

The anatomy of government failure

William R. Keech¹ · Michael C. Munger^{1,2}

Received: 9 March 2015 / Accepted: 29 April 2015 / Published online: 22 May 2015
© Springer Science+Business Media New York 2015

Abstract Government failure is a much bigger problem than its contemporary treatment implies. Setting aside natural disasters, most of the great catastrophes of human history have been government failures of one sort or another. We argue that many so-called market failures are government failures because government defines the institutions in which markets succeed or fail. The concept of government failure has been trapped in the cocoon of the theory of perfect markets. Narrowly defined deviations from market perfection have been designated market failures, for which government corrections may or may not really be a solution. Government failure in the contemporary context means failing to resolve a classic market failure. We propose an alternative approach for evaluating whether government fails: the Pareto standard. If an available Pareto improvement is not chosen, or is not implemented, that is a government failure. We organize government failure into two types: *substantive* and *procedural*. Substantive failures include the inability or unwillingness to maintain order, to maintain sound fiscal and monetary policies, and to reduce risks of transaction costs, which we classify as corruption, agency and rent-seeking. Procedural failures are inadequacies of available social choice mechanisms, causing collective decisions to be arbitrary, capricious, or manipulated. We conclude with some reflections on human rationality and the implications of behavioral economics.

Keywords Market failure · Government failure · Competitive equilibrium theory · Collective choice mechanisms

JEL Classification A10 · A12 · D02 · D6 · H1

✉ Michael C. Munger
munger@duke.edu

William R. Keech
wrk3@duke.edu

¹ Department of Political Science, Duke University, Durham, USA

² Department of Economics, Duke University, Durham, USA

1 Introduction

1.1 Overview

Aside from natural disasters like tornadoes, floods, and earthquakes, the greatest catastrophes of human history have been institutional failures for which government, or the state, carries the basic responsibility. We call these “government failures.” Government failures are more frequent, and frankly more fundamental, than the “market failure” problems on which many economists focus.

Government failure has an intuitive meaning. For example, World War I and Mao’s Great Leap Forward both were government failures, because their consequences were (predictably) perverse, but were initiated by governments anyway, leading to enormous destruction and loss of life.¹ Our standard of failure, then, is not just whether state action achieves its object, but also whether that object is consistent with justifications of political authority that rest on solving problems of collective action. More simply, government may fail to do things it should do, or government may do things it should not do.

This approach differs from the subordinate treatment government failure receives in most contemporary economics. That treatment begins with a theory of perfect markets, follows with a theory of market failures, and follows *that* with conjectures about actions government might take to resolve the market failure. By that reasoning, government failure occurs when state interventions exacerbate, or simply fail to resolve, a market failure.

Such a focus on markets is not wholly misplaced, of course. Price systems can and do play an important role in coordinating economic activity, and in improving human well-being. They sometimes fail in a way that is captured accurately by the theory of market failure.² But

¹ Rummel (1997, chap. 1) estimates the total “demicides” or “government mass murders” during the twentieth Century as being well over 150 million worldwide.

² For a seminal example, see Bator (1958). There are many examples of the notion of government failure preceding or causing market failure, such as de Mariana (Mariana 1839, pp. 232–236). Juan de Mariana noted that there had been a disruption in markets, causing dissatisfaction with merchants, but that the “real” cause was the sovereign printing far too many coins of debased metals. One of the earliest explicit comparisons is Coase (1964, p. 195; emphasis added), who said:

Contemplation of an optimal system may provide techniques of analysis that would otherwise have been missed and, in certain special cases, it may go far to providing a solution. But in general its influence has been pernicious. It has directed economists’ attention away from the main question, *which is how alternative arrangements will actually work in practice*. It has led economists to derive conclusions for economic policy from a study of an abstract of a market situation. It is no accident that [economists] find a category “market failure” but no category “government failure.” Until we realize that we are choosing between social arrangements which are all more or less failures, we are not likely to make much headway.

Demsetz (1969, p. 1; emphasis in original) called this kind of comparison the “nirvana” fallacy:

The view that now pervades much public policy economics implicitly presents the relevant choice as between an ideal norm and an existing “imperfect” institutional arrangement. This nirvana approach differs considerably from a *comparative institution* approach in which the relevant choice is between alternative real institutional arrangements. In practice, those who adopt the nirvana viewpoint seek to discover discrepancies between the ideal and the real and if discrepancies are found, they deduce that the real is inefficient. Users of the comparative institution approach attempt to assess which alternative real institutional arrangements seems best able to cope with the economic problem; practitioners of this approach may use an ideal norm to provide standards from which divergences are assessed for all practical alternatives of interest and select as efficient that alternative which seems most likely to minimize the divergence.

For other important background and problems with consistency in the “typical” economics approach see Backhouse and Medema (2012).

whether markets succeed or fail, they operate in a larger context of institutional rules that are set up and enforced by government. Markets may fail because governments interfere too much, regulate too little, or act badly, distorting the processes of production and exchange.

Public discourse sometimes touches market failures and government failures construed more broadly. In our view, such discussions often posit a false dichotomy—more government versus more markets—in ways that are influenced superficially by ideology and political values. For example, little agreement has emerged about whether the recent financial crisis (2007–2009) was a government failure or a market failure. In our view, it was a system failure, with elements of both. But if our claim that government defines the rules under which markets operate is valid, then the government failure is causally prior. The real question is not whether we need more or less regulation, but how the rules of the game should be structured in terms of incentives, property rights, and constraints.

The underlying problem is that government must be strong enough to maintain order, and to pass and then enforce laws and property rights systems that encourage coordination and solutions to collective action problems.³ Governments that fail are generally *both* too big, *and* too small, at the same time. By doing too much, attention and resources are spread too thinly across too many activities. But then regulations or policies that should serve important functions receive too little enforcement emphasis. Perhaps worst of all, other regulations created out of momentary political alignments or reactions to emergencies—regulations that have long outlived their original purpose—survive because citizens or industries have become dependent on the protection of artificial rents.

For these reasons, our approach begins with the simplest realistic choice setting. We will assume that a “dictator”—a unitary actor with well-defined objectives for solving collective problems—is in charge of the state. We then evaluate outcomes using the Pareto criterion, which anchors our approach in the grand tradition of competitive equilibrium theory (CET) in economics. CET is a theory of market success, and a benchmark for comparison. There is, as far as we know, no analogous theory of government success to use as a benchmark. But allowing a dictator to strive to limit outcomes to Pareto optima, and then to allow the selection among Pareto optima using side payments or compensation allows an (idealized) metric that identifies government failures with considerable analytic precision.

Our “dictator” is an analytic device used to solve three different problems: incentives, information, and capacity. As a consequence, dictators are variously assumed to be (respectively) benevolent, omniscient, or omnipotent.

Benevolence means that the goals are social as well as individual well-being, and in this paper, we will use the Pareto standard—with compensation—to define benevolence. Policies that fail to achieve this result are then government failures. We will posit the existence of a benchmark we will call the “Optimal Pareto Optimum” (OPO) and assume that achieving the OPO is the dictator’s objective.⁴

³ See Cooter and Siegel (2010) for such an interpretation of the U.S. Constitution. On the other hand, see Winston (2006) for some evidence that this may not be what states actually do.

⁴ This is an extreme assumption, and might be quarreled with. But we are trying to establish a benchmark against which success or failure might be measured. Three things should be said: (1) These assumptions are no more unrealistic than those used to generate CET. But when those assumptions fail to be valid, we blame markets, not the assumptions. (2) This approach allows us to finesse problems of “second best,” though of course this comes at the price of assuming that a single optimum obtains and that preferences are monotonic, so that movements toward the optimum are improvements. (3) An alternative, one we endorse for future work, might be an extension of Dixit-Stiglitz “constrained optima,” and the associated preferences (Dixit and Stiglitz 1977). We thank Robert Grafstein for clarifying our thinking on this point. Some related arguments are reviewed and discussed by Buchanan (1983).

Omniscience is another matter. In our view, no one person—benevolent or not—could possibly have all the relevant knowledge about values, demand, and supply, to maximize welfare. One version of this question is the famous “socialist calculation debate,” settled for many people by Hayek’s (1945) equally famous argument that only markets can provide a robust discovery process to aggregate dispersed knowledge. This may be true, and in fact that could be a reason why a benevolent dictator might choose to use markets as a practical (though imperfect) substitute for omniscience or synoptic knowledge. But it is not necessary to argue that markets *can* solve the problem to assert that a dictator *cannot* solve the problem, at least not alone. No central authority could be omniscient, so even a benevolent dictator must choose among institutional arrangements as a means of eliciting and aggregating information (Hayek 1945, Myerson 1979).

Omnipotence is as problematic as omniscience. In a large and complex society, even a benevolent dictator needs the help of agents in implementing Pareto optimal policies. But it is fatally unrealistic to think that her benevolent policies will be carried out as intended. Some agents may be corrupt, or shirk because they are lazy or do not share the benevolent leader’s goals, or may redirect their energies toward rent-seeking—all likely sources of government failure.

It is useful to pause to consider what now might be called “The Tullock Objection,” or the assumption that people in the public sector (elected officials, bureaucrats, interest groups, voters, and so on) are motivated by anything other than self-interest.⁵ In simple terms, there is no such thing as a “benevolent dictator.” Buchanan started this “objection” when he (re)discovered Knut Wicksell in the University of Chicago’s Harper Library in 1947. Buchanan read German, and when he found “A New Principle of Just Taxation” he said that he felt as if “the scales had fallen from my eyes” (Buchanan 2001). Wicksell (1896, p. 82 emphasis added) said:

[W]ith very few exceptions, the whole theory of [public finance] still rests on the now outdated political philosophy of absolutism. The theory now seems to have retained the assumptions of its infancy...when absolute power ruled most of Europe...

Even the most recent manuals on the science of public finance frequently leave the impression of *enlightened and benevolent despotism*...

Public choice has developed this insight, particularly in the work of Gordon Tullock, by relying on the assumption of behavioral symmetry, meaning that elected officials and bureaucrats are no worse, but no better, than private citizens, because they are no different. Public officials are no more benevolent than anyone else, and therefore the idea of a central authority that seeks single-mindedly to advance the public interest is fanciful.⁶

⁵ We thank Georg Vanberg for helping us to develop this line of argument. For the Tullock side, see Tullock et al. (2002), particularly as it was extended by Boettke et al. (2007). An interesting counterpoint on “public interest” is Becker (1983).

⁶ A parable often attributed to Gordon Tullock is the story of the “Roman Emperor.” Here we use the version told by Boettke et al. (2007):

The tale of the Roman Emperor, often invoked by Gordon Tullock, summarizes the problem of pre-public choice political economy well. According to this tale a Roman Emperor is asked to judge a singing contest between two contestants. Upon hearing the first contestant sign, the Emperor awards the prize to the second singer under the assumption that she clearly cannot be any worse than the first. But the Emperor’s assumption is quite possibly off the mark; the second signer could in fact be much worse. This parable highlights the proposition that imperfect markets do not necessarily justify government intervention. The consequences of this methodological demand for behavioral symmetry were therefore damaging to the government-as-corrective conclusion of the previous generation of public economists, and augmented the unintended consequences critique of government intervention.

Nonetheless, we adopt the convention of a benevolent dictator. One reason is that the assumption remains common, more than a century after Wicksell's objections and 50 years since attempts at refutation by development of the Public Choice literature. The central reason is not that we think that the benevolent dictator approach is useful or verisimilous. Instead, assuming a benevolent dictator allows us to finesse the problems of collective choice that plague any group endeavor, as those are not our main focus. To put it another way, even if the patently absurd assumption of a benevolent dictator is granted, the failures of government we catalog all still occur, and are in fact unavoidable.

1.2 Why the Pareto criterion is appropriate as a benchmark

The “Pareto criterion” allows the comparison of different social states of the world based on the evaluations of individuals. The difficulty in selecting “good” policies is both ethical and practical; the Pareto criterion gives primacy to individual judgments about what is better or worse, and asks if there are some moves on which we can all agree.

Pareto *improvements* are based on a comparison between two alternatives, the status quo Q and a feasible new social state S . We say S is “Pareto superior” to Q if *everyone* in the society prefers S to Q (strong form) or if *at least one person* prefers S to Q and no one prefers Q to S (weak form).

Pareto *optimality* is based on a comparison between a status quo Q and any other alternative S . We say Q is “Pareto optimal” if there exists no feasible S that is a Pareto improvement over Q .

Pareto comparisons have three key elements, and a proviso⁷:

1. *Autonomy*: The process defers to the individual's subjective judgments about better or worse. If the individual believes herself to be better off, she is, and in the absence of other compelling facts that belief is given deference.⁸ So no one has to answer “why” or give reasons. Instead, the utility function of the individual is the judge. In practical terms, autonomy requires *actual* consent to proposals that are believed to “improve” the welfare of the individual.
2. *Unanimity*: The Pareto criterion is a means of obtaining consent to public or collective decisions. The strong form requires that all people in the society agree that they are better off. The weak form requires only that no one is worse off, in line with Mill's “Harm Principle” (Mill 2002).
3. *Compensation*: Pareto comparisons suffer from the weakness that no one must be harmed. But many comparisons create total benefits far larger than the relatively small costs they create. This suggests augmenting Pareto comparisons with side payments or compensation. Importantly, this suggests a means of adjudication among Pareto optima, because some Pareto optima are “better” (create more total value) than others. That is, if the benefits to choosing S (for the gainers) are large enough to create a surplus value that allows the compensation of those who prefer Q (who would, in effect, “lose” by moving from Q to S), unanimity can still be achieved. The gainers compensate the losers, and the result is a Pareto improvement, because (after the

⁷ For technical conditions on the ranking of Pareto optima, see Das (1999) and Keeney and Kirkwood (1975).

⁸ The presumption in favor of autonomy is part of the standard “liberal” precepts. The question is what conditions must be met to rebut the presumption of autonomy. For an excellent discussion, see Gaus (2005).

compensation is paid) everyone prefers S to Q. This “compensation-augmented Pareto comparison” suggests the existence of at least one “Optimal Pareto Optimum” or OPO. The OPO is that Pareto optimum compared to which no Pareto improvement is possible, even allowing compensation (and assuming, because of the actions of the benevolent dictator, no hold-up or other strategic bargaining tactics).⁹

Proviso: There is an obvious, and potentially crippling, tension between autonomy and compensation.¹⁰ As we will see there are objections to the notion that someone else can accurately gauge what a citizen values and how that trades off for compensation. Further, there are bargaining problems because the citizen might be better off but hoping to “hold up” the process by refusing to give actual consent. Thus, there are at least three problems: (a) information; (b) incommensurability; and (c) hold-up. Each of these problems might legitimately prevent agreement. But we are only going to consider the information problem, and investigate the extent to which different mechanisms are effective “discovery processes.” That is, we will assume that all (relevant) issues are commensurable, making compensation possible, and we will assume that people are cooperative and willing to accept a proposed solution (if it is in fact better from their perspective) even if strategic hold-up might give them an even better outcome. The reason for this proviso is obvious, we hope: even if these (significant) simplifications are allowed, governments still fail. Without these simplifications, failure can be systemic and catastrophic.

1.3 Outline: substance, process, and behavior

We apply this apparatus to the problem of government failure in three steps. The first step is explained in Sect. 2: *substantive government failure*, or the failure to solve basic problems of organization and credible commitment. In these instances, which we examine in the first major section of the paper, we consider problems of organizing a functional system of governance, to make credible commitments to control debt or the temptation to debase the currency, and to control rent-seeking. All of these problems, if unsolved, leave the society at an outcome that is not Pareto optimal. In practical terms, this means that at least one feasible alternative exists that could make every member of the society better off, and yet the change is not made.¹¹

This is precisely the logic of the Hobbesian “war of all against all” (Hobbes 2012). In the absence of any order, no investment, property, or law is worthwhile. Everyone in the society would be better off agreeing to some system that offers governance, predictability, and freedom from coercion by outsiders or other citizens, so we look at the “absence of order.” If the state cannot impose predictable rule of law without corruption, then state activities are diverted from securing order and producing basic public goods to lining the

⁹ Clearly, the claim of the existence of the OPO rests on a (dubious) assumption that all policy moves are commensurable, and can be compensated. But as we discuss below it is far from clear that this is true. If we are considering whether to outlaw, allow, or subsidize abortion, it is unlikely that citizens who think abortion should be illegal would be satisfied with side payments, or have their mind changed by compensation.

¹⁰ This tension is captured in an old joke. A radical standing on a soapbox giving a speech shouts “Come the revolution, you’ll all have milk and honey to eat!” A bystander shouts back, “What if I don’t like milk and honey?” The radical glowers and yells, “Oh, friend, come the revolution, you will like milk and honey!”

¹¹ One could also make intergenerational comparisons as part of a Pareto comparison with compensation, but it is not clear how future generations could negotiate, much less effect, an optimal payment.

pockets of officials, so we look at “transaction costs” such as corruption, agency, and rent-seeking.

If the society secures a foundational and workable consensus on the substance of governance—a “constitution”—then the problem of escaping Hobbes’ “state of nature” is solved, and some Pareto optimum is selected. The failure to reach a Pareto optimum is therefore how we define *substantive government failure*.¹²

The problem of *procedure* is quite different. Procedure is the problem of choosing the “best” from *among* a set of Pareto optima. As Hardin (1990) argued, this step requires justifying the procedures to be used by an actual polity. To gain analytic purchase on this problem, we employ a simplifying procedure: imagine that only two Pareto optima are on the table, one of which is better than the other under the “potential Pareto with compensation” comparison. That is, allowing for cardinal utilities and side payments, one of the two Pareto optima being considered is better than the other.

We investigate whether any of five general categories of procedures—voting, interest group pluralism, pricing-exchange processes, bargaining, or cost-benefit analysis—can be expected consistently to select the correct answer.¹³ It is important to note that we do not consider this decision context to be realistic; quite the contrary. Our goal is to assume away the Arrovian problems of social choice (by having only two alternatives) and problems of ethics, ideology, and political intransigence (by having cardinal utilities, considering only compensable issues selected by purely rational agents, and allowing side payments). What we find is that even in these artificially favorable circumstances all five categories of procedures fail to select the OPO.

In Sect. 4, we consider some extensions and possibilities for the elaboration of government failures under more realistic assumptions about human behavior. These observations are necessarily more suggestive than dispositive, because work in psychology and behavioral economics has only begun to illuminate the problems of decisions and collective action and the limits of even the best procedures to produce the desired outcomes.

2 Substantive government failure

2.1 Justifying the state

It useful at the outset to summarize the justification for the existence and centrality of the state in society. Economists tend to think of markets as institutions for maximizing social welfare. But there is nothing “social” about maximization: Robinson Crusoe faced maximization problems when he was alone on his island. On Thursday (i.e., before Friday), Crusoe had to decide whether to build a shelter, find food, find water, or rest. The optimum

¹² It may be that we are under-counting the costs of coercion, either in transition or *per se*. The premise is that there are equilibria that are not Pareto optimal. And we posit that there are Pareto optima, possibly desirable on other grounds, that are not equilibria. The only way to keep the society “at” the preferred Pareto optimum is by force. In the case of a Prisoner’s Dilemma setting, of course, citizens might endorse being forced in this way, since the outcome is preferable to what they could accomplish on their own. But the presumption of an “optimal Pareto optimum” is that all societal choices have this property. Thanks to Scott Ainsworth for helping us clarify our thinking on this point.

¹³ For optimistic versions of this approach, see Fishkin (2011) and McKean (1965). Somewhat less optimistic is North (1994). For a general discussion of “non-market” failures of institutions see Wolf (1979, 1993).

choice among alternative activities meant the difference between life and death. Consider Crusoe's account:

I was not idle...I considered the keeping up a breed of tame creatures thus at my hand would be a living magazine of flesh, milk, butter, and cheese for me as long as I lived in the place, if it were to be 40 years... [I also grew grapes], which I never failed to preserve very carefully...they were not only agreeable, but medicinal, wholesome, nourishing, and refreshing to the last degree...

But now I come to a new scene of my life. It happened one day, about noon, going towards my boat, I was exceedingly surprised with the print of a man's naked foot on the shore, which was very plain to be seen on the sand. I stood like one thunderstruck... I came home to my fortification, not feeling, as we say, the ground I went on, but terrified to the last degree, looking behind me at every two or three steps, mistaking every bush and tree, and fancying every stump at a distance to be a man. (Defoe 1968, chap. XI)

When Crusoe later encountered Friday, their first problem was to arrive at a “bargain” wherein neither would kill the other. Markets are premised on a particular form of cooperation called *exchange* (which is a social act) rather than solely promoting *optimization* (which is an individual act, as made clear by Crusoe's choices of growing grapes and keeping goats). But for precisely that reason, markets require a *prior*—though perhaps unspoken—social agreement that we not kill each other, and (possibly) in addition that we will cooperate for mutual defense.¹⁴

The justification for cooperative agreement is thus prior to markets, at least in terms of the way we usually think of markets as institutions for reducing the transactions costs of impersonal exchange. Purely voluntary *bilateral* contracts among individuals cannot solve the problem of circumscribing coercion, and voluntary *multilateral* contracts are much too costly to negotiate and enforce, at least not over large areas and over long time spans. Consequently, *collective* institutions for controlling coercion are a necessary pre-condition for the very existence of the kind of enforceable *voluntary* contracts that markets simply assume. Buchanan's (1987) metaphor of “politics as exchange” suggests that groups can constitute agreements that are Pareto-superior cooperative solutions to the failure of decentralized bilateral transactions.

But this means that Hobbes was right in an important sense. Coercive institutional arrangements exist that are Pareto improvements over a ‘state of nature’ wherein no state institutions are in place. And, as Brennan and Buchanan (1985) note, these rules can become nearly self-enforcing, and therefore much cheaper to enforce, if they are followed and internalized. F. A. Hayek, no fan of the state, concedes as much when he says:

The question is ... what properties [social] rules must possess so that the separate actions of the individuals will produce an overall order. Some such rules all individuals of a society will obey because ...they will be part of their common cultural tradition. But there will be still others which they may have to be made to obey, *since although it would be in the interest of each to disregard them*, the overall order on which the success of their actions depends will arise only if these rules are generally followed. (Hayek 1973, vol. I, p. 45; emphasis added)

¹⁴ One could argue that this agreement need not be explicit, or even conscious, but might instead be the equilibrium result of an indefinitely repeated game, based on the “folk theorem.” See, for example, Myerson (1991). It's still a recognizable pattern of behavior where people who might kill each other do not, in recognition of the individual benefits of acting in this way.

But even if we accept the *existence* of at least one (unknown) Pareto-improving governance structure, all that would establish is that *some* form of state can be justified. *Which* form? And once the state exists, what should it do, and how can the state be induced to do that? What should it not do, and how can it be prevented from doing that? There is an obvious logical contradiction in claiming that since some (unspecified, possibly unique) governance structure can be justified that therefore *any* state is better than a state of nature. What is needed, and what we propose here, is a *guide* to choosing among feasible Pareto optima.

That choice will be implied by the capacity of varieties of actual institutions to choose what we will call the *Optimal* Pareto Optimum (or, again, the OPO). One virtue of this approach, in our view, is that even under these strikingly restrictive (and advantageous) assumptions about the decision context government still exists. But another is that it is closely related to the approach traditionally taken to defining “market failure.” The concept of market failure goes well back into the nineteenth century, with John Stuart Mill, Henry Sidgwick and Arthur C. Pigou (Medema 2009, chaps. 2 and 3). As Ledyard (2008) put it, “The best way to understand market failure is first to understand market success.” The problem is that “market success” cannot be defined without a set of implicitly assumed, usually unnoticed, assumptions about the existence of a functioning state.

There are exceptions, of course. A. C. Pigou is (rightly) seen as an ur-text for the argument that the state might be able to improve policy outcomes, given market failures. But he was far from naïve about the limits of government action (Pigou 1912, pp. 247–248):

It is not sufficient to contrast the imperfect adjustments of unfettered private enterprise with the best adjustments that economists in their studies can imagine. For we cannot expect that any State authority will attain, or even whole-heartedly seek, that ideal. Such authorities are liable alike to ignorance, to sectional pressure, and to personal corruption by private interest. A loud-voiced part of their constituents, if organized for votes, may easily outweigh the whole.

Later, in “State Action and Laissez-Faire,” Pigou (1935, p. 124) again sounded a note of caution¹⁵:

In order to decide whether or not State action is practically desirable, it is not enough to know that a form and degree of it can be conceived, which, if carried through effectively, would benefit the community. We have further to inquire how far, in the particular country in which we are interested and the particular time that concerns us, the government is qualified to select the right form and degree of State action and to carry it through effectively.

Nonetheless Pigou argues that the distinction between states and markets is fundamentally misleading, because the state comes first:

No defender of so-called laissez-faire desires that the state should do absolutely nothing in matters relevant to economic life. The most ardent believer in the economic harmonies, that are supposed to flow from the unimpeded pursuit by individuals of their private interests, argues that these harmonies will not emerge unless robbery at arms is restrained by law, fraud repressed, and contracts which have been formally accepted enforced. (Pigou 1935, p. 109).

In our “anatomy” of procedural government failure, we identify three problems that all governments, authoritarian and democratic systems alike, have to solve: the absence of

¹⁵ These references for Pigou (1935) were found by Medema (2010, pp. 49–50).

order, abuse of the purse, and the failure to control transaction costs. Those in political power may or may not care whether their societies suffer these problems. But states that exhibit these failures will be weakened, and are more vulnerable to revolution or external threat than they would be otherwise. More importantly, for our purposes, such failures make states “failures,” in the sense of failing to achieve the optimal Pareto optimum that justified their creation and continued existence.

2.2 The absence of order

The logic of Hobbes (1651) is that even a Leviathan state, an autocratic monarch with unlimited power, would be better than no government at all. In a similar vein, Huntington (1968, p. 1, emphasis added) has said that:

The most important political distinction among countries concerns not their form of government but their *degree* of government. The differences between democracy and dictatorship are less than the differences between those countries whose politics embodies consensus, community, legitimacy, organization, effectiveness, stability, and those countries whose politics is deficient in these qualities

Huntington included the USSR with the United States in the former group, countries that are extensively governed by a kind of consensus. But he was wrong: The collapse of the Soviet Union suggests that repression masked severe deficiencies in “consensus, community, legitimacy, organization, effectiveness and stability.” The United States itself may not have the “degree” of government it seemed to have in 1968.

Nonetheless, we endorse Huntington’s point: the absence or extreme weakness of government is a failure, a failure that has profound implications for the way people are able to live their lives. It may be possible, in principle and some cases in historical fact, for an anarchist group to govern itself through consensus on the obligations and limits of citizenship, to organize itself, and to maintain stability (Taylor 1982 gives examples). Such situations are rare in the contemporary world, and—for whatever reason—have not survived long when they do exist.¹⁶ Thus, when we say “governance” we will be considering government by state institutions, and when we say “government failure” we will mean the failure of state institutions.

It is common to argue whether the state is “too strong” or “too weak.” But this is a misleading dichotomy. The citizens have selected certain competencies or jurisdictions for the state to exercise power, usually a monopoly on the legitimate use of force. In that context, the state must be strong enough to carry out its functions within those competencies, but much too weak—because of institutional restraints—for its powers to metastasize outside the jurisdiction legitimated by citizen consent (in a democracy) or established custom and convention (in a stable autocracy). A state that lacks either institutional powers or practical capacity within its core jurisdictions suffers the most fundamental government failure: the absence of order. These are failed states.¹⁷

¹⁶ See, for background and examples, Friedman (1989), Hayek (1948), Narveson and Sanders (1996), Rothbard (1978), and Scott (2009).

¹⁷ See <http://ffp.statesindex.org/indicators>. See Ezrow and Frantz (2013, chap. 1) for a discussion of the concept of state failure, and a review of measures and their limitations. The Fund for Peace maintains an index of failed states at <http://www.failedstatesindex.org>, which is based on twelve indicators, four social (demographic pressures, refugees, group grievance, and human flight and brain drain), two economic (uneven economic development, and poverty and economic decline), and six political and military (state legitimacy, public services, human rights and rule of law, security agreements, and external intervention).

In their book on *Failed States and Institutional Decay* (2013), Ezrow and Frantz advocate a focus on the quality of state institutions, which they classify as administrative, judicial, security, and political. Good institutions promote political stability and economic performance. For “high quality” administrative institutions they advocate meritocratic recruitment and promotion, and autonomy (among other things). For high quality judicial institutions, they advocate independence, integrity, and transparency. For high quality security institutions, they advocate professionalism and a centralized military command with professional norms of deference to civilian authority. Ezrow and Frantz (2013, chaps. 4–7 and Conclusion) do not advocate any features as characteristic of high quality political institutions, because they suggest that no features of such institutions uniformly promote political stability and good economic performance.

An element of conventional wisdom in political economy is that limitations on government are a fundamental requirement for economic growth. Still, as Huntington (1968, p. 8) put it, “authority must exist before it can be limited.” Hanson (2014) distinguishes investment-led growth, in which a strong state may be necessary for building an industrial base, from innovation-led growth, where institutional veto players may be more constructive. Hanson argues that at early stages of development, state authority is more important than constraints on rulers for promoting total factor productivity and growth than it is when states are closer to the technological frontier, when constraints become more important.

The absence of a government with the capacity to provide for the security and well-being of its citizens is the most basic government failure. And it is, to use the terminology of Weimer and Vining, a passive failure, but a passive failure so catastrophic as to make the state unsustainable as a distinct political and geographic entity.¹⁸ As Weimer and Vining (2011, p. 206) put it:

[Passive failure] can include outcomes that are attributable to the government not diagnosing market failures correctly as well as situations in which the lack of intervention derives from more concrete causes, such as the active influence of organized interest groups that successfully block efforts to correct market failure.

Some such failures to act may be sustainable, at least in the short run. For example, an effort at pollution control that is blocked by supporters of the emitting industry simply imposes net economic costs that could have been avoided by government action. But passive government failure can extend to a breakdown in the collection of taxes, the failure to pay for and organize police and military protection, and an inability to maintain basic infrastructure. Failed states have economic costs.

A number of theorists recently have addressed the problem of government failure at the level of systemic collapse. Diamond (2005) argues that what he calls “collapse” hinges on institutions that fail to solve collective actions, and may even create negative feedback loops. More recently, Acemoglu and Robinson (2012) argue that a strong central

¹⁸ David Weimer and Aidan Vining have kept the existing theory of government failure alive through five editions of their successful textbook, with at least one chapter on government failure in each of the five editions from 1989 through 2011. They begin with the (perfectly standard) notion of “a perfectly competitive economy” (Weimer and Vining 2011, p. 54) as the benchmark, and then gauge government failures in terms of deviation from CET, a conception we have argued is too narrow. Nonetheless, it is important to recognize that Weimer and Vining’s chapters on government failure are central to the way that academic policy analysts and economists have conceived the issue. For a broader review of academic treatments of government failure to date, see Eyzaguirre et al. (2014). An interesting view of the asymmetry of treatment of market versus government failure is Ferrarini et al. (2011).

government is a prerequisite for growth and prosperity. They put a nearly exclusive emphasis on institutions, with a focus on institutions that reward innovation and produce shared prosperity. They are skeptical of the “geographic determinism” explanations given by Diamond, and argue that the most important feature of institutions is whether they create sustainable prosperity and a shared sense of cultural success over long periods.

Scott (1998, 2009) takes a more pessimistic view. He believes that two tendencies are evident in the pattern of state development over time. The first is the consolidation of power in the central government. The second is the general weakening and decline of highly centralized states, echoing in some ways Olson (1982). He claims that “high modernist ideology” focuses on centralizing state power at the expense of the factors that made societies grow and prosper in the first place. Scott (1998, pp. 88–89) argues that, to the extent that these two factors are observed in all governments over time, there is a focus on rules, procedures, and bureaucracy that suppresses what he calls “the indispensable role of practical knowledge, informal processes, and improvisation in the face of unpredictability.”

2.3 Abuse of the purse

Governments have spending obligations that can range from small to large, depending on the values and beliefs of those who control the government. Governments can cover the spending with taxes, borrowing, or monetization. Having a monopoly of legitimate force, governments can tax their citizens, and they can borrow from them as well as from other potential creditors outside their boundaries. Prudence and rule of law substantially reduce the interest rate required for borrowing, and the ability to raise taxes through voluntary compliance likewise makes it possible for states to borrow now against that implied revenue source in the future.¹⁹

Governments may or may not follow a “fiscal constitution” defining conditions under which it is appropriate to borrow rather than tax.²⁰ The risks of too much borrowing include “debt intolerance,” which “is a syndrome in which weak institutional structures and a problematic political system make external borrowing a tempting device for governments to employ to avoid hard decisions about spending and taxing” (Reinhart and Rogoff 2009, p. 21). One possible result from this syndrome is default, and an inability to borrow in the future.

Of course, this discussion has presumed—implausibly—that the value of money remained constant over time. But the capacity of nations to depreciate their currencies is an additional danger. In fact, debasement of the currency is simply another means of indulging fiscal indiscipline. A government that inflates its currency can “repay” its debts with less actual value than what it borrowed originally. This was possible under metallic standards—for example, a gold standard or silver standard—by reducing the amount of precious metal in the currency (Reinhart and Rogoff 2009). Bernholz (2003) shows that in general, inflation has been lowest in metallic regimes, and highest in fiat money regimes. Selgin and White (1987) give an even less optimistic view of the effects of fiat money.

But this historical fact—more fiat money, more inflation—confuses cause with effect. A scrupulously maintained gold standard indicates a commitment to sound money, not vice

¹⁹ North and Weingast (1989) argue that the fact that the English king was bound by legal obligations actually increased the capacity of the state to borrow, whereas the relatively greater power of the French monarch, unbound by such restrictions, prevented the French from having access to capital by selling bonds.

²⁰ See White (2014) for a description of the establishment and demise of an informal fiscal constitution in the United States. More generally, see Buchanan and Wagner (1977) and Wagner (1983, 1989) for generic difficulties with using transfers and debt as policy tools.

versa. Exactly the same result could be obtained with fiat money system under a tightly enforced monetary constitution, if a monetary constitution could be tightly enforced. And lacking a commitment to a monetary constitution, a gold standard is no protection, because the value of gold, or—what amounts to the same thing—the amount of gold contained in or backing the currency, could be adulterated or manipulated. The key is not the composition of the currency, but the commitment of the government authority to sound money.

Failed states are unable to solve the time inconsistency problem, and either borrow excessively or monetize debt excessively in the short run in a way that ultimately denies access to loans at risk free interest rates. Abuse of the purse is more than a technical failing, implying a failure of the state to be able to finance basic services such as defense and enforcement of property rights.

2.4 Transaction costs

There are many generic failures that can happen in either democratic or autocratic governments at the level of implementation and enforcement. Corruption, agency loss, and rent seeking are problems that involve transaction costs, extortion, or strategic contracting problems in one way or another. This problem highlights again our claim that governments are logically prior to markets, in the sense that institutions create the set of rights and the transaction costs involved in exchanging rights and the predictability of rewards that determine the vitality of a political economy.

The simplest of these failures is *corruption*, which, in a governmental context, is the private use of public resources by the government officials in control of those resources. Corruption can range from direct embezzlement of government funds to the solicitation of bribes in return for government contracts or public policies. Such corruption usually is unambiguously bad (because corruption is Pareto inferior to rule of law), and it is almost always—by definition—technically illegal, though it may be winked at or even encouraged by the authorities.²¹

Agency loss occurs in a governmental context when an agent substitutes his or her goals for those of the principal. Agency loss can range from simple shirking of duties to substitution of the preferences of the agent for the goals of the principal. An example of this is government regulation of industry, when the regulators become more sympathetic with the regulated industries than with the proponents of regulation (Stigler 1971). But agency loss can also result from an unwillingness to devote effort or impose needed reforms, as with the problems the United States has recently had with the Veteran's Administration hospital system, or with the failure of the Department of Health and Human Services to prepare adequately for the web interface problems in signing up for Obamacare. These problems may be a lack of capacity in some nations, but in the U.S. case, it appears that the failure primarily was a lack of implementation and political will. Benefits of progress that seem substantial in principle can be entirely consumed by agency costs.

The baseline for *rent seeking* is more subtle than that for corruption or agency loss, but at root it involves competing for a resource that already exists. Rents are defined as “rewards and prizes not earned or not consistent with competitive market returns” (Congleton et al. 2008, p. 1). As Buchanan (1983) points out, the more government intrudes into directing resources in society rather than protecting individual rights and enforcing contracts, the more opportunities there will be for rent seeking. The point is that

²¹ If the legal regime involves rent-seeking arrangements, a bribe to cut through red tape may be defensible on utilitarian grounds. See Wedeman (2012, chap. 2).

efforts to secure benefits that are defined by government in excess of the competitive market returns implied by the opportunity costs of the resources in question will fail to produce benefits, and can impose substantial net social costs.

2.5 Substantive government failure

The theme of the “failures” outlined in this section is the inability of actual states under some circumstances to achieve Pareto optimal solutions to problems of cooperation. The three large scale failures described here, the failure to maintain basic order, failure to manage the purse, and failure to manage institutions that reduce transaction costs of exchange, all prevent citizens from achieving feasible cooperative results.

To the extent that the facilitation of capturing the gains from cooperation is, in our account, the justification for the creation and existence of the state, the failure to achieve a Pareto optimum is the most fundamental kind of failure. In a sense, the difficulties we have discussed here are meta-problems, failures so fundamental that the very justification of political authority is impeached. In the next section, we take up the problem of procedure, assuming that the problems discussed here have been solved, or at least are controlled. Procedures are means for choosing among Pareto optima, when citizens disagree about which outcome is preferable. Choosing procedures is one of the fundamental problems of politics.

3 Procedural government failures

3.1 Theoretical problems of governance: five mechanisms for choosing and implementing policy in real states

Governments make choices for society using some collective choice rule. Democracies use some form of voting procedure, and authoritarian governments, even the narrowest dictatorships, are likely to use some system of weighting the welfare, productivity, and value of different segments of the population.²² Then it is the nature of government choice that

²² One could object, as have James Buchanan and others, that this statement of the problem is simply incoherent. As Buchanan put it, “There are obviously many reasons why the taxpayer may not evaluate alternative tax instruments in the same way that the applied welfare economist evaluates them” Buchanan (1979, p. 54). That would apply equally well to technocrats or politically appointed bureaucrats, of course. Another view, an extension, really, is Wagner (1985, pp. 386, 388):

[It is a] curious piece of vocabulary that affixes the designation “individualistic” to an analytic construction in which people are manipulated as objects at the disposal of some type of despot, who is presumed to be benevolent by virtue of being named “Social Welfare Function.”

[P]olicy outcomes are assessed against some transcendent criterion of goodness, independent of any consideration of what the participants might or might not work out among themselves.

The point is that since the values we place on things are subjective, no justification is available to expect that the way a benevolent dictator, or a public finance analyst working for the Army Corps of Engineers, will value either the costs or benefits of a public action the same way that any given individual will. This seems like an epistemological problem, perhaps an aspect of the “knowledge problem,” but the real objection is that it is literally impossible to “know” what people want at a high level of aggregation. This is the problem that the various “procedures” we outline must solve, by collecting information from individuals. If the dictator were not benevolent, of course, he could ignore what the people want. But we are assuming that the dictator is benevolent, *but not omniscient*. Thus, he has no way of figuring out how individuals (as opposed to a central technocrat) actually value costs and benefits. Similar discussion, though with some important differences, can be found in Dolfma (2013), Shepsle and Weingast (1984), and Simmons (2011). We thank Georg Vanberg for helping us work out ways to try to address this difficulty.

the single policy or rule selected is binding on all: one defense budget, one national flag, and so on.

Exceptions arise if rules or policies are segmented. In the case of federal systems, disjoint subsidiary units may have different rules; less benignly, government may discriminate among races, castes, or regions, selecting some subset of citizens to tax, or oppress, compared to others whom the rules privilege.

But these exceptions are complications of the general observation that governments must choose policies or budgets and that then this budget applies to the entire geographic area relevant to that choice. Discrimination based on morally arbitrary or purely invidious difference might be categorized as a separate kind of “government failure” on its own, but we will not discuss it further here. Our approach has been to place government institutions in the most favorable light possible, and then to make comparisons based on the Pareto criterion, assuming the “dictator” is not racist or evil.

So, ruling out purely harmful or discriminatory government actions, we describe five mechanisms or algorithms for making collective decisions. And, again, even authoritarian states may use at least one of these mechanisms, unless the state is ruled purely by the random whim of the autocrat. Still, rule by autocrats is likely to be idiosyncratic, almost by definition.

3.1.1 *Democracy*

Democracy, in our view, has two essential—and possibly conflicting—features: popular rule based on voting to elicit the “will of the people,” and limitations on governments’ ability to control or coerce individuals, even if that is the will of the people as revealed by the voting rule. In this sense, democracy is any of a broad class of institutions that allow groups of people to achieve a balance, avoiding both the absence of order and the government failures of excessive or coercive action listed in the previous section. Elections, for example, allow a group to choose a leader to achieve the gains of cooperation in providing public goods (Samuelson 1954; Strotz 1958; North et al. 2012), while also affording an opportunity to get rid of bad leaders, without violence or revolution. Constitutional limitations on government block bad policies or the metastasis of coercive power into aspects of private lives to which citizens have not given their consent. Obviously, neither elections nor constitutions are foolproof.

The first two mechanisms or procedures we consider—voting and interest group pluralism—are standard political decision mechanisms of democracy. We consider these to be “centralized” mechanisms, because even though the unit of the choice—the individual voter or interest group—is capable of independent judgment and registers an independent preference, the result is a single choice for the entire group. In some cases, such as choosing a defense budget or a speed limit, this single choice for all citizens may make sense. In other cases, such as restaurant menus, reading or television viewing choices, or the decoration and layout of homes, decentralized and individual choices may be preferable.

3.1.2 *Generic mechanisms*

The next two procedures (decentralized price-signaling and bargaining) are generic ways in which a democratic or authoritarian society can make decisions that approximate Pareto improvement on a disaggregated basis, rather than making top-down decisions for the whole society. An autocratic government could choose—as appears to be the case now in Singapore, or for decades in Hong Kong, before the changeover, for example—to use

market processes as the result of a conscious choice to protect property rights and control corruption. Alternatively, a system of “bargaining” among elites, either as members of a “Politburo” or aristocracy or among powerful warlords or barons, might either be chosen or simply emerge as a conventional means of making allocation decisions. One might argue that markets and bargaining “just happen,” of course. But the efficiency and stability of such a system depends heavily on the ability of the government to make credible commitments to respect the results and protect property rights.

We will also categorize as bargaining such small group reflective decision mechanisms as “deliberative democracy.” The style of bargaining differs, of course, in that participants are trying to persuade others to take account of other principles or ethical considerations, but the essence of bargaining is that participants change or at least adjust their positions in response to “offers” by others.²³

3.1.3 *Technocracy*

The final mechanism, cost-benefit analysis, is a technocratic solution that cuts across economics and politics. The premise of cost-benefit as an approach to allocation decisions is to add up the total benefits to the gainers, and then subtract the total costs to the losers. Cost-benefit may well be observed in any bureaucratic setting, in either democracies or autocracies, privileging the (possibly flawed) estimates by experts over the (possibly flawed) judgments of individuals in decentralized settings or preference aggregation mechanisms such as voting.

3.2 The nature of the comparison

3.2.1 *Centralized, decentralized, and hybrid mechanisms*

The comparison is conducted under artificially favorable conditions, as it is useful to ignore the classical problems of collective choice that arise with ordinal, subjective utilities and multiple alternatives.²⁴ We limit the alternatives to two Pareto optimal choices, and assume cardinal, interpersonally comparable, utilities, not because these assumptions are plausible but because these assumptions allow us to focus on the intrinsic flaws in choice mechanisms even under artificially favorable conditions. More simply, what we illustrate is that government fails regardless of which of the five choice mechanisms is selected, even if the social choice problem is assumed away.

To clarify the nature of the comparison we wish to make among the five alternative mechanisms for choosing or organizing cooperative human activity, we present a Venn diagram in Fig. 1.

We stipulate, as a starting point, from the previous section, the desirability of achieving Pareto optima. After all, if a social choice process selects an outcome that is *not* Pareto optimal then some or all members of the society can be made better off without harming anyone else. The real problem, which economists have tried to avoid, is the selection *among*

²³ The authors acknowledge the assistance of Thomas Schwartz for discussions clarifying our thinking on this point.

²⁴ This is true not least because the usual criticism of social choice mechanisms, the “Arrow Theorem” and its variants, was originally directed at the conclusions reached about welfare results in *market* settings. Thus, there is little use in highlighting the inability of voting mechanisms to solve the social choice problem, since markets can’t do it either. Arrow (1963), see Mueller (2003) for a review.

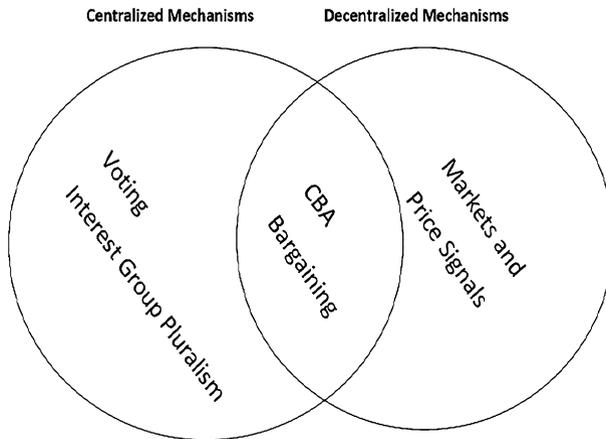


Fig. 1 Centralized and decentralized mechanisms for organizing groups of humans

Pareto optima when people disagree about the “best” outcome. But this is a problem which has a solution, if one accepts the utilitarian calculus of adding up gains to the gainers and losses to the losers.²⁵

Public or collective choice is characterized by processes that admit of only one choice from among many (such as a speed limit, a defense budget, a policy on same-sex marriage) which then must be obeyed by all in a given jurisdiction. Consequently, almost all choices involve (depending on the status quo or reversion point) an improvement for some people and harm for others. The Pareto criterion, as such, has nothing to say about such problems. But society has to choose *something*, and some ways of choosing are better than others, in particular circumstances.

In asking the question this way, we are connecting to a pre-Arrovian tradition in welfare economics. In some ways, the Arrow Impossibility result, and subsequent work in social choice, has been a diversion. The older debate in social science was over the relative performance of different collective choice mechanisms, given a benchmark for comparison. The best known work in this area was that by Bergson (1938) and Samuelson (1947).²⁶ The “market failure” (Bator 1958) view was on a parallel track, premised on a comparison with an *optimal* Pareto optimum. When market processes fell short of this, state regulatory action was required.

A second set of results, often simply ignored by “market failure” view, was the earlier “socialist calculation” debate of the 1930s and 1940s. Oskar Lange and Abba Lerner (among others, Lange and Taylor 1938; Lerner 1944) argued that it was possible to “solve” the problem of allocating resources so as to achieve a Pareto optimum. Further, since the government could also redistribute benefits, it was possible to achieve the optimal Pareto

²⁵ It is important to note, for now and later in the paper, that we cannot really mean fully utilitarian aggregation. Aside from the measurement problems, which would disqualify the utilitarian calculus on their own, there is also the problem of “standing,” or deciding which preferences (in effects, which pleasures and which pains to *count*). A person might honestly enjoy cannibalism, and want to pay a suicidal person (almost indifferent between life and death, and so willing to “sell” at a low price) to be eaten. We will simply rule out, by prior side constraint, all those Pareto improvements that are not allowed by the standards of morality and human rights in the society. Thanks to Keith Dougherty for pointing out the importance of restricting the extent of the application of the utilitarian claim.

²⁶ For a review of the history, and intellectual development, of this concept see Mueller (2003, chap. 19).

optimum. Ludwig von Mises and F. A. Hayek (again, among others; von Mises 1920; Hayek 1945), argued that the problem was not just difficult, but actually unsolvable, without unrestricted market processes to give accurate and dynamically updated price signals.²⁷ These “Austrian” objections are shunted aside in an approach that focuses on “market failure,” but come back into focus when government failure is the object of analysis.

3.2.2 Cardinal utilities, compensable choices

The approach we propose is restricted to policies that are compensable, or for which trade-offs exist for the citizen. This is not always true, of course. If a person thinks abortion is morally wrong, it is unlikely she will accept side payments or compensation to drop this belief. This sort of “all or nothing” issue resists compromise, and is singularly ill-suited to treatment by the usual tools of economists.

It is possible to consider even uncompensable policies in consequentialist terms, of course. Donohue and Levitt (2001) argued that a substantial material benefit flowed from legalizing abortion, because crime rates fell quite sharply after *Roe v. Wade* in 1973. More importantly, crime rates fell most, cross-sectionally, in states that had the highest rates of abortion.

This “result,” *even if it is completely correct*, is simply irrelevant to the concerns of abortion opponents. The objection to abortion is not that it has bad consequences for the society. The objection is that abortion has bad consequences for the fetus, and for the moral character of the mother and the health care workers committing the—as opponents see it—murder.

What activities might be compensable? We are thinking of mostly economic decisions: given that a road is going to be built, should we use route A or route B? In each case, eminent domain will be used to take property for a public purpose, and the owners will receive compensation, though less than they might demand to sell their property on the private market.

Alternatively, we might be thinking of basic regulations. Suppose the group is faced with a choice between two speed limits, again A and B. In this case, no compensation for the “losers” is usually offered. But it is possible to imagine achieving unanimous consent by *offering* compensation, provided the gains to the gainers exceed the losses to the losers in material terms. The problem is made more difficult by the necessity of specifying a *status quo* (the present policy, which if unchanged is preserved) or a *reversion point* (the

²⁷ The details of that debate are complex, but the core question was whether the state was capable of achieving the goal of selecting the optimal Pareto optimum. The difficulty faced by the state was aptly stated by—of all people—Leon Trotsky:

If a universal mind existed, of the kind that projected itself into the scientific fancy of Laplace—a mind that could register simultaneously all the processes of nature and society, that could measure the dynamics of their motion, that could forecast the results of their inter-reactions—such a mind, of course, could a priori draw up a faultless and exhaustive economic plan, beginning with the number of acres of wheat down to the last button for a vest. The bureaucracy often imagines that just such a mind is at its disposal; that is why it so easily frees itself from the control of the market and of Soviet democracy. But, in reality, the bureaucracy errs frightfully in its estimate of its spiritual resources...The innumerable living participants in the economy, state and private, collective and individual, must serve notice of their needs and of their relative strength not only through the statistical determinations of plan commissions but by the direct pressure of supply and demand. (Trotsky 1933).

Caplan (2004) asks whether the “knowledge problem” is really the unique, or even central, problem with socialist or other command economies. He claims—plausibly, in our view—that incentive problems are likely to be at least co-equal in hindering economic development. Zeckhauser (1973) raises the question of whether voting systems can solve the problem reliably, even with “honest” preferences.

alternative policy that will be selected if the groups fails to come to an agreement (Romer and Rosenthal 1978, 1979).

Economists generally have taken one of two approaches to solving the problem. First, they have assumed that the status quo is not a Pareto optimum, and then discussed means of moving to “the” Pareto optimum.” In the case of speed limits, we might start with “Zero.” That is, no cars are allowed to move. This is terribly inefficient. Then it is easy to imagine that someone proposes “Thirty.” Everyone prefers 30 to 0, and a Pareto improvement is achieved. It is quite likely that some choosers want a higher speed limit, but others may prefer a lower speed limit. Therefore, no other Pareto improvements are possible, compared to the new status quo point of 30 mph, and the choice is settled. But the same could be said for moving from zero to 25, or from zero to 40. Almost any collective decision is better than a bad reversion point. But that doesn’t mean that any collective decision is an OPO.

The second approach is to specify something like a *social welfare function* (SWF). This approach—once discredited after the results published by Arrow (1963) and others—is making a strong comeback. It is simply assumed that there is some coherent and mathematically tractable means of comparing alternatives on purely utilitarian grounds. Besley (2006), for example, cites Sen (1977) as a justification for assuming that a SWF exists and is known by the state and will be used by the state to make choices. All that is necessary is utilitarianism: “If interpersonal comparisons are allowed, social welfare functions can be derived which satisfy Arrow’s axioms” (Besley 2006, p. 21).

The claim is true, but the assumption of utilitarianism is not innocuous. What is meant is that if two alternatives A and B are available, then “we” know that A is better even though some people prefer B. Further, once “we” have the SWF, this knowledge is *ex ante* and objective; no additional information is required. Given two alternatives, one can be selected.

3.3 An example: choosing among Pareto optima

To make the problem clear, we consider a “toy” example, a group of people choosing between two alternatives. We will consider a setting wherein the starting/reversion point is not Pareto optimal, and the group must choose between two Pareto optima, each better (by definition) than the starting point. Consider Fig. 2, which depicts the situation, with the amount of Good 1 produced on the horizontal axis, the amount produced of Good 2 on the vertical axis, and the “Production Possibilities Frontier” marked in the heavy *dashed* line.

This familiar diagram “shows” that two alternatives, points A and B, are clearly better, for everyone, than Start. A third alternative, C, might be better than Start, but it would depend on how people trade off Good 1 against Good 2. Some people might prefer the extra Good 1, and would be willing to give up the amount of Good 2 required to move from A, or B, to C. But others might not. The SWF is used to choose for the society, because it can make interpersonal utility comparisons.

The SWF (actually the relevant indifference curve of the SWF) in Fig. 2 is the heavy *dotted* line. The SWF—because it can make interpersonal utility comparison—summarizes all the individual tradeoffs faced by each member of the society, adds up all goods and the utilities, and reveals that alternative B is the best Pareto optimum. A is Pareto superior to Start, and B is Pareto superior to Start, but B is socially preferred to A, because the SWF says so.

The problem is that the SWF is oracular; it is not obvious how real institutions might generate values for this sort of interpersonal welfare comparison. To make this clearer, let’s assume that the group has somehow already selected outcome “A,” which is on the production possibilities frontier for the society and is therefore Pareto optimal. Of the five citizens, some number prefer a move to “B,” and the rest prefer to remain at “A” (no one is

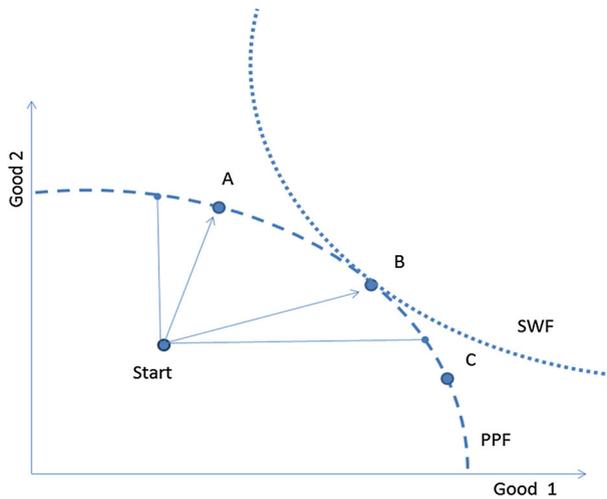


Fig. 2 The analytical problem of choosing the optimal Pareto optimum

indifferent). What should they do? If they actually had a SWF, the answer would be “move to B.” But how might the information that could be used to reach such a conclusion be gathered?

The answer in the social choice literature is to assume (a) that “we” know all preferences, of all individuals, and (b) that—as Sen and Besley suggest—interpersonal utility comparisons are allowed. We do something like this all the time, in effect, by using public policies to answer “needs” using money taken in taxes from others. Or when we build roads or dams, or when we impose regulations requiring motorcycle riders to wear helmets. In every case, we are weighing the harms, and benefits to different people, and trying to choose the policy that maximizes the sum of the benefits after subtracting the sum of the harms.

The problem is that the first assumption—perfect information about preferences and the consequences of policies—is absurd. Rather than assuming a SWF as an input to the policy process, real policy processes are imperfect attempts to elicit the information that would make approximating a SWF possible, as an output. Information is the *result* of various (imperfect) discovery procedures, not something we can assume policymakers have at the outset.

This suggests a highly simplified SWF, one that uses a cardinal measure of interpersonally comparable utility. Economists use a variety of names for this concept: willingness to pay, consumer surplus, producer surplus, and so on. Consequently, we will define the units of satisfaction as “dollar-utiles,” based on the willingness to pay of the participants. For any two points A and B, an individual either prefers A, prefers B, or likes them equally well. And if A is preferred to B, then “willingness to pay” tells us by exactly how many dollars A is better, for that individual.²⁸

²⁸ A difficulty in this approach is that “willingness to pay” might depend on *ability* to pay, as is the case for Coasian bargaining in the following section. A very poor person can’t “want” anything in a market setting, simply because she lacks money. At the other end of the scale, a very rich person “wants” everything. Our answer is that the approach requires assuming that all citizens have more or less the same wealth, since this translates to more or less equal weight in the SWF, regardless of their actual wealth position. Alternatively, if utility exhibits sharply diminishing returns in income differences in wealth would cause effects that would distort our conclusions. We rule out both of these considerations by assumption, but only to make the problem easier. Allowing for problems in payment and diminishing returns is akin to social choice problems in that the society may not be able to make a determinate choice in the first place. We are grateful to William Shughart for conversations on this point.

3.4 Five alternative mechanisms

3.4.1 Overview

Barring the existence of a “utilitron,” a Benthamite calculating machine (Bentham 1907) capable of adding up pain and pleasure, we are restricted to the five decision mechanisms already described as a means to discover the SWF:

- (a) elections or democratic voting,
- (b) interest group pluralism,
- (c) decentralized market processes, using prices,
- (d) bargaining (both Coasian and deliberative),
- (e) cost benefit analysis.

These five apparently “different” mechanisms differ only in how they count or weigh preferences and gather information about utilities. Returning to the social choice problem with just two alternatives, we ask: do the gains from moving to B benefit those who prefer B more than the losses of moving to B harm those who prefer to remain at A?

To extend the example, suppose that the group has five members, named Person 1 through Person 5. And suppose (again) that for some reason we have exact information on how each person values alternatives A and B, summarized in Table 1. Our example, then, starts with the answer, as if this were a Monte Carlo experiment in statistics: the group “should” move to B, assuming a utilitarian ethic and no other considerations, and requiring that all policy tradeoffs be compensable. The “units” in Table 1 are arbitrary, but cardinal, “utils.” Later we will treat them as if they were dollars, for ease of making comparisons and calculating side payments. But it is important to remember that these are actual utilities, and do not depend on willingness to pay based on wealth.²⁹

Why does this choice process not run afoul of the Arrow problem, or the Condorcet problem? The answer is that we are assuming just two alternatives, cardinal utilities, full information, and side payments. Interpersonal utility comparisons are therefore possible, and easily implemented through bargains or calculations of total willingness to pay.

The same determinacy of the “best” choice would actually be true with arbitrarily many choosers. The utilitarian optimum is the alternative that produces net positive dollar-utils, summed across all choosers in the society. We will call this the *optimal Pareto optimum* (OPO), and it is the dollar-utile maximizing Pareto optimum selected by the SWF. This is, to be sure, a highly restrictive assumption. But the assumption is useful because it reduces the complexity of the problem to one of implementation. It turns out that even if one can assume the Arrow problem away, and even if utilities are cardinal and interpersonally comparable, no mechanism can consistently and correctly select the OPO. Our claim in no way impeaches the importance of the Arrow (1963) result. Rather, this difficulty is stacked *on top* of Arrow as a barrier to the aspirations for “perfect” social choice. Even if utilities were cardinal and interpersonally comparable, the problem is intractable.

In the next section, we consider how each of the five decision mechanisms might fare in reaching the “correct” conclusion, with our quote marks around the word “correct”

²⁹ Otherwise, the difficulty is that the person with the lowest “willingness to pay,” person 1, may care a lot, but has a total wealth of only \$25. The person who seems to have the largest utility, in terms of willingness to pay, is person 5. But that might be Bill Gates, who doesn’t really care at all about the choice between A and B, but never carries anything smaller than \$100 bills, and hates to make change. Thanks to Sean Ingham for pointing out the importance of the distinction.

Table 1 Choosing alternatives A or B

	Will pay to stay at A	Must be paid to move to B
Person 1: prefers A	100	–100
Person 2: prefers A	50	–50
Person 3: prefers A	25	–25
Person 4: prefers B	–250	250
Person 5: prefers B	–300	300
Total utilitarian “value”	–375 (inferior)	+375 (superior)

serving as a warning that the problem is perhaps not as easily solved as we have pretended so far. But before we begin, it is important to note again that we have imposed enormous simplifications on the problem, eliminating many of the difficulties that usually confuse such proceedings. There is no proposal power, there are no additional alternatives to be proposed, and Arrow’s “Impossibility” theorem does not apply, since there are only two alternatives and utilities are cardinal.

3.4.2 Elections or democratic voting

In this subsection, and in the other subsections of Sect. 3, we refer to Table 1, with its five choosers (Persons 1–5) and two Pareto optimal alternatives (A and B). The example helps clarify the comparison, because all the social choice mechanisms are being evaluated in the same decision context.

Since only two alternatives are at hand, all *voting* systems (Borda, Condorcet, Approval, and so on) reduce to simple majority rule. Assuming that the only considerations are those listed in the utilitarian calculus in Table 1, then Persons 1, 2, and 3 vote for A, and Persons 4 and 5 vote for B.

One of the most famous arguments in favor of majority rule is May’s Theorem (May 1952; Mueller 2003). May showed that majority rule is the only decision rule that satisfies four (arguably) desirable conditions:

- *Decisiveness*: the implied SWF is single-valued for any given set of preferences, returning a result of A wins, B wins, or they tie.
- *Anonymity*: the identities of the voters can be hidden, and is in fact irrelevant to the outcome. Any permutation of the ballots, or random restacking, results in the same outcome.
- *Neutrality*: If one alternative W defeats (respectively, ties, loses to) another alternative V, and all voters have the same relative rankings of A and B that they have for W and V, then A defeats (ties, loses to) B.
- *Positive Responsiveness*: If the outcome is a tie, or is decided by a single vote, and one voter changes her mind, the outcome changes. Thus individual votes “matter” in close races.

May’s Theorem is an important characterization of majority rule, but one of the positive “features” turns out to be a bug. The fact that majority is neutral and anonymous means that information about intensity of preference is blocked; all that choosers can record is the direction of their preference, in this case “stay at A” or “move to B.” Consequently,

majority rule ignores valuable information, information that would be crucial to solving the utilitarian problem of choosing among Pareto optima.³⁰ To illustrate the nature of the problem, suppose that a slight majority has a weak preference to enslave a minority (alternative A). The minority has an intense preference not to be enslaved (alternative B). Popular elections are a poor method to handle such issues, precisely because everyone’s opinion counts the same, regardless of intensity.

3.4.3 Interest group pluralism

The interest group pluralism approach, as articulated in Bentley (1908) and Truman (1951), held out a promise of dealing with the intensity problem. These classical theorists actually had a very specific goal in mind: Describe the actual (as opposed to theoretical) process of decision-making in democracies, and evaluate its qualities compared to those claimed for abstract majority rule.

Bentley (1908) argued that a group is simply a distinct portion of society; any one person may be a “member” of many different groups. Further, as he famously claims in Chapter 7:

There is no group without its interest...The group and the interest are not separate...There is no political process that is not a balancing of quantity against quantity. There is not a law that is passed that is not the expression of force and force in tension (Bentley 1908, p. 37).

To simplify greatly, the advantage of groups in Bentley’s view is that groups allow intensity of preference to register in policy-making, whereas majoritarian institutions suppress information about intensity.

The difference is subtle, but the example in Table 1 makes it clear. Three people prefer policy A; the remaining minority of two people would prefer B. As we saw in the previous section, A would win with 60 % of the vote (3/5) if a vote is taken. But, in Bentley’s view, it is better to think of the two groups in a kind of tug-of-war. Those who favor A are pulling at a force equal to their willingness to pay—\$100, \$50, \$25, respectively—for a total force of 175 in dollar-utiles. Those who favor B pull a force of \$250 and \$300, for a total force of 550 dollar utiles. Clearly, the team pulling for B wins the tug-of-war. To Bentley’s credit, this is the “correct” answer, in utilitarian terms.

This is hardly surprising, because Bentley intentionally and explicitly invokes Jeremy Bentham’s “greatest good for the greatest number” maxim. Bentley’s claim is that only group competition can calculate this quantity exactly, because group action adds up utilities, whereas democracy adds up votes and that’s wrong, as we saw in the previous section. More generally, Bentley’s idea is to use W_i as the measure of the utility for the advocates for alternative A and W_j as the measure of utility for the advocates for B, where N_A and N_B are the numbers of supporters for each alternative:

$$\sum_{i=1 \text{ to } N_A} W_i \geq \sum_{j=1 \text{ to } N_B} W_j.$$

³⁰ This insight is developed much more deeply by Saari (2008); for a review, see Penn (2011). Saari (2008, p. 21) says of the Arrow result and other “voting paradoxes” that “This ignored information, which by not being used...makes these negative conclusions obvious rather than surprising...”.

In our example, $N_A > N_B$ ($3 > 2$) but the sum of the “pulls” W_j exerted by those who prefer B (\$550) exceeds the pulls W_i (\$175) exerted by those who favor A. So, B wins, as it should.

Truman (1951) articulated a process by which shared characteristics could be translated into organized groups. According to Truman, individual outlooks derive from their experiences in groups, and these group experiences are the primary ways that the individual knows the society. Shared characteristics plus interaction lead to the formation of groups. He writes of equilibrium in terms of stability and uniformity, but when this equilibrium is disturbed, new groups are formed and institutionalized in ways that make new claims on society. We may think of individuals with shared characteristics that are “potential groups,” subject to becoming organized in response to disturbances.

These ideas are presented in a descriptive framework, and no specific claims are made for Pareto optimality. But there is a kind of complacency, even optimism, in the identification of a mechanism by which shared attitudes, which constitute the interests, might be disturbed and then mobilized. Truman’s disturbance/reequilibration process fills in the mechanism by which Arthur Bentley’s vision of a process that reflects intensity of feeling might actually work.

Olson Jr. (1965) forever destroyed this complacency in *The Logic of Collective Action*, a book that argued that *group* costs and benefits are nearly irrelevant to the question of whether an effective group actually forms. Instead, Olson emphasized three factors that determine whether a group will form:

- Size.
- Concentration of benefits.
- Selective incentives.

Olson’s critique of the classical group theorists, and anyone else who wants to rely on group rent-seeking for benefits to society, is truly fundamental. *There is no necessary connection between the measured intensity of group preference and the underlying distribution of intensity of preference profiles in the population.* Most “groups” don’t form. But even worse, those groups that do manage to form are organized by an entrepreneurial rent-seeking leader, or around selective incentives that may have little to do with the group’s putative public purpose. Consequently, politically powerful groups may simply be exploiting even their own members, as well as other citizens outside the group who are not organized. The total dollar gains by the “winners” may be far less than the total dollar loss by losers, and yet the will of the winners is implemented as public policy.

In a later book, *The Rise and Decline of Nations*. Olson (1982) argued that the long term effects of an accretion of interest groups engaged in rent-seeking could be an actual decline of national productivity and efficiency. He suggested that the losers of World War II, such as Japan and West Germany, were doing better than some of the winners, such as Britain. The reason was that the losers saw their interest group structures wiped out by military defeat, which gave them an opportunity to start over.

3.4.4 Decentralized market processes, using prices

In a properly functioning market system, idealized to rule out externalities and public goods, individuals engage in fully informed *euvoluntary*³¹ exchanges. Exchange makes

³¹ For background on the conditions under which exchange is “truly,” or *euvoluntary*, see Guzmán and Munger (2014).

both parties better off, by definition. If there are unexploited opportunities for exchanges the result is deadweight loss. In any Pareto optimum situation there is no deadweight loss, by definition.

If each member of society is free to buy and sell resources, the result is not only that utility will be maximized, but also that resources will move to higher valued uses in each exchange. A buyer may acquire a commodity intending to consume it, but then learns that another buyer values it even more highly. A single commodity item may be exchanged many times, in each case moving to a higher valued use, and increasing the utility of both the buyer and the seller at each stage.

Consequently—if exchanges are *evolutionary* and transaction costs are negligible—commodities move ultimately to their highest valued uses. Exchange does not stop so long as any unexploited mutually beneficial exchanges remain unconsummated. So, when no more voluntary exchanges are possible, all the choosers and the economy as a whole are in equilibrium. But this equilibrium must have two features: no more exchanges are left in the decentralized market system that could raise the total value of dollar-utiles for the society, and no resource could be reallocated so as to make anyone better off without making at least one person worse off. These conditions, stated and proved in more rigorous fashion, are called the First Welfare Theorem and the Second Welfare Theorem of general equilibrium economics.

This result, an equilibrium that satisfies the two welfare theorems, is exactly equivalent to the existence and selection of an OPO. The result assumes (as do the other results) that the pre-existing distribution of wealth and primary goods is in some sense just, and that the property “owned” by all of the choosers in the society is legitimately acquired and can be sold to capture utility gains. So, given the resources possessed by each chooser, and the consequent effects on “willingness to pay” for a move between A and B, the result would simply emerge.³²

No formal negotiation of any kind would need to take place. The two citizens who prefer B would buy resources from the three citizens who would prefer A (and the bargaining space for doing so is non-empty, by definition, because of the surplus created by moving B), and everyone would wake up one morning and society would have “chosen” point B. If we could ask the three citizens would prefer A before if they wanted to return to A, they would answer, “No!” They were compensated in amounts greater than the amount of their dollar-utile preference for remaining at A. Moving back to A would harm not only Persons 4 and 5, but also Persons 1, 2, and 3.

This interpretation of the welfare theorems is quite old. Walras was interested in the problem in the 1870s, and a wide variety of attempts have been made to extend and clarify the meaning of the result (for a review, see Stiglitz 1991). Interestingly, this result also defines the standard of performance against which actual market processes are compared in the “so-called market failure paradigm.” The categories of market failure are simply the violations of the assumptions that together guarantee the existence of the general equilibrium. These are often described as

- Private goods only.
- No externalities.
- Perfect, common information sets.
- Agents are price-takers; no one can influence a price by bargaining.

³² This interpretation of equilibrium could be debated. See, for example, Arrow (1951) and Debreu (1959). For a clear interpretation of the intuitive meanings one might attach to equilibrium, see Dürpe and Weintraub (2010).

If all of these conditions are met, market processes discover and implement the OPO (though see Fisher 1983, for a cautionary note). But if any of the conditions is not met, market processes implement something other than the OPO, and may fail to choose *any* Pareto optimum, causing the economy to settle inside the production possibilities frontier.

3.4.5 Bargaining

Recalling our example of comparing A and B as outcomes, the problem is to find a way to move from A to B, because B is the OPO. The problem would be simple if only one person preferred B strongly, because that one person could simply pay off the damages caused to the other citizens, and still enjoy some or all of the surplus created in the move from A to B.

That logic is the essence of the “Coasian Bargaining” solution, usually conceived as a way of dealing with externalities or with assigning property rights. Coase’s fundamental insight was summarized in Coase (1960), but what got Coase thinking along these lines was the problem of assigning rights to sections of the radio spectrum (Coase 1959).

The public policy question being asked was how best to ensure that particular parts of the radio spectrum (“frequencies”) be awarded to the private entities (“stations”) that would value them most. The analysts at the Federal Communications Commission were treating the assignment of frequencies to stations as a complex matching problem, thinking about it as engineers. The amount of information required to solve the problem this way was enormous, and the information that could be obtained was likely to be unreliable.

Coase (1959) pointed out something that he thought was nearly self-evident, but which turned out to be (depending on one’s viewpoint) either profound or controversial. Coase argued that the *initial* assignment of stations was almost irrelevant to the question of how to ensure that the stations were matched to frequencies in a pattern that maximized the total value of the entire arrangement. Of course, some stations would at first be assigned frequencies that later would turn out to be more valuable in someone else’s hands. But if transaction costs are not high and it is possible to buy and sell frequencies freely, then over time the (re)assignment from bilateral bargaining and exchange would match the frequencies with the stations that valued them most.

The logic turns out to be quite general: if transaction cost are not prohibitive, the initial assignment is (nearly) irrelevant, because bargaining will “find” the highest valued solution, and implement it. In particular, this kind of bargaining works also in the cases of externalities, as Coase (1960) more famously argued. One potential problem is that this “invariance” (regardless of the initial assignment of property rights, the final allocation is efficient) property depends on some additional assumptions.³³ Coase was analyzing business decisions, where rights create profits and can be borrowed against, but such claims may be less plausible for individuals.

³³ The three most important assumptions about the Coasian bargaining process are (a) transactions costs are not prohibitive; (b) wealth effects do not prevent offers from being made, or force offers to be accepted; and (c) negotiations are not crippled by hold-up problems or bargaining asymmetries. We explicitly consider (a) below. Assumption (b) is a problem of the pre-existing distribution of wealth and power, rather than a problem with bargaining itself. As discussed earlier, our approach assumes that everyone has more or less the same wealth endowment. If that is not true, then there are *even more* problems with using government or market process. Finally, (c) is the condition we have been imposing from the outset. We thank David Weimer for helping to clarify our understanding of these problems, though he is not to blame for our interpretation.

For our example from Table 1, the “Coase Theorem” would work like this: Start at Point A, and suppose that transaction costs of organizing are not prohibitive. Then Person 4 and Person 5 will contract together to make compensating side payments to Persons 1, 2, and 3. Persons 1, 2, and 3 will therefore consent to move from Point A to Point B, which will in that case be a Pareto Improvement.

What exactly will be the side payments 4 and 5 offer to 1, 2, and 3? That cannot be predicted; the specific deal will be a product of the particular pattern of negotiations. But if for some reason we start with Point A as the status quo, we will move to Point B, the OPO.

Of course, we might start instead at Point B; what then? Persons 1, 2, and 3 will meet together, and make offers to 4 and 5 to try to persuade them to move to Point A. But given that Point B is worth \$275 more in total value than Point A, Persons 1, 2, and 3 will fail to sign a contract to move to Point A. So, if we start out at Point B, we remain at Point B.

Since Point B is (in our example) the OPO, that completes the “proof” of the Coase Theorem. The initial assignment of the status quo is irrelevant to the final outcome in the “social choice” problem, because—in the absence of prohibitive transaction costs—Coasian Bargaining will discover the OPO, and if it differs from the status quo the society will contract to move to the OPO. Where you start out only matters for the wealth positions of the participants, and has no effect on the overall distribution of productive resources.

The problem is the necessity that transaction costs not be prohibitive. In fairness to Coase, zero transaction costs was never an assumption he was remotely interested in. In fact, Coase (1988) argued that the effects of transaction costs were really his *primary interest*, when it came to the study of institutional arrangements.

Given what we know about group formation, and the Olson problem above, it seems dubious to conclude that a group (Persons 4 and 5) really could organize to persuade everyone else (Persons 1, 2, and 3) in the society to move to Point B, if for some reason we start at Point A as a status quo. If just two people were trying to persuade three people, using side payments, they might solve the problem.³⁴ If there were two thousand people, trying to negotiate with three thousand people, organizing to negotiate, collect, and disburse the side payments necessary to achieve unanimous consent seems much harder, and probably impossible. People on the side favoring B would try to free ride, and people on the side favoring A would hold out for larger payments. The genius of Coasian bargaining is that it achieves unanimous consent, and therefore euvoluntary exchange, simply by allowing the participants to discover a point in the bargaining space (which is nonempty, by assumption). The transaction costs associated with this bargaining are simply prohibitive in most cases wherein an actual ‘social’ choice is required.

Of course, we have also included deliberation under our heading of “bargaining,” but in that case the bargaining is much more about persuasion based on shared principles rather than side payments. Still, the principles are strikingly similar. A group of people gather. The group is small enough to be able to converse, hear claims, and respond to claims. Their goal is to reach unanimous agreement on the choice between two alternatives. Some authors (e.g., Dryzek 2010; Thompson 2008; Fishkin 2011; Gaus 2011; Steiner 2012) have claimed that the groups can reach consensus if their members are allowed to deliberate at length and share information and persuasive normative arguments.

³⁴ Note, however, that Coase’s examples (Coase 1960) generally are one-on-one, or at most among small groups, so that the group formation problem is simply finessed, although the bilateral monopoly that results is likewise indeterminate, though for different. An interesting extension of the “bargaining” or compensation solution is Groves (1979).

Interestingly, however, the mechanism on which Coasian bargaining focuses, side payments or compensation, would not be allowed.

If the group is small, all speak the same language, and share cultural values in common, it is quite plausible that at least some choices could be made in just the way that deliberative theorists claim. But if the group is large, and the “transaction costs” of agreement are high, because of language skills or the inability to make side payments, then as an empirical matter bargaining fails.

As far as we know, this way of categorizing “bargaining,” and the shared transaction costs problems of reaching agreement, have not previously been recognized. But the problem is implicit in the “decision costs” calculus of Buchanan and Tullock (1962) and the “transaction costs” approach of North (1990).

3.4.6 Cost-benefit analysis

The problem of CBA is, in terms of its conception of the problem, isomorphic with both Coasian bargaining and classical, Bentley-style interest group pluralism (Munger 2014). Specifically, the question is whether:

$$\sum_{i=1 \text{ to } N_A} W_i \geq \sum_{j=1 \text{ to } N_B} W_j.$$

The ethical foundations of the CBA approach were two papers (Kaldor 1939; Hicks 1939) that were in part a response to the intellectual cul de sac into which welfare economics had wandered. If in fact “interpersonal utility comparisons” were not allowed, as the subjectivist paradigm (Jevons 1871; Menger 1871; Pareto 1927) required, then how could economists say *anything* about choosing among Pareto optima? Just one year earlier, Harrod (1938, pp. 396–397); quoted in Chipman and Moore (1978, p. 547) asked the question this way:

Consider the Repeal of the Corn Laws. This tended to reduce the value of a specific factor of production—land. It can no doubt be shown that the gain to the community as a whole exceeded the loss to the landlords—but only if individuals are treated in some sense as equal. Otherwise how can the loss to some—and that there was a loss can hardly be denied—be compared with the general gain?

The defense that Kaldor and Hicks offered was that policy recommendations were generally impossible without some metric that goes beyond the simple Pareto criterion, because ultimately policy always has redistribute consequences.

This is exactly the same problem as choosing among Pareto optima, without compensation³⁵: If society moves from A to B, some citizens benefit and others are harmed. In effect, utility is taken from the losers and transferred to the gainers. The transfer might take the form of actual redistribution, with physical money or assets being moved around. But the policy might also have gains and losses that are simply utility differences, with no visible transfers of any kind. What is required is that these utility differences can be converted into a cardinal dollar metric, using willingness to pay for the *winner*s and the necessity of compensation to return the *loser*s to the same indifference curve as before.

So, Kaldor–Hicks proposed that cost-benefit analysis, performed by an expert agent of the state or a bureaucrat, was the answer. Add up the net losses to the losers, and the net

³⁵ For a discussion of this problem, see Stringham (2001) and Munger (2014).

gains to the winners. If the sum of the gains to the gainers exceeds the losses to the losers, then it is possible to transform the comparison of two Pareto optima to the standard Pareto improvement structure: the gainers gain more than enough to compensate the losers and still be better off after paying the compensation.

Overall, it is clear why CBA was taken to be a significant advance in welfare economics, particularly as a template for public project or program analysis. Requiring that economic experts be mute in any setting other than pure Pareto improvement leaves the passage of many potentially beneficial policy changes up to voting.

What is different about the Kaldor–Hicks approach to CBA is the mechanism by which the information about the values of gains and losses is acquired. Under Coasian bargaining, the participants speak for themselves, but the transaction costs of organizing all those participants makes a unanimous solution unlikely. In cost-benefit analysis, someone else always “speaks” for both the gainers and the losers. The process is centralized, greatly reducing the transaction costs of estimating a solution. Analyst-mandarins estimate measures of willingness to pay and desire for compensation. Further, analysts must decide whose gains and losses are allowed to “count,” solving the problem of *standing*.³⁶ So, while Coasian bargaining failed on transaction costs grounds, Kaldor–Hicks faces the difficulty of obtaining information in the face of the absence of any “discovery” procedure that can rely on prices or sincere bidding.

Interestingly, this problem has been encountered in a variety of collective choice contexts, as the “preference revelation” problem.³⁷ The generic problem of measuring or discovering sincere valuation is connected to the fact that no compensation or payment scheme is incentive compatible (Myerson 1979, 2008). Either survey responders know that the information they provide will determine their tax shares, in which case they free ride by under-reporting the value, or the information is unrelated to their tax shares, in which case they free ride by over-reporting the value. Consequently, just as with interest group pluralism, there is no reason to expect the intensity measures that derive from Kaldor–Hicks to match the true measures of utility on either side of the A versus B question. The reliance on experts, even if the experts are genuinely interested in neutral professionalism and the finding the right answer, is simply not justified by the results (Tetlock 2006).

Kaldor–Hicks is isomorphic with interest group pluralism in another way: *distributive consequences*. The Kaldor–Hicks approach is often called the (potential) ‘compensation principle’ because there is no requirement that the compensation actually be effected. So long as net gains are positive, and compensation is *possible*, the policy action is indicated and the OPO is identified. As Stokey and Zeckhauser (1978, p. 279) characterize this approach, “Make a change when total income increases; don’t worry about how it is distributed.” But this means that the outcome is coercive, in the sense that no consent is obtained from those who lose from the “improvement,” even if the premise of the move to improve is that the surplus created is enough to allow compensation. Kaldor–Hicks “improvements” need not be Pareto improvements, though the outcomes selected by Kaldor–Hicks must make Pareto improvement, or unanimous consent, *possible*.

³⁶ The issue of standing is quite complex. For a review of the issues, and some important arguments about who should “count,” see Whittington and MacRae (1986). They note that entities may be given standing by participating directly in decisions, by having their preferences counted, by having their welfare counted, or by representation by others whom they do not choose. Putting locks on doors clearly hurts the earnings potential of robbers; should we credit this very real cost in choosing public policy?

³⁷ See, for example, Tideman and Tullock (1976), Groves and Ledyard (1977). For a review, see Mueller (2003, pp. 162–168).

Overall, the CBA approach requires substantial compromises with both problems of information asymmetry (the authority doesn't know enough) and autonomy (the authority imposes what people should want, rather than what they do want, and they may all want different things). Nonetheless, CBA is the core procedure for centralized choice, and it is often indispensable if the choice has been made to use the state as the mechanism for delivering services or imposing regulations.

3.5 Overview

This section has examined five procedures for choosing and implementing one rather than another Pareto optimum. Such procedures are partly “discovery” processes, since the correct answer may not be known in advance. But procedures are also themselves policy choices, because some are likely to perform more reliably than others in different circumstances.

What we have argued is that the five procedures (voting, pluralism, markets, bargaining, and CBA) each have fundamental flaws that prevent it from being a universal solution. More simply, government procedures fail, in the sense that outcomes deviate from the optimal Pareto optimum systematically. We have embedded “markets” as one possible choice of the government authority to give market failure its due place among procedures that are imperfect. But in making this comparison we highlight the fact that market failures derive from the same essential problems, and for analogous reasons, as other types of failures of procedures.

4 Realism in behavioral premises

4.1 Beyond rationality

Until now, our “anatomy of government failures” has focused on aggregate entities. We looked first at *substantive* failures, or the fundamental problems of function and operation that face any state apparatus. We then turned to *procedural* failures, or the difficulties that collectivities face in trying to obtain information from, and organize the activities of, large groups of *homines economici*.

But people are not *homines economici*; they are people. Behavioral economics and psychology have raised important questions regarding the ability of citizens to apply rational decision theory to the problems they face as consumers and investors. And these are valid questions. The problem is the person who is too boundedly rational to decide what kind of car to buy does not become fully informed when he enters the voting booth. The corporate CEO who has a time horizon that is “too short” to invest for the future does not become more patient when she faces, as a member of the U.S. House of Representatives, an “up or out” election every 2 years. If anything, the problems of ignorance and bounded rationality limiting “correct” choices apply to political settings with even greater force.

Still, this work is only beginning to be developed, and models of human brain architecture and cognition are in their infancy. Consequently, we can do little more than plant a few signposts, perhaps highlighting some directions for further research and inquiry.

4.2 Cognition and decision

Daniel Kahneman, a psychologist who was awarded the Nobel Prize in Economic Sciences in 2002, argues that two modes of thinking are possible. “*System 1* operates automatically and quickly, with little or no effort and no sense of voluntary control. *System 2* allocates attention to the effortful mental activities that demand it, including complex computation. The operations of system 2 are often associated with the subjective experience of agency, choice and concentration” (Kahneman 2011, pp. 20–21).

These two systems correspond to what Haidt (2012, p. 53) calls the *rider* “(controlled processes and reasoning why)” and the *elephant* “(automatic processes, including emotion, intuition ...).” Haidt explicitly links his metaphor to Kahneman’s (2011, p. 385) two systems: System 2 is the rider trying to control the elephant, and perhaps to rationalize and defend its behavior. System 1 is the elephant. Haidt suggests that—as the metaphor itself implies—the elephant is much more powerful than the rider. For our purposes, these considerations are important, but need to be raised as refinements of the more general problems we have discussed earlier in this paper. Institutions and procedures matter, and affect the way that elites act and citizens choose. But neither Kahneman nor Haidt distinguish elites and masses, or consider institutional contexts. This is appropriate for their purposes, because they are analyzing human judgment and decision making in general. Our purposes require that we consider the connections between institutional arrangements and procedures and the problems raised in the purely individual-level work of Ariely (2008), Lodge and Taber (2013), Haidt, Kahneman, and others.

One of the key insights of Public Choice is that poorly informed and imperfect consumers do not become omniscient geniuses when they go to vote. They are essentially the same person, though since the institutional settings and incentives differ the actual behavior may be different. The question is, given the way real human beings think about choices, and choose, what can be said about correspondence between analysts’ hopes for democratic processes and the actual outcomes we observe?

In a way, our inquiry echoes that of Riker (1982). Riker asked whether the faith people have in democracy is justified, in the sense that “fair” decision rules must lead to “good” outcomes. He argued that this might not be so, though on procedural grounds like those we discussed in the previous section.

There is another potential problem with the “fair decision rules lead to good outcomes” paradigm, however, if the behavioral assumptions that underlie the claim are not borne out empirically. That is, if voters—acting in the way voters really act—choose alternatives that voters themselves, as a group, would not willingly select, then that is a government failure. The complicating, and possibly aggravating, factor is that elites may recognize voter frailties and foibles, and manipulate the outcome. In that case, the result may be worse than failing to select the best outcome. The “best” outcome may not even be available, because elites are able to restrict the set of alternatives for their own benefit.

As models of the behavior of mass publics, a more verisimilous view of human psychology raises important additional questions about democracy, which roots government policy in the consent of the governed, and about autocracy, which roots government in the desire and capacity of technocrats to make wise choices. Consequently, in the rest of this section we will restrict our attention to the implications of the behavioral psychology for the effectiveness of democratic government. Two major sets of complicating factors are considered: (A) Information problems, and (B) Loss aversion.

4.3 Information problems

Information problems are ubiquitous in human decision settings. In fact, “information asymmetry” is one of the key causes of “market failure,” and is used to justify government interventions such as the Food and Drug Administration, or occupational licensing for physicians or attorneys. Markets do have some responses to the value of information, of course, providing information about electronics or appliances in *Consumer Reports* or labels from Underwriters’ Laboratories. And if you are not sure whether you will like a new movie, you can go to *Rotten Tomatoes* and find out what other people think.

The problem is that recognizing that *consumers* may be poorly informed and subject to manipulation, provides no cause for optimism that *voters*, who are after all the same people, will be informed and immune to manipulation. In fact, there are good reasons for believing that citizens-as-voters may be even less informed than citizens-as-consumers.

- The information required to make political choices is complex and aggregated, since candidates or parties are bundles of issues.
- The feedback mechanism for evaluating choice is murky, separated in time from the choices (elections happen long before policy effects are felt, and sometimes long after policy choices were made), and subject to ideological filtering.
- The act of acquiring accurate information is a “public good,” because knowing the right answer does not allow a voter to select an outcome, but merely to be one voice among many.
- The act of participating itself is a “public good,” because if a citizen becomes informed and votes the positive effects are available also to those who did not vote. When we consider that even the best informed vote counts no more than an ill-informed vote, and is not likely to affect the outcome of the election, it becomes clear that this public good, just as in a market setting and for the same reasons, is likely to be undersupplied. Even worse, perhaps, as Brennan (2011) argues, the negative externality of uninformed voting will likely be overproduced.

The problem is clear: Markets underprovide public goods, and overprovide public bads, but so do democracies, and for the same reasons exactly: no individual has an incentive contribute to the public good, because (a) *his contribution* doesn’t make the preferred outcome much *more* likely, and (b) *failing to contribute* doesn’t make the preferred outcome much *less* likely.

Not surprisingly, conventions and institutions evolve to help solve these problems, but only imperfectly. The information problem is (partly) solved by heuristics or short-cuts, substituting a habit or reputation for cognitive wear and tear. In broad political contexts these heuristics are called “ideologies,” which may take the form of vague value statements (Downs 1957) or more concrete but still amorphous reduced-dimensional spaces (Hinich and Munger 1994). It is important to note that these are not ideologies in the integrated, cognition-intensive sense meant by Converse (1964), but rather connections to a vague left–right dimension that may “fit” the actual issue space quite imperfectly.

The turnout problem is (partly) solved by the institution of political parties, or “teams” of elites seeking elective office by endorsing candidates and organizing electoral mass infrastructures. As Aldrich (2011) points out, the connection between the ideological motivation and the partisan delivery of policy together constitute the psychological basis for the “party identification” felt, and acted on, by voters.

The difficulty is that parties may not be very competitive, particularly in first-past-the-post or plurality systems. It may be difficult for candidates to gain access to the ballot, and without partisan endorsement outside candidates are prevented from competing effectively unless they have unlimited sources of funds, as was the case with Ross Perot in 1992. Given the slippage between the ability of voters to discipline candidates or parties they vote for and the actual policies implemented, there is no close connection between the results voters want—if voters even knew what that was—and the policies implemented and then carried out by political elites. Voter behavior may serve individual goals, but the collective outcome is different from what they want. And elites are able to manipulate this separation between promises (in elections) and informed feedback (in particular policy settings).

According to Downs (1957, pp. 212–214), rationality requires that citizens choose selection principles that deliver (in expectation) their desired ends. Further, a rational voter—at least in the Downsian view—is willing to update selection principles when confronted with confounding evidence, even if this requires that he or she must revise core beliefs. But this claim is implausible, given the complexity of such information and the clear incentives of elites to obscure the process of evaluation. As North (1990) put it:

the actors would always possess true models or if they initially possessed incorrect models the information feedback would correct them. But that version of the rational actor model has simply led us astray. The actors frequently must act on incomplete information and process the information that they do receive through mental constructs that can result in persistently inefficient paths.

The question is whether political institutions—meaning ideologies and parties—are capable of solving, or substantially mitigating, the information problem facing voters. Downs (1957, p. 83) himself strikes a cautionary note:

As soon as uncertainty appears, the clear path from taste structure to voting decision becomes obscured by lack of knowledge... [Some voters] are highly uncertain about which party they prefer. They are not sure what is happening to them, or what would happen to them if another party were in power. They need more facts to establish a clear preference. By providing these facts, [politicians] find an opportunity to become effective.

[But political elites] are not interested per se in helping people who are uncertain become less so; they want certainty to produce a decision which aids their cause. Therefore they provide only those facts that are favorable to whatever group they are supporting...

This problem means that what Downs (1957) called “persuaders” may be different from neutral and disinterested informers, which would act more like news media and report (at most) facts. Persuaders are free to use all of the classical tricks of rhetoric and framing, portraying their candidates in the best possible light and selectively attacking their opponents. Candidates and campaign consultants can use advertising to prime projection (Conover and Feldman 1982), framing (Tversky and Kahneman 1981), or latent prejudices (Mendelberg 2001) in ways that bear no resemblance to the “rational” model of voter choice.

This disconnect between voter information, voter choice, and political outcomes means that no anchor is available in policy discussions for voters to use to evaluate candidates. Worse, there may be no reason to expect that voter ignorance will even be “rational,” in the limited Downsian sense of updating when confronted with new information. This view

was argued persuasively by Zaller (1992) in his “Receive, Accept, Sample” model. Elites are more likely to receive messages, but less likely to accept them if the message contradicts the views the person already holds. And less politically sophisticated voters either fail to receive information or else sample only very shallow and contradictory sources, preventing them from having accurate views of candidates or other political phenomena.

Johnston et al. (2014) distinguish between *instrumental* and *identity* politics. As the instrumentally oriented rational choice tradition makes clear, it is rational not to invest much energy in informing one’s self about politics, because the chances that one’s single vote will make a difference in the election is infinitesimal. However, politics is not just instrumental. It also is an expression of one’s identity. Johnston et al. (2014, p. 75; see also Lenz 2012) say that “In mass politics, the quest for positive distinctiveness is displayed by a tendency to express partisan allegiance through the adoption of party-approved beliefs, attitudes, and behaviors.”

This view reverses the classical Downsian spatial model. Instead of voters’ policy preferences driving the choice of candidates and parties, identity with candidates and parties drives policy preferences. If this view is correct, or even partially correct, it presents a serious challenge not just to full-information models of rational choice, but even to the “rational ignorance” view. Along these lines, Achen and Bartels (2015) undermine the common assumption that voters have policy preferences that drive candidate choice. They suggest that causality runs the other way, as well, with voters having policy preferences that are driven by partisan identity. If party drives policy preference instead of preferences driving partisan choice, the possibility of political agreement is undermined. Only if parties agree will voters agree in this context.

In the worst case, such “ignorance” can lead to persistent errors that depend more on voters’ expressions of identity and ideological belonging than any kind of cognitive process. Brennan and Lomasky (1997) noted that the expression of preference is divorced from aggregation. So voters choose to express the viewpoint or moral value that brings them credit from people around them. The fact that this view is inconsistent with facts, or that the policy does not actually achieve its objectives, is a public goods problem. Each voter would prefer better policy, but each voter is better off individually expressing the popular and estimable view. Munger and Schaller (1997) used this model to explain the apparent paradox that most people opposed Prohibition privately but voted for it publicly, at least for the 18th Amendment the first time around.

This critique is pushed even further by Caplan (2007) and Somin (2013). At best, in their view, voters might be “rationally irrational” (Caplan 2001, p. 4). Far from updating beliefs in response to evidence, voters are free to choose their beliefs in accord with what facts they would like to be true about the world. They pay no penalty for persistently believing and acting on absurdly false facts, and are rewarded by others in the group with which they identify for having “good” beliefs.

Thus, political affiliations are more like sports team loyalties than rationally held beliefs about good policies. If you identify with party A, then you adopt and defend the policy positions advocated by party A. This claim, if true (contra Aldrich 2011; Lupia and McCubbins 1998), completely reverses the Downsian logic, in which partisanship is an information shortcut leading voters to choose parties that are closer to their policy preferences. Voters choose a party based on the attractiveness of its basic value statements, and then adopt the implied policy positions. The causal link assumed by democracy enthusiasts connecting choice and good policy is completely broken. There is no reason to expect democratic means to lead to good government, and there is no way to recognize or correct bad government when it happens.

4.4 Loss aversion/endowment effects

An asymmetry arises between losing something one already has, and the prospect of gaining something new. A wide variety of experiments in human psychology show that if we are threatened with, or have suffered, a loss we will expend far more resources to recoup the loss than we would have spent to obtain an equivalent gain. This is different from simple “diminishing marginal utility of income,” where a given loss hurts more than the same gain would help. In this case, the endowment or sense of right to possession or control dominates our rational sense of calculation. What that means is that we will spend more effort, more resources, and suffer more risk, than recouping the loss is “worth” in simple material terms.

This effect, called “the endowment effect” (Kahneman et al. 1990; Kahneman 2011, pp. 189–199) has been found to explain a wide variety of apparently “irrational” choices and behaviors, in different contexts. An interesting recent application is Berejikian and Early (2013), who use the endowment effect to explain “foreign policy resolve,” or the willingness of nations to expend substantial resources to win back concessions, territory, or favorable policies that are perceived as belonging to the nation. They seem much less willing to expend resources or take risks to obtain equivalent new benefits of comparable or even larger size.

This raises an important point: rather than satisfying the rationalist, “realist” conception of foreign policy, national choices are driven by domestic political considerations. And voters, being *homines politici*, vote based on loss aversion. It is unacceptable for a national leader, or a party, to give up territory or favorable treatment from another nation without great loss of face.

Perhaps even more importantly, the problem of loss aversion connects with Olson’s (1965) collective action problem. Olson claimed that policy asymmetry emerges because small groups enjoying concentrated benefits would find it much easier, in transaction costs terms, to organize for benefits. But a surprising asymmetry emerges after the fact, also, one which Olson later (Olson 1982) called “institutional sclerosis.”

It is a truism that a program, once started, is very hard to get rid of. The explanation often given is that the benefits are concentrated and the costs are diffuse. But more importantly, the loss of rent (public works pensions in California cities, farm price supports, and so forth ad infinitum) is opposed fiercely by those who have come to count on it as part of their income, and promoted only weakly by those who might expect to gain.

To oversimplify, suppose that four people now enjoy a rent, and two people would benefit from ending the program. Imagine that the cost of the rent is larger than the amount of the gain or benefit it creates. The standard “logic of collective action” might dictate that the smaller group would win: there are only two of them, and the gains are larger than the losses. On purely “collective action” grounds we would expect the program to be ended.

But the “endowment effect” would imply that the four people who already “owned” the rent would fight hard to preserve the program, dissipating much of the value of the rent every year. The two who would benefit from ending the program would only have a conjectural, and not yet achieved, benefit. Consequently, the sizes of the groups and their abilities to solve collective action problems are no longer central, *once the program is in place*. For this reason, we expect institutional sclerosis to be a very severe and intractable problem, even if it imposes large deadweight losses, and even if it cuts against the usual “logic of collective action.”

Loss aversion thus exhibits a key connection with rent-seeking, which was discussed earlier. A group that succeeds in obtaining for itself a rent, through political action or the

award of a property right, comes to perceive that rent and the returns it attracts as part of their endowment. Farm price supports, taxi medallions, and public pensions far in excess of what municipal governments can afford all persist longer than appears “rational,” given the size of the costs to the public. But the groups that “won” these rents would suffer a loss, the prospect of which leads them to organize feverishly, and with much more energy and purpose than the impulse to organize on the other side, which would stand to receive a benefit they do not yet possess or perceive to be part of their endowment.

In this section we have examined some additional possible causes of government failure. A large part of the democratic tradition focuses on the willingness and capacity of the state to act on “the will of the people.” We have tried to simplify the problem of identifying the will of the people in previous sections by eliminating fundamental problems of social choice.

But two difficulties remain, and we have discussed each of them in this section. The first problem is information: can voters learn what they need to know? And does any reason exist to expect that voters will voluntarily provide the public goods of becoming informed and participating in politics? The answer is at least a qualified “no.” Voters’ reasons for choosing one alternative have more to do with identity, projection, and framing than with any understanding of the effects of policy on their welfare, or of the connections of outcomes with their moral values. Instead, their moral values are used to select identities and parties; those policies then are the one the voter “supports.” The connection between good voter choice and good policy outcomes is disrupted, and there is no reason for optimism as politics grow larger and less connected by traditional social ties.

The second problem is loss aversion or the endowment effect. This result from behavioral psychology would be important enough in and of itself, because it explains much of the foreign policy and domestic paradoxes that seem like simple stubbornness or irrationality. But when combined with the collective action problem, loss aversion becomes a powerful tool for explaining institutional sclerosis and the persistence of programs that are, from a rational perspective, wasteful, inefficient, or obsolete.

5 Final words

The concept of government failure is a much bigger problem than its contemporary treatment has admitted. The concept has been trapped in the cocoon of competitive equilibrium theory. CET is one of the greatest achievements of the social sciences, but “failures” using CET as a benchmark do not do justice to either the damage that unconstrained government can do, or the importance of effective government for well-working markets, not to speak of general well-being. Whether the cocoon produces a beautiful butterfly or a destructive moth, we hope that we can now move beyond a simplistic “government failure is the failure to correct market failures” approach.

We began by asserting that many of the greatest catastrophes of human history have been government failures. While our examples were illustrative rather than comprehensive, our point is that government failure is a sufficiently important problem to be considered in its own right, rather than as derivative of, and secondary to a theory of perfect markets.

We have laid the groundwork in three steps for a conception of government failure that does not depend on even the theoretical existence of market perfection. The first step was a basic list of *substantive* failures that all governments may yield to if they do not cultivate and maintain appropriate institutions. These institutions can be formal, written rules, or unwritten practices that constitute the “rules of the game.” Their purpose is to assure that

government is strong enough to preserve order and liberty, to discipline the temptations to have an irresponsible and unsustainable fiscal and monetary policy, or to use public power for private purposes. This list is not exhaustive, and it concentrates on risks that are systematic and continuing rather than situational.

The second step was to show the limitations of a comprehensive set of *procedural* alternatives for making collective decisions. We extended Arrow's basic distinction between government and markets into five more precise procedures for choosing among alternatives: voting, interest groups, markets, bargaining, and cost-benefit analysis. We assumed away two major problems of social choice: the problem of cyclical majorities, by narrowing the alternatives to two, and the problem of interpersonal comparison of utility, by assigning cardinal utilities. With these unrealistic assumptions, we showed that all five mechanisms still had crippling limitations. They either arbitrary outcomes or they break down with the participation of large numbers.

Third, we have acknowledged the validity of the critique of the classical model of consumer choice and markets process arising from new work in behavioral economics and psychology. Human beings are only partly rational, in the literal sense that only parts of our brains, and only a segment of our decisions, fit that model. The calculating and reasoning part of us coexists with another part that is more instinctive and emotional. Moreover, there is reason to believe that the rational side of human behavior and choice is influenced if not driven by the more instinctive and emotional side. These basic facts, along with some more narrowly defined patterns of behavior like loss aversion and vulnerability to heuristics and framing, make individual choices and decisions less dependable and more manipulable than has commonly been assumed in fields influenced by rational choice models. But our answer is that this critique, while valid, applies with equal and probably greater force to models of "rational" choice in politics. If consumers are too limited cognitively to choose toasters, why are voters able to choose detailed programs for picking winners and losers in the economy of the future? If markets are flawed, politics is flawed for the same reasons, and with more risky consequences.

Ultimately, the conclusion one might draw from the work we have reviewed here is that it is a mistake to think of questions of policy and regulation as "more" government versus "less" government. It is possible for government to fail either by doing too little or by doing the wrong things with great energy and efficiency. Given the possible permutations of the varieties of problems we have identified, it may seem surprising that government does not fail more than it does. But that suggests a more optimistic note: like market failures, the problems we have identified are potential problems, and in many cases the "problems" we have talked about are simply answers to questions such as "Why does a particular institutional arrangement exist, and not some other arrangement?" Government institutions, like market institutions, may well be solutions to problems that no one understands until decades later. Our hope is that a more complete anatomy of the potential failures of governments in the abstract will spur a deeper understanding of political institutions in the real world of political economy.

Acknowledgments With apologies to Francis M. Bator (1958) regarding the title. Earlier versions of this paper were presented at the 2012 meetings of the Public Choice Society, the 2014 meetings of the Association for Public Economic Theory, the 2014 meetings of the American Political Science Association, and faculty seminars at the University of Georgia, Stanford Law School, Istituto Bruno Leoni, Milan, Italy, Australian National University in Sydney, and Masaryk University in Brno, Czech Republic. We acknowledge helpful comments of Christopher Achen, Scott Ainsworth, Jonathan Anomaly, Larry Bartels, Michael Becher, Peter Boettke, Geoffrey Brennan, Brett Carter, Keith Dougherty, Robert Grafstein, Sean Ingham, Libby Jenke, Christopher Johnston, David Levy, Leonidas Montes, Sam Peltzman, Richard

Salsman, Thomas Schwartz, William Shughart, Georg Vanberg, David Weimer, Barry Weingast, Samuel Wilson, Stanley Winer, and Leonidas Zelmanovitz on previous drafts or in useful discussions. We offer particular thanks to Carl Simon, who spent hours listening to early versions and helping us make more sense. But the interpretations offered here, as well as the errors, are the authors' alone.

References

- Acemoglu, D., & Robinson, J. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. New York: Crown Business Publishing.
- Achen, C. H., & Bartels, L. M. (2015). *Democracy for realists* (draft manuscript).
- Aldrich, J. (2011). *Why parties? A second look*. Chicago: University of Chicago Press.
- Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions*. New York: Harper-Collins Publishers.
- Arrow, K. J. (1963). *Social choice and individual values*, 2nd edn. New York: Wiley.
- Arrow, K. J. (1951). An extension of the basic theorems of classical welfare economics. In *Proceedings of the second Berkeley symposium*. Berkeley: University of California Press.
- Backhouse, R. E., & Medema, S. G. (2012). Economists and the analysis of government failure: Fallacies in the Chicago and Virginia interpretations of Cambridge welfare economics. *Cambridge Journal of Economics*, 36, 981–994.
- Bator, F. M. (1958). The anatomy of market failure. *Quarterly Journal of Economics*, 72(3), 351–379.
- Becker, G. (1983). A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics*, 98, 371–400.
- Bentham, J. (1907). *Introduction to the principles of morals and legislation*. Oxford: Oxford University Press.
- Bentley, A. (1908). *The process of government*. Chicago: University of Chicago Press.
- Berejikian, J., & Early, B. R. (2013). Loss aversion and foreign policy resolve. *Political Psychology*, 34(5), 649–671.
- Bergson, A. (1938). A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, 52, 314–344.
- Bernholz, P. (2003). *Monetary regimes and inflation: History, economic and political relationships*. Northampton, MA: Edward Elgar Publishing.
- Besley, T. (2006). *Principled agents? The political economy of good government*. Oxford: Oxford University Press.
- Boettke, P., Coyne, C., & Leeson, P. (2007). Saving government failure theory from itself: recasting political economy from an Austrian perspective. *Constitutional Political Economy*, 18, 127–143.
- Brennan, H. G., & Buchanan, J. M. (1985). *The reason of rules: Constitutional political economy*. New York: Cambridge University Press.
- Brennan, H. G., & Lomasky, L. (1997). *Democracy and decision: The pure theory of electoral preference*. New York: Cambridge University Press.
- Brennan, J. (2011). *The ethics of voting*. Princeton: Princeton University Press.
- Buchanan, J. M. (1979). *Cost and choice: An inquiry in economic theory*. Chicago: University of Chicago Press.
- Buchanan, J. M. (1983). The achievement and the limits of public choice in diagnosing government failure and in offering bases for constructive reform. In H. Hanusch (Ed.), *Anatomy of government deficiencies* (pp. 15–25). Berlin: Springer-Verlag.
- Buchanan, J. M. (1987). The constitution of economic policy. *Science*, 236(June), 1433–1436.
- Buchanan, J. M. (2001). *The intellectual portrait series: A conversation with James M. Buchanan*. Interviewed by H. Geoffrey Brennan. Indianapolis: Liberty Fund. Retrieved from <http://oll.libertyfund.org/titles/1030>.
- Buchanan, J., & Tullock, G. (1962). *The calculus of consent*. Ann Arbor: University of Michigan Press.
- Buchanan, J., & Wagner, R. (1977). *Democracy in deficit: The political legacy of Lord Keynes*. New York: Academic Press.
- Caplan, B. (2001). Rational ignorance vs. rational irrationality. *Kyklos*, 54(1), 3–26.
- Caplan, B. (2004). Is socialism really “impossible”? *Critical Review*, 16, 33–52.
- Caplan, B. (2007). *The myth of the rational voter: Why democracies choose bad policies*. Princeton: Princeton University Press.
- Chipman, J., & Moore, J. (1978). The new welfare economics. *International Economic Review*, 19, 547–584.

- Coase, R. H. (1959). The federal communications commission. *Journal of Law and Economics*, 2, 1–40.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1–44.
- Coase, R. H. (1964). The regulated industries: Discussion. *American Economic Review*, 54(2), 192–197.
- Coase, R. H. (1988). *The firm, the market, and the law*. Chicago: University of Chicago Press.
- Congleton, R., Hillman, A., & Konrad, K. (Eds.). (2008). *Forty years of research on rent seeking*. New York: Springer-Verlag Academic Publishers.
- Conover, P. J., & Feldman, S. (1982). Projection and the perception of candidates. *Western Political Quarterly*, 35, 228–244.
- Converse, P. E. (1964). The nature of belief systems in mass publics. In David Apter (Ed.), *Ideology and discontent* (pp. 219–241). New York: Free Press.
- Cooter, R., & Siegel, N. (2010). Collective action federalism: A general theory of Article I, Section 8. *Stanford Law Review*, 63, 115–186.
- Das, I. (1999). A preference ordering among various Pareto optimal alternatives. *Structural and Multidisciplinary Optimization*, 18(1), 30–35.
- de Mariana, J. (1839). *Historia general de España* (Vol. 3). Barcelona: D. Francisco Oliva.
- Debreu, G. (1959). *Theory of value: An axiomatic analysis of economic equilibrium*. New Haven, CT: Cowles Foundation/Yale University Press.
- Defoe, D. (1968). *The life and adventures of Robinson Crusoe*. Boston: Houghton Mifflin.
- Demsetz, H. (1969). Information and efficiency: Another viewpoint. *Journal of Law and Economics*, 12(1), 1–22.
- Diamond, J. (2005). *Collapse: How societies choose to fail or succeed*. New York: Penguin Books.
- Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3), 297–308.
- Dolfsma, W. (2013). *Government failure: Society, markets and rules*. Cheltenham: Edward Elgar.
- Donohue, J. J., & Levitt, S. D. (2001). The impact of legalized abortion on crime. *Quarterly Journal of Economics*, 119, 249–275.
- Downs, A. (1957). *An economic theory of democracy*. New York: Harper and Row.
- Dryzek, J. (2010). *Foundations and frontiers of deliberative governance*. New York: Oxford University Press.
- Düppe, T., & Weintraub, E. R. (2010). *Finding equilibrium: Arrow, Debreu, McKenzie and the problem of scientific credit*. Princeton, NJ: Princeton University Press.
- Eyzaguirre, H., Ferrarini, T., & O’Roark, J. B. (2014). Textbook confessions: Government failure. *Journal of Private Enterprise*, 29(3), 159–175.
- Ezrow, N., & Frantz, E. (2013). *Failed states and institutional decay: Understanding instability and poverty in the developing world*. New York: Bloomsbury Academic Press.
- Ferrarini, T., Gwartney, J. D., & Morton, J. S. (2011). Advanced placement economics: The good, the bad, and the ugly. *Econ Journal Watch*, 8(1), 57–75.
- Fisher, F. M. (1983). *Disequilibrium foundations of equilibrium economics*, Econometric Society Monographs in pure Theory. New York: Cambridge University Press.
- Fishkin, J. (2011). *When the people speak*. New York: Oxford University Press.
- Friedman, D. (1989). *The machinery of freedom: Guide to a radical capitalism*. LaSalle, IL: Open Court Publishing.
- Gaus, G. F. (2005). The place of autonomy within liberalism. In J. Christman & J. Anderson (Eds.), *Autonomy and the challenges to liberalism* (pp. 271–306). New York: Cambridge University Press.
- Gaus, G. F. (2011). *The order of public reason: A theory of freedom and morality in a diverse and bounded world*. New York: Cambridge University Press.
- Groves, T. (1979). Efficient collective choice when compensation is possible. *Review of Economic Studies*, 46, 227–241.
- Groves, T., & Ledyard, J. (1977). Some limitations of demand-revealing processes. *Public Choice*, 29, 107–124.
- Guzmán, R. A., & Munger, M. C. (2014). Euvoluntariness and just market exchange: Moral dilemmas from Locke’s *Venditio*. *Public Choice*, 158(1–2), 39–49.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon Books.
- Hanson, J. (2014). Forging and taming Leviathan: State capacity, constraints on rulers, and development. *International Studies Quarterly*, 58(2), 380–392.
- Hardin, R. (1990). Rationally justifying political coercion. *Journal of Philosophical Research*, 15, 79–91.
- Harrod, R. F. (1938). Scope and methods of economics. *Economic Journal*, 48, 383–412.
- Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519–530.
- Hayek, F. A. (1948). *Individualism and economic order*. Chicago: University of Chicago Press.

- Hayek, F. A. (1973). *Law, legislation, and liberty*. Chicago: University of Chicago Press.
- Hicks, J. (1939). The foundations of welfare economics. *Economic Journal*, 49(196), 696–712.
- Hinich, M., & Munger, M. C. (1994). *Ideology and the theory of political choice*. Ann Arbor: University of Michigan Press.
- Hobbes, T. (1651, 2012). *Leviathan*. Oxford: The Clarendon Press.
- Huntington, S. P. (1968). *Political order in changing societies*. New Haven: Yale University Press.
- Jevons, W. S. (1871, 2013). *The theory of political economy*. Houndmills: Palgrave Macmillan.
- Johnston, C. D., Lavine, H., & Federico, C. M. (2014). *Personality, parties and the foundations of economic opinion*. Unpublished manuscript, Duke University Department of Political Science.
- Kahneman, D. (2011). *Thinking: Fast and slow*. Farrar, Straus and Giroux: New York.
- Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental test of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98(6), 1325–1348.
- Kaldor, N. (1939). Welfare propositions in economics and interpersonal comparisons of utility. *Economic Journal*, 49(195), 549–552.
- Keeney, R. L., & Kirkwood, C. W. (1975). Group decision making using cardinal social welfare functions. *Management Science*, 22(4), 430–437.
- Lange, O., & Taylor, F. M. (1938). *On the economic theory of socialism*. Minneapolis: University of Minnesota Press.
- Ledyard, J. O. (2008). Market failure. In S. N. Durlauf & L. E. Blume (Eds.), *The new Palgrave dictionary of economics* (2nd ed.). New York: Palgrave Macmillan.
- Lenz, G. (2012). *Follow the leader? How voters respond to politicians' policies and performance*. Chicago: University of Chicago Press.
- Lerner, A. (1944). *The economics of control: Principles of welfare economics*. New York: Macmillan.
- Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. New York: Cambridge University Press.
- Lupia, A., & McCubbins, M. (1998). *The democratic dilemma: Can voters learn what they need to know?* New York: Cambridge University Press.
- May, K. O. (1952). A set of independent, necessary, and sufficient conditions for simple majority decision. *Econometrica*, 20, 680–684.
- McKean, R. N. (1965). The unseen hand in government. *American Economic Review*, 55(3), 496–506.
- Medema, S. G. (2009). *The hesitant hand: Taming self-interest in the history of economic ideas*. Princeton: Princeton University Press.
- Medema, S. G. (2010). Pigou's "Prima Facie Case": Market failure in theory and practice. In R. E. Backhouse & T. Nishizawa (Eds.), *No wealth but life: Welfare economics and the welfare state in Britain, 1880-1945*. New York: Cambridge University Press.
- Mendelberg, T. (2001). *The race card: Campaign strategy, implicit messages, and the norm of equality*. Princeton, NJ: Princeton University Press.
- Menger, C. (1871, 1981). *Grundsätze der Volkswirtschaftslehre (Principles of Economics)*. New York: New York University Press.
- Mill, J. S. (2002, 1859). *On liberty*. New York: Dover Publications.
- Mueller, D. (2003). *Public choice III*. New York: Cambridge University Press.
- Munger, M. C. (2014). Kaldor-Hicks coercion, Coasian bargaining, and the state. In J. Martinez & S. Winer (Eds.), *Coercion and social welfare in public finance: Economic and political dimensions* (pp. 117–135). New York: Cambridge University Press.
- Munger, M. C., & Schaller, T. (1997). The prohibition-repeal amendments: A natural experiment in interest group influence. *Public Choice*, 90(4), 139–163.
- Myerson, R. (1991). *Game theory: Analysis of conflict*. Cambridge: Harvard University Press.
- Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, 47(1), 61–74.
- Myerson, R. B. (2008). Mechanism design. In: S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*, 2nd edn. New York: Palgrave.
- Narveson, J., & Sanders, J. (Eds.). (1996). *For and against the state*. Lanham, MD: Rowman & Littlefield.
- North, D. (1990). *Institutions, institutional change, and economic growth*. Cambridge: Cambridge University Press.
- North, D. (1994). Economic performance through time. *American Economic Review*, 84(3), 359–378.
- North, D., Wallis, J., & Weingast, B. (2012). *Violence and social orders: A conceptual framework for interpreting recorded human history*. New York: Cambridge University Press.
- North, D., & Weingast, B. (1989). Constitutions and commitment: The evolution of institutions governing public choice in seventeenth century England. *Journal of Economic History*, 49, 803–832.
- Olson, M, Jr. (1965). *The logic of collective action*. Cambridge: Harvard University Press.
- Olson, M, Jr. (1982). *The rise and decline of nations*. New Haven: Yale University Press.

- Pareto, V. (1971/1927). *Manual of political economy*. New York: A. M. Kelley. (Translated from the French edition of 1927).
- Penn, E. M. (2011). Impossibility theorems and paradoxes in collective choice theory. In J. J. Cochran (Ed.), *The Wiley encyclopedia of operations research and management science*. New York: Wiley.
- Pigou, A. C. (1912). *Wealth and welfare*. London: Macmillan.
- Pigou, A. C. (1935). State action and laissez-faire. In *Economics in practice: Six lectures on current issues* (pp. 107–128). London: Macmillan and Co.
- Reinhart, C., & Rogoff, K. (2009). *This time is different: Eight centuries of financial folly*. Princeton: Princeton University Press.
- Riker, W. H. (1982). *Liberalism against populism: A confrontation between the theory of democracy and the theory of social choice*. San Francisco: W.H. Freeman and Company.
- Romer, T., & Rosenthal, H. (1978). Political resource allocation, controlled agendas, and the status quo. *Public Choice*, 33(4), 27–43.
- Romer, T., & Rosenthal, H. (1979). Bureaucrats versus voters: On the political economy of resource allocation by direct democracy. *Quarterly Journal of Economics*, 93(4), 563–587.
- Rothbard, M. N. (1978). *For a new liberty: The libertarian manifesto*. New York: Collier Macmillan.
- Rummel, R. J. (1997). *Death by government*. New Brunswick: Transactions Publishers.
- Saari, D. G. (2008). *Disposing dictators, demystifying voting paradoxes: Social choice analysis*. Cambridge: Cambridge University Press.
- Samuelson, P. (1947). *Foundations of economic analysis*. Cambridge: Harvard University Press.
- Samuelson, P. (1954). The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4), 387–389.
- Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. New Haven: Yale University Press.
- Scott, J. C. (2009). *The art of not being governed*. New Haven: Yale University Press.
- Selgin, G., & White, L. (1987). The evolution of a free banking system. *Economic Inquiry*, 25(3), 439–457.
- Sen, A. K. (1977). On weights and measures: Informational constraints on social welfare analysis. *Econometrica*, 45, 1539–1572.
- Shepsle, K., & Weingast, B. (1984). Political solutions to market problems. *American Political Science Review*, 78(2), 417–434.
- Simmons, R. (2011). *Beyond politics: The roots of government failure*. Oakland, CA: The Independent Institute.
- Somin, I. (2013). *Democracy and political ignorance: Why smaller government is smarter*. Palo Alto: Stanford University Press.
- Steiner, J. (2012). *Foundations of deliberative democracy: Empirical research and normative implications*. Cambridge: Cambridge University Press.
- Stigler, G. J. (1971). The theory of economic regulation. *Bell Journal of Economics and Management Science*, 2(1), 3–21.
- Stiglitz, J. (1991). The invisible hand and modern welfare economics. Working Paper No. 3641. National Bureau of Economic Research, Cambridge, MA. Retrieved from <http://www.nber.org/papers/w3641.pdf>.
- Stokey, E., & Zeckhauser, R. (1978). *A primer for policy analysis*. New York: W.W. Norton and Co.
- Stringham, E. (2001). Kaldor Hicks efficiency and the problem of central planning. *Quarterly Journal of Austrian Economics*, 4, 41–50.
- Strotz, R. H. (1958). Two propositions related to public goods. *Review of Economics and Statistics*, 40(4), 329–331.
- Taylor, M. (1982). *Community, anarchy, and liberty*. New York: Cambridge University Press.
- Tetlock, P. E. (2006). *Expert political judgment: How good is it? How can we know?*. Princeton: Princeton University Press.
- Thompson, D. F. (2008). Deliberative democratic theory and empirical political science. *Annual Review of Political Science*, 11, 497–520.
- Tideman, N., & Tullock, G. (1976). A new and superior process for making social choices. *Journal of Political Economy*, 84, 1145–1159.
- Trotsky, L. (1933). The soviet economy in danger. New York: Pamphlet Pioneer Publishers. Uncopyrighted. Originally published in *The Militant* during the months of November and December. Retrieved from <http://www.marxists.org/archive/trotsky/1932/10/sovecon.htm>.
- Truman, D. B. (1951). *The governmental process*. New York: Alfred A. Knopf.
- Tullock, G., Seldon, A., & Brady, G. (2002). *Government failure: A primer in public choice*. Washington, DC: Cato Institute.

- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
- von Mises, L. (1920). *Economic calculation in the socialist commonwealth*. 1990 reprint. Auburn: Ludwig von Mises Institute.
- Wagner, R. E. (1983). *Public finance: Revenues and expenditures in a democratic society*. Boston: Little, Brown.
- Wagner, R. E. (1985). Normative and positive foundations of tax reform. *Cato Journal*, 5, 385–399.
- Wagner, R. E. (1989). *To promote the general welfare: Market processes vs. political transfers*. San Francisco: Pacific Research Institute for Public Policy.
- Wedeman, A. (2012). *Double paradox: Rapid growth and rising corruption in China*. Ithaca, NY: Cornell University Press.
- Weimer, D. L., & Vining, A. R. (2011). *Policy analysis* (5th ed.). New York: Longman.
- White, B. (2014). *America's fiscal constitution: Its triumph and collapse*. New York: Public Affairs.
- Whittington, D., & MacRae, D, Jr. (1986). The issue of standing in cost-benefit analysis. *Journal of Policy Analysis and Management*, 5(4), 665–682.
- Wicksell, K. (1896). *Finanztheoretische Untersuchungen*. Jena: Gustav Fischer Verlag. (Partially reprinted as Wicksell, K. (1958). A new principle of just taxation. In R. A. Musgrave & A. T. Peacock (Eds.), *Classics in the theory of public finance* (pp. 72–118). London: Macmillan).
- Winston, C. (2006). *Government failure versus market failure: Microeconomics policy research and governmental performance*. Washington: The Brookings Institution.
- Wolf, C, Jr. (1979). A theory of non-market failure. *Journal of Law and Economics*, 22(1), 107–139.
- Wolf, C, Jr. (1993). *Markets or governments: Choosing between imperfect alternatives*. Cambridge: MIT Press.
- Zaller, J. (1992). *The nature and origins of mass opinion*. New York: Cambridge University Press.
- Zeckhauser, R. (1973). Voting systems, honest preferences and Pareto optimality. *American Political Science Review*, 67(3), 934–946.