause was to get it commended to them by their near kinsfolk.” Also, and more relevant for my purposes, the injunction “Distrust thyself” has no clear institutional implications. Although the idea of distrust points in the right direction, it needs to be spelled out in a different way.

My proposal is to consider procedural accounts of good institutional design. I shall discuss three versions of this idea, and opt for the last and least ambitious.

First, we might ask whether institutions can be designed to select good decision makers. “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.” In later chapters, I discuss some proposals to this effect. Toward the end of Chapter 2 I discuss procedures intended to select competent jurors. Chapter 5 as a whole is devoted to a device – cross-voting – that has been adopted at various times and places to generate impartial representatives from an electorate of partial voters. Although the idea is intuitively attractive, it has rarely worked well. Trivially, members of expert bodies such as central bank committees are selected on grounds of competence. In this book, however, I am not concerned, except occasionally, with expert decision making, but with democratic decisions.

15 Pensée 78, ed. Sellier; my italics. See also the comments in Chapter 1 on the tendency of the French revolutionaries to think or to claim that the only disinterested attitudes were counterinterested ones.
16 Montaigne (1991), p. 1057. He believes, of course, that it can’t be done.
Second, we may ask whether institutions can be designed to eliminate bad decision makers, such as the biased, the prejudiced, and the incompetent. Trivially, this is what happens when children are not allowed to vote. In criminal jury selection, the aim is not to eliminate jurors who would be unsuitable in any trial, but only those with biases and prejudices that might be triggered by particular aspects of a given case or a given defendant. Allegedly, female jurors are hard on female defendants, middle-class jurors soft on middle-class defendants, and overweight jurors either hard or soft on overweight defendants. One famous American defense attorney, Clarence Darrow, preferred older jurors: they “are generally more charitable and kindly disposed than young men.”

Another, Samuel Leibovitz, preferred them young: “They’re not set in their ways. They are tolerant.”

As these examples suggest, there are no hard facts in this area. If jury selection were an exact science, the challenges by prosecutors and defense attorneys would tend to cancel each other in their effects, the only effect being wasteful costs and delays. If, as I believe, it is not, it does not even have the virtue of neutrality.

Third, we may ask whether decision making by ordinary citizens can be improved by structuring their relation to the environment. This process involves three main procedures: preventing the decision makers from learning certain things (ignorance), preventing others from learning certain things about them (secrecy), and enabling others to learn certain things about them (publicity). These are central notions throughout the book. As we shall see, secrecy and publicity are relative to an audience and to a point in time. A vote by a juror, for instance, may be known to other jurors but not to the public at large. It may be unknown to other voters at the time of voting but be made known to them afterward.

In addition, rotation can serve the function of insulating decision makers from corrupting influences, a technique used, among other places, in the Roman Empire, the Chinese empire, and in the French ancien régime. 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.” Rotation may need to be supplemented by a mechanism – election or randomization – that makes it difficult to predict who will replace the departing decision maker. In fact, “rapid rotation of assignments can lead to ‘judge shopping’ by litigants through the mechanism of delaying a case until judicial assignments have been rotated.” Yet, as I shall argue,

---

18 Cited in ibid., p. 199.
19 For these cases, see, respectively, Olivier-Martin (2010), pp. 64, 106; Skinner (1975), p. 341; and Olivier-Martin (2010), pp. 267, 626. In China, rotation was supplemented by the “law of avoidance” that kept officials from serving in their native provinces.
because rotation can also undermine epistemic competence it can be difficult to tell when it is appropriate.

In the cases I shall consider, the requirement of ignorance applies mainly to jurors. As we shall see in Chapter 2, members of a jury are not supposed to know certain facts about the case, the defendant, or the law. (Sometimes they are also, more paradoxically, requested to act as if they did not know some facts they do know.) Although the jury seems to be the most important institution in which a group of individuals is asked to make decisions behind an artificial “veil of ignorance,” other examples can be cited. Holding the assets of politicians in blind trust offers one illustration. The use of need-blind committees for college admission and of gender-blind committees for musical auditions relies on the same principle. Double-blind line-ups in criminal investigations and double-blind medical studies, while important, do not involve collective decisions.\(^2\)

The main reason for inducing ignorance in the jury is the belief that knowledge would trigger passions, biases, or prejudices that are likely to have a negative impact on the decision. A standard example is knowledge of a defendant’s previous criminal record. A secondary reason is that the admission of certain kinds of evidence might have undesirable incentive effects on cases other than the one that is being decided. A standard example is the ban on eating from the fruit of the poisoned tree.

The use of secrecy aims at eliminating undesirable incentives – bribes, threats, social ostracism – for deciding one way rather than another. To be effective, secrecy typically has to be mandatory rather than optional, since a demand for secrecy often allows others to infer how the person requesting it will vote. Also, if a person has received a bribe to vote in a certain way, she can hardly then opt for secrecy.

In juries, deliberations are shrouded in secrecy, always during the trial and often afterward. Except when unanimity is required, the identity of those who voted for and against conviction is usually also kept secret. The jury itself may practice secret voting. In some countries, this procedure is mandatory; in others it is up to the jurors to decide whether they shall decide by public or secret voting. Little seems to be known about optional jury practices, except that it is apparently common to begin the deliberations by a secret straw ballot.

In elections, even when secrecy is mandatory, it is rarely perfect, or believed to be perfect. The micro-technology of the secret ballot can be decisive.\(^3\) In

\(^2\) A veil of ignorance can also be a useful device for preventing data-mining and curve-fitting: “you put half the data in cold storage, and look at it only after deciding which models to fit” (Freedman 2005, p. 64). For some reason, this salutary but costly procedure (also called out-of-sample testing), which is “commonplace in the physical and health sciences,” is “rare in the social sciences” (ibid.). Not only may social scientists be subject to ordinary self-deception when they mine the data, but even to second-order self-deception (“I never deceive myself”) when they believe they have no use for a tool that would keep them honest.

\(^3\) For Imperial Germany, Leemann and Mares (2011) demonstrate the importance of the color, shape, and transparency of ballots as well as the design of urns. Pierre (1893) offers numerous examples from France.
Colombia, electoral fraud is made possible by the use of mobile phones to take a picture of the ballot and send it to the briber. “In Chiapas, Mexico, during the 2000 elections … voters worried that the ruling party had a satellite watching the ballot box or that ballots could be linked to them by finger prints.”

In 2010, the Icelandic Supreme Court invalidated the elections to the Constitutional Council on the grounds that the secrecy of the vote had been jeopardized in five different ways. Yet compared to previous centuries, the secret ballot has undoubtedly reduced the power of the state, political parties, local communities, or powerful individuals to shape voting behavior. One might also extend the logic of the secret ballot to campaign contributions, by imposing donor anonymity.

In some American states this idea has been used, with mixed success, in the election of judges.

It may be worthwhile emphasizing the difference between secrecy and privacy. Carl Schmitt confuses these two notions when he argues against the secret ballot in elections by contemplating a possible extension of the practice:

It is fully conceivable that one day through ingenious discoveries, every single person, without leaving his apartment, could continuously express his opinions on political questions through an apparatus and that all these opinions would automatically be registered by a central office, where one would only need to read them off. That would … provide a proof that the state and the public were fully privatized. It would not be public opinion, for even the shared opinion of millions of private people produces no public opinion … In this way, no common will arises, no volonté générale; only the sum of all individual wills, a volonté de tous, does.

Today, of course, that apparatus exists in the form of the Internet. The vote in the privacy of one’s apartment could not always, however, be kept secret. A vote buyer could send his agent to the home of the voter to verify that she votes for him; and ten agents might reach enough voters to tip the balance in a local election. The very publicity of the voting process makes the secret vote possible. Together with the requisite technology, voting of officials recruited from different political parties, monitoring the voters and each other, are the guarantors of the secret ballot. Schmitt compounds his mistake when he asserts that

---


85 Axelsson (2011). The Court did not assert that the secrecy had actually been violated but, using the tenuous possibilistic reasoning discussed in Chapter 1, claimed that it could have been. In all likelihood, the Court’s decision was politically motivated (Gylfason 2012).

86 As suggested by Ayres (2000), who also discusses some formidable technical obstacles to the implementation of the idea. In this context, donor anonymity means that neither the candidate nor the public at large would know the identity of the donors. Alternatively, as argued by the Supreme Court in the otherwise disastrous Citizens United decision in 2010, one could require full public disclosure so that voters could decide for themselves “whether elected officials are ‘in the pocket’ of so-called moneyed interests.” The current situation is inferior to both options.

87 Banner (1988).

88 Schmitt (2008), p. 274. Schmitt’s implicit view that Rousseau was an opponent of secret voting does not seem to have a firm basis in the texts (see note 64 in Chapter 1).
“the current electoral and voting secrecy is not a genuine secret at all. According to their discretion, the voters can decide whether to disclose and make this secret public; its preservation is only a right, not a duty of the state citizen.” 29 From the point of view of democratic theory, however, the crucial question is whether the citizen can credibly communicate her vote to a prospective vote buyer, since if she cannot he will not pay for it. She can do so in the privacy of her home, not in the public space of the polling booth. 30

In contemporary assemblies, secrecy is exceptional. Citizens can listen to the debates in public galleries, and votes are public and usually recorded. In most assemblies, discussions are reported verbatim, allowing only for correction of grammatical mistakes and elimination of redundancies. In the American Congress, representatives can alter their speeches before they go into the Congressional Record, but in recent years revised remarks have been printed in a different typeface. Historically, assemblies have not always granted publicity of their debates. 31 They have often denied the public access to their proceedings and also used secret voting to allow their members to shield themselves from the public, from the executive, or from the political parties to which they belong. As I argue in Chapter 4, secrecy may in fact be a desirable feature of constituent assemblies. In ordinary legislatures, though, publicity is usually essential to prevent self-dealing. 32

Let me illustrate the pressure on assemblies to deliberate and vote in public. In 1401, Sir Arnold Savage complained that certain members of the House of Commons reported their deliberations to the King. 33 In 1532, Henry VIII forced a division in the House of Commons, that is, a physical separation of the members, that enabled him to identify who voted his way and who voted against him. “Several who had earlier been in opposition joined the ‘yeas’ for fear of the king’s indignation, and a majority was obtained.” 34 In 1633, Charles I experienced

29 Ibid.
31 See Vermeule (2004) for overview and discussion of American practice. Pierre (1893), pp. 824–25, 834–37, has a full discussion of the practices in the French Senate and National Assembly. The principle of publicity was often undermined by the transformation of the assembly into a “secret committee” at the request of a small number of deputies (ranging from 5 to 20). Before 1870, the names of the requesters were not inserted into the record. Similarly, before the abolition of the right to demand a secret vote, the insertion of the names of the requesters into the record was refused as being contrary to the principle of the secret ballot (ibid., p. 1010). Demands for secrecy were to be kept secret. Conversely, when deputies or senators demanded a public vote (written or viva voce from the tribune), their names had to be entered into the record so that “it would not be possible to impose a long and fastidious process on the assembly in secondary matters without incurring accountability” (ibid., p. 1027). Demands for publicity had to be made publicly.
32 Ignorance can achieve the same aim, 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.
33 Adams and Stephens (1914), p. 172.
difficulties getting his most cherished measures through the Scottish Parliament, despite the fact that he was present in the assembly and “ostentatiously took down the names of those who spoke against them.” In 1641, “[The] trial of Strafford . . . produced the first exercise in parliamentary history of the coercive use of publicity of members’ votes in the chamber. The names of fifty-nine M.P.s who had voted against Strafford’s attainder were posted up on the Royal Exchange under the title, ’The names of those men, who to save a Traytor, would betray their Country.’ It is hardly surprising that attendance in Parliament began to fall.”

In two constituent assemblies, demands for public voting came from representatives branding the threat of popular action. In 1789, radical members of the Assemblée Constituante insisted on roll-call voting on controversial issues to scare their opponents into voting against their conviction. Although no framers were killed or physically injured, many feared for their lives and for those of their families. To explain why some members voted against bicameralism, a liberal deputy for the nobility wrote to his constituency, “Some deputies from the third estate have told me, I do not want my wife and children to have their throats cut.” When radicals in the 1848 Frankfurt constituent assembly used the same tactic, “the work of parliament was . . . threatened by the hunt of the crowd for unpopular members of the assembly.” One member of the Center Right was beaten up and two killed.

In modern political systems, parties are usually able to enforce discipline on deputies, since voting in Parliament is almost invariably public. Until 1988, the Italian Parliament was a rare exception to this rule. In that year, “party leaders promoted the abolition of secret vote as it prevented them from monitoring organized intra-party factions that used secret voting as a tool to undermine incumbent governments.”

To my knowledge, when the publicity of debates has been induced by pressure, it has always been by popular demand. The principal naturally wants to know not only how its agents are voting but also what they are saying. Reporting of debates and access of the public to assemblies were nevertheless contentious issues in British and American history for centuries. Deputies claimed they were fighting a two-front war against the executive and the people and needed to be protected from both. Progress was interrupted by reversals until by the early 19th century publicity imposed itself irreversibly.

In modern societies, publicity and the related norm of transparency impose themselves in an increasing number of domains. Historically, the secrecy of

---

37 Lally-Tolendal (1790), p. 141.
38 Eyck (1968), p. 312.
39 Gianetti (2010).
40 Pole (2008), chs. 4 and 5.
41 There does not seem to be a well-established distinction between publicity and transparency as features of decision-making processes. In an administrative context, when the reasons for a decision are made known to the individual whom it concerns but not to anyone else, intuition...