Mechanism Choice

Jonathan B. Wiener
Duke University School of Law

Barak D. Richman
Duke University School of Law

Follow this and additional works at: https://lsr.nellco.org/duke_fs

Part of the Administrative Law Commons, International Law Commons, Policy Design, Analysis, and Evaluation Commons, Public Law and Legal Theory Commons, and the Public Policy Commons

Recommended Citation
https://lsr.nellco.org/duke_fs/178

This Article is brought to you for free and open access by the Duke Law School at NELLCO Legal Scholarship Repository. It has been accepted for inclusion in Duke Law School Faculty Scholarship Series by an authorized administrator of NELLCO Legal Scholarship Repository. For more information, please contact tracy.thompson@nellco.org.
Mechanism Choice
Jonathan B. Wiener*
Barak D. Richman**

Chapter outline
I. Introduction
II. Normative Mechanism Choice
   A. In General
   B. The Regulator’s Toolbox
   C. Criteria for Optimal Instrument Choice
III. Positive Mechanism Choice
   A. Optimal Choice
   B. Suboptimal Choice
   C. Public Attitudes
   D. Positive Choice at the International and Supranational Level
IV. Future Research

I. Introduction

Mechanism choice can generally be described as the selection of some way to structure rules for social behavior. Nobel Laureate Eric Maskin recently described a mechanism as “an institution, procedure, or game for determining outcomes” (Maskin 2008: 568).

In the realm of public law, mechanism choice is synonymous with “instrument choice” or policy design. The selection of the policy instrument can be as important to success or failure as the intended policy outcome. Good intentions or objectives are not enough: the choice of tools matters. A large and growing literature in instrument choice and mechanism design examines both the normative criteria for correcting market failures, matching optimal instruments to different types of problems, minimizing costs, and overcoming incomplete information; and also the positive political factors that may influence the actual selection of instruments, and the pattern of such choices across issue areas, governance systems, and time.

Public policy instruments are selected and designed by public bodies –legislatures, executive agencies, and courts – that are comprised of individuals with their own policy preferences, and that are subject to pressures from private interests through lobbying.

* Perkins Professor of Law, and Professor of Environmental Policy and of Public Policy Studies, Duke University; University Fellow, Resources for the Future (RFF).
** Professor of Law, Duke University, and Faculty Associate, Health Sector Management, Fuqua School of Business. We thank Dan Farber for helpful comments on a prior draft.
campaign contributions, and elections. Thus, it is no surprise to the student of public law that the mechanisms actually selected to implement public policy are not necessarily the ones that best pursue the public interest.

This chapter begins with a brief summary of normative mechanism choice, including the legal literature on instrument choice and the economics literature on mechanism design. It then moves to a more detailed discussion of positive mechanism choice, also called public choice, political economy, or positive politics. This positive literature explores how political institutions and pressures shape the selection of mechanisms to implement policy, notably when the selected instrument departs from the normative ideal. The positive study of mechanism choice not only informs how political processes shape policy outcomes, but also sheds useful insights into those processes themselves.

II. Normative Mechanism Choice

A. In General

Mechanism choice is a kind of social “engineering,” the task of designing optimal instruments to achieve social objectives (Maskin 2008: 567). But as the history of engineering demonstrates (and as mechanism design theory corroborates), there is no such thing as a perfect design. Every design involves choices among features that correspond to tradeoffs among competing objectives (Petroski 2004).

In economics, a core concern is mechanism design in the face of incomplete information. The significance of research addressing this concern is was reflected in the award of the 2007 Nobel Prize in economics to Eric Maskin, Roger Myerson, and Leonid Hurwicz. If preferences were known and outcomes were controllable, the designer could simply mandate actions and results. But if the designer has incomplete information, s/he needs some mechanism to achieve optimal results by eliciting from actors their private information (honest preference revelation) via a mechanism design that is “incentive compatible.” (Maskin 2008: 568, 571; Myerson 2008: 586-87). Neoclassical theory argues that markets are presumptively best at this information-elicitation task, so long as they are competitive, are free of significant externalities, and without information asymmetries (Hayek 1945; Maskin 2008: 572).

But if markets are flawed or incomplete, due to problems such as externalities, transaction costs, free riding, or incomplete information, then the theory of mechanism design seeks to choose the best form of government intervention to correct the market failure (Baliga & Maskin 2003). Such market failures in competition, externalities, and information, and the need to correct for those failures, are among the basic rationales for public law (Stiglitz 1989).

At the same time, government policies to correct market failures can pose their own problems, that is, government failures. Incomplete information about preferences and outcomes also confronts government decision makers. Eliciting voters’ true preferences
through voting may be elusive or even unattainable (classic studies include Arrow 1963; Gibbard 1973; and Satterthwaite 1975; more recent treatments include Barberá et al. 1997; Benoit 2000; Reny 2001). Foreseeing policy outcomes involves risk and uncertainty, so government incurs the costs of information and analysis as it tries to assess regulatory impacts in advance, and the costs of flaws in policy choices as they arise over time. Government failures in policy design can include policies that are excessively costly, that disproportionately benefit parochial constituencies at public expense (rent-seeking) or government bureaucracies themselves (internalities), that allocate burdens unfairly (inequity), or that induce adverse side effects (derived externalities or risk-risk tradeoffs) (Wolf 1993; Graham & Wiener 1995; Mueller 2003). The challenge for optimal policy is thus to minimize the sum of market failures and government failures.

One response, not uncontroversial, to mitigating the problem of government failure is to add a layer of supervisory oversight of the institutions developing policy interventions – that is, a system of regulatory review of those regulating the public. In the United States, this has been implemented through both judicial review (at least since the Administrative Procedure Act of 1946) and White House oversight of agency rulemaking (at least since the administration of Jimmy Carter). President Carter issued Executive Order 12044 and signed the Paperwork Reduction Act of 1980, creating the Office of Information and Regulatory Affairs (OIRA) within the Office of Management and Budget (OMB). President Reagan gave OIRA the power to apply benefit-cost analysis to regulatory review under his EO 12291, and President Clinton reaffirmed that approach in his EO 12866 (with notable improvements). (See generally Breyer 1993; Kagan 2001; Graham 2007. For a recent critique, see Revesz & Livermore 2008). In early 2009, President Obama called for a updating the system of regulatory review, and nominated legal scholar Cass Sunstein to lead that effort. In the European Union, similar oversight mechanisms have recently been established through the “Better Regulation” initiative that created the Impact Assessment Board (Lindseth et al. 2008; Wiener 2006; Renda 2006) and greater use of judicial review of administrative action (Lindseth et al. 2008; Alemanno 2008). Internationally, regulatory oversight occurs in institutions including, notably, the World Trade Organization (WTO) and its dispute resolution bodies.

Such oversight of regulatory design is a partial answer to the Roman poet Juvenal’s ageless question, “Quis custodiet ipsos custodes?” (“who will watch the watchers,” or “who will guard the guardians?”). It is only a partial answer because the additional oversight introduces another layer of costs (especially delay) and because the oversight body itself may require watching. To avoid an infinite regress of oversight bodies, a system of transparency and of checks and balances, including public participation through elections and perhaps litigation, can help those being watched (that is, the public, regulated entities, and advocacy groups) also help watch the watchers (the oversight bodies) of the watchers (the regulators) (Hurwicz 2008).

B. The Regulator’s Toolbox
Together, these contending concerns and institutions shape the choice of regulatory mechanisms. Policymakers and academics collectively have developed a large menu or “toolbox” of instruments to correct for market failures.

Perhaps the most common regulatory mechanism in historical practice has been the imposition of conduct standards, in which government specifies the behavior or technology that firms must adopt. Examples include administrative regulations mandating technological standards (such as pollution filters or types of fish nets). But the long history of research on mechanism design has developed many alternatives to conduct-specifying regulations, including taxes and subsidies (Pigou 1920; Weitzman 1974), property rights (Coase 1960; Demsetz 1967; Hardin 1968; Libecap 1989), tradable allowances or “cap and trade” (Crocker 1966, Dales 1968, Montgomery 1972; Tietenberg 1990; Wiener 1999a; Stavins 2003; Tietenberg 2006; Tietenberg 2007), tax-allowance hybrids (Roberts & Spence 1976), information disclosure (Hamilton 2005), process-enhancing procedures (Richman & Boerner 2006), regulatory “nudges” that guide cognitive heuristics (Thaler & Sunstein 2008), and bureaucratic control (Williamson 1999).

Each of these policy instruments has associated pros and cons; none is a costless intervention or perfectly efficient solution. The challenge is to develop criteria and evaluation methods that identify the policy instruments that would be comparatively superior to achieve particular social objectives under particular market conditions. We discuss these criteria further below.

To flesh out the menu of mechanism choices in the regulator’s toolbox a bit further, it includes at least the following six types of instruments:

1. Conduct rules

These instruments involve government commands to firms and individuals, prescribing specific technologies, practices, methods, or behaviors that must be employed or must be avoided. These instruments mandate conduct, rather than outcomes. (Related labels have included “command and control,” design standards, and technology-based regulation.) Examples include administrative regulations mandating prescriptive design standards for technology to employ (e.g. pollution filters, catalytic converters, types of boilers or fuels, standards in information and communications technology) or technology to avoid (e.g. types of fish nets).

The judicial version of conduct rules was reflected in the traditional negligence standard in civil tort liability, in which “due care” is defined as particular conduct or technology that private actors must adopt. This is still the approach in some areas of law, such as medical malpractice. The more modern version of the negligence standard defines due care as “reasonable” conduct, evaluated by a benefit-cost test that asks if the marginal costs of precaution are justified by
marginal benefits of precaution, but leaving flexibility to private actors in how best to undertake precautions.

2. Quantity/Property rules

These instruments limit the use of a resource or entry into a market by setting a quantity limit on such use or giving a property rightsholder the right to exclude. If the market failure is overuse of an open-access resource (such as a grazing commons, a fishery, or the atmosphere) that imposes external harms on others, then quantity/property instruments can remedy such failures by limiting access to the socially optimal amount.

Such limits might involve spatial parceling of a resource (“private property”), or limited permits to use a resource, or limits on entry into a market. They include property rights (such as rights to land, objects, minerals, airspace, and patents and copyrights), ambient standards (limiting the concentration or abundance of use or of pollution), emissions or effluent standards (limiting the amount of additional pollution that can be added in a period of time), and use or extraction licenses (such as for patents, copyrights, fishing and hunting licenses, airplane tickets, and broadcast spectrum licenses). Quantity use limits can be fixed (non-transferable) performance standards (e.g. fixed pollution limits, highway speed limits, or airplane tickets); or they can be transferable among users in a system of tradable permits or tradable allowances (also called “cap and trade”; e.g. pollution emissions allowances, fisheries catch quotas, broadcast spectrum licenses, and aviation landing slots).

Tradable allowances can be issued by the government to private actors for free (such as “grandfathered” in proportion to historical users), or they can be sold or auctioned by the government to raise revenues and help offset the cost of the regulatory restriction (Goulder et al. 1997; Parry 1995), or some mixture of free and auctioned allowances can be designed. A significant focus of the economics of mechanism design involves the design of auctions to try to elicit honest preference revelation by bidders (e.g. Athey & Levin 2001).

Another key consideration in cap and trade policies, constraining costs, is discussed in the next section regarding the comparison between taxes and tradable allowances.

Critics of incentive-based instruments such as tradable allowance mechanisms have worried that these instruments may lead to “hotspots” of concentrated pollution (e.g. if one firm purchases many allowances), may soften the stigma associated with polluting by “licensing the right to pollute,” and may have high administrative costs. Advocates of tradable allowances respond that hotspots are unlikely if the aggregate cap is stringent, that hotspots depend on local emissions causing local harms (which is not the case, for example, with most greenhouse gases), and that hotspots can be limited through the design of the allowance.
trading system to restrict high ambient levels. They add that far from licensing the right to pollute, incentive instruments like taxes and tradable allowances make polluters pay for every unit of pollution (whereas command technology standards and fixed performance standards allow residual pollution to be emitted for free), and do better at motivating innovation, thus reducing pollution more effectively. And advocates argue that the administrative costs of incentive instruments (chiefly monitoring emissions and collecting taxes or tracking allowance trades) are no higher than for command technology standards (developing engineering analyses of best technologies and litigating those choices), and in any event are dwarfed by the social cost savings delivered by incentive instruments.

3. Price/Liability rules

These instruments limit use of a resource not by limiting the quantity or spatial terms of use, but by setting a price on use. If the market failure is overuse of the resource, then price/liability instruments can remedy that problem by setting a price for use that equals the social marginal harm of use. Such price instruments include taxes on use, subsidies for abatement, and strict liability in tort.

A classic contrast pits between taxes versus tradable allowances. Under a quantity-based system of tradable allowances, government sets the aggregate quantity of use, and the market then determines the price of an allowance. By contrast, under a tax, government sets the price for use, and the market then determines the quantity of use. As Weitzman (1974) showed, if the government knows the true marginal cost of abatement, then it can set either the quantity or the price and achieve the same result. But under uncertainty about the true marginal cost curve, the quantity instrument risks fluctuations in the cost of use, and the price instrument risks fluctuations in the quantity of use. Which instrument is preferable depends (all other things equal) on whether one is more concerned about the costs escalating or the use escalating (that is, on the relative slopes of the marginal cost and marginal benefit curves).

If cost escalation is a key concern (favoring a tax), but a quantity limit on total use is desirable or taken as given, then a quantity-based tradable allowance system can be designed to constrain costs in several ways: setting the stringency of the quantity cap appropriately (not too tight); enabling firms to enjoy “how” flexibility in methods of abatement, “where” flexibility in the location of abatement across users, and “when” flexibility through banking and borrowing over time; ensuring a broad and competitive market for allowances; and setting price ceilings (“safety valves”) at which extra allowances will be sold (i.e., by creating a tax-allowance hybrid instrument) (Roberts & Spence 1976), perhaps coupled with price floors that maintain at least some incentive for innovation and also help constrain price volatility in both directions (Burtraw et al. 2009).

Taxes raise revenues; as noted above, tradable quantity allowances can also raise revenues by being sold or auctioned.
A key difference between taxes and tradable allowances is their ability to engage participation by users, especially where the voting rule for adoption of the policy requires consent to be bound (as it does at the international level). This feature is discussed further in Part III below.

Subsidies also act as price instruments, by offering users a payment to abate (hence a foregone payment from failure to abate), or by supporting basic research and development (R&D) in which private firms would not invest adequately. Subsidies for abatement can reduce resource use at the margin, but unlike taxes or tradable allowances, they may also reduce the average cost of operating in the subsidized industry and thereby attract greater investment that perversely increases output in that sector (Oates 1990). Subsidies for basic R&D can help overcome the market failure in incentives for innovation, but unless the price of the new technologies can be driven down below the price of conventional technologies, the R&D subsidy will need to be paired with some instrument (such as a tax or tradable allowance system) to correct the market failure in resource use by internalizing the external costs of the conventional technologies and thus to encourage diffusion and adoption of the new technologies (Jaffe et al. 2005; Pezzey et al. 2008). And government subsidies for the deployment of particular technologies run the risk of picking a losing technology.

4. Information disclosure rules.

If the market failure involves asymmetric information, or obstacles to bargaining because of incomplete information, then an information disclosure instrument can help remedy the problem and facilitate more efficient market transactions. Information, transparency, labeling, and related instruments are widespread. For example, information disclosure is currently required in sales of securities, sales and leases of real estate, loans, sales of and consumer products (such as appliances, motor vehicles, food, prescription and over-the-counter drugs). Some of these disclosure requirements are highly detailed. Additional versions of information instruments include Environmental Impact Statements, Regulatory Impact Assessments, the OSHA Hazard Communication Standard, and the Toxics Release Inventory (Hamilton 2005). The function of regulatory oversight itself involves disclosure of decisionmaking rationales and evidence – by the agency in its notice and comment rulemaking process, by OIRA, and by the courts through judicial proceedings.

5. Government ownership

If regulation of market transactions is not sufficient, government can acquire ownership of a resource or enterprise. Government may acquire ownership through negotiated purchase, or through the power of eminent domain to take property (and pay compensation). Public parks and lands, public works and utilities, and various services (including the Postal Service, Amtrak, and air traffic
control) are currently operated by the public sector in the USA; various sectors are government-owned in other countries. Market failures in financial markets have recently led to increased government ownership of banks, lenders, and insurers. Government ownership may enable wider public access to resources, but it may also replace market incentives to invest in conservation of resources with bureaucratic procedures that let resource values dissipate or become captured by private users. Government acquisition or compensation in response to private losses can create an expectation of relief that generates “moral hazard,” excessive risk-taking by the insured (Akerlof & Romer 1994).

6. Private ordering

Industry custom, trade association codes of conduct, professional codes, and group property management systems are all examples of mechanisms that private actors may adopt to regulate conduct. Where actors face low transaction costs of developing, agreeing on, monitoring and enforcing such codes, they can be attractive mechanisms (Ostrom 1990; Ellickson 1990). The question of instrument choice remains: private codes of conduct and social norms typically take one of the mechanism design options described above.

The relationship between government policy and private ordering may vary. In some cases, government adoption of regulatory policy may supersede or crowd out private codes and norms. In these cases, private codes are sometimes meant to fend off government regulation; an example is the history of self-regulation by the movie industry meant to avoid government-imposed content ratings or restrictions. In other cases, government policy may serve an expressive function that helps establish and strengthen social norms.

Other instruments are also possible. For example, Dijkstra (1999) classifies education and training as a regulatory instrument.

C. Criteria for Optimal Instrument Choice

As a general matter, there is no single universally optimal mechanism or policy instrument, because the choice is pragmatic, involves several considerations, and is largely driven by the attributes of the specific market failure (Breyer 1982; Baumol & Oates 1988). The economics and legal literatures have offered useful, though not exhaustive, theories for optimal mechanism choice, the process of selecting the optimal instrument to correct a particular problem.

Perhaps the most common method of normative mechanism selection, at least as advocated by economists and invoked by OMB regulatory review, is benefit-cost analysis (Stokey & Zeckhauser 1978; Munger 2000). The debate over benefit-cost analysis is extensive; it goes beyond the scope of this chapter, because benefit-cost analysis seeking to answer how much regulation is desirable, whereas mechanism choice typically seeks to
select the best policy instrument to achieve a given degree or objective of regulation (with that goal having been given by the legislature). Thus, mechanism choice in regulatory policy often involves cost-effectiveness analysis rather than full benefit-cost analysis.

In either case, a full analysis of benefits or effectiveness would include both targeted and ancillary consequences (Graham & Wiener 1995; Revesz & Livermore 2008); and a full analysis of costs would include compliance costs, general equilibrium social costs (Hazilla & Kopp 1990), and administrative costs (Ackerman & Stewart 1988). A full analysis of benefits and costs would also include dynamic considerations, such as the degree to which an instrument spurs innovation (Jaffe, Newell & Stavins, 2003) or encourages absorbing new information and adapting to changing circumstances. Distributional criteria, such as fairness or justice, can also inform a benefit-cost comparison of alternative mechanisms and policy outcomes. Benefit-cost analysis can aim to maximize aggregate net benefits (the Kaldor-Hicks criterion), or can be applied to a more stringent test that would maximize net benefits while also ensuring that no individual is made worse off (Pareto-improving), by combining the optimal choice with compensatory side payments from its winners to its losers. Broader versions of benefit-cost analysis embrace all of these considerations, and both qualitative as well as quantitative analysis, in a “cognitive” or “warm” approach to pragmatic decision making that sees benefit-cost analysis as a tool, not a rule, to inform sound judgment by responsible officials (Sunstein 2000; Farber 1999; Wiener 2006).

On these criteria, many economists argue that for most externality problems (such as pollution), incentive-based instruments such as taxes and tradable allowance (cap and trade) systems are superior to conduct rules that specify behavior or technology. Incentive-based instruments achieve results at lower cost, through “how,” “where” and “when” flexibility (as described above). Moreover, incentive-based instruments stimulate continuous innovation, while command technology standards may stagnate innovation once the government-selected technology has been adopted (Jaffé, Newell & Stavins 2003). (For surveys of instrument choice comparisons across diverse criteria, see Baumol & Oates 1988; Wiener 1999a.)

Law and economics scholars have developed additional theories to guide optimal instrument choice. For example, Calabresi & Melamed (1972) developed a template for when property rules are superior to liability rules, focusing on transaction costs and judicial errors. Weitzman (1974) identified tradeoffs between price instruments (such as taxes) and quantity instruments (such as tradable permits) when the decision maker is uncertain about costs and benefits. Breyer (1982) described how matching the type of policy instrument to different types of problems can help solve market failures. Stewart (1986) emphasized the basic choice in public law between replacing markets with bureaucratic controls, versus “reconstituting” markets through incentive-based regulations that exact performance while preserving substantial flexibility to market actors. Baumol & Oates (1988) synthesized the economics literature on optimal instrument choice for environmental protection. Williamson (1999) employed new
institutional economics to suggest when economic and political transactions require administrative management, including regulation or public agencies.

In sum, normative evaluations of policy instruments rest on a wealth of valuable theoretical criteria. Yet actual policy choices often depart from normative guidance. It is therefore important to understand the positive political processes that select actual policy instruments.

III. Positive Mechanism Choice

The main question posed by the robust political economy literature on positive mechanism choice is whether positive political forces actually choose the instruments deemed optimal by normative analysis, or instead depart from the optimal choice to employ some other (suboptimal) instrument. If positive politics do affect the process of mechanism choice, then political forces predictably constrain policy options or raise the costs of certain policy instruments. Accordingly, understanding these political constraints (which may be impossible or too costly to change) should inform the normative instrument choice analysis. As James Buchanan advised, one should understand the political system before prescribing normative instrument choice (Buchanan 1987).

Views of the political process vary widely, and instrument choice predictions depend heavily on the underlying model of the political system. Here we focus on instrument choice in the United States given its particular set of political institutions and players. In this section, we begin with optimistic models in which positive instrument choice emulates optimal normative choice. Then we move to more pessimistic models in which positive instrument choice departs from optimal normative choice. At each step, we assess the relevant empirical evidence and consider complications to positive models; we also address the politics of both administrative regulation and property rights. Further, we examine the roles of public attitudes that might produce results that are more mixed and nuanced than either the highly optimistic or highly pessimistic models. Last, after focusing on positive instrument choice in the United States, we comment on positive instrument choice in international treaties and in the European Union, in order to offer some comparative perspective on the role of basic institutional structures (such as voting rules) in shaping positive choices.

A. Optimal Choice

If all individuals’ interests are effectively represented in the political system, positive mechanism choice may emulate the normative criteria for maximizing net benefits (and perhaps even compensating losers). Lowi (1979) advanced a vision of pluralist democracy in which all citizens’ interests are expressed through organized interest groups, and government decisions thus optimally aggregate and reconcile citizens’
preferences. Aidt (1998) arrived at the same socially optimal outcome through a more formal model, in which every individual is effectively represented by an interest group.

Even if all individuals are not effectively or equally represented by interest groups, optimal choice is still possible. Becker (1983, 1985) showed that if the political process is a frictionless competition among interest groups, organized interests will bargain for their desired policy outcomes and, in order to secure adoption of their preferred programs, will propose policy mechanisms that reduce the costs (prices) of achieving their objectives. This competitive bidding process dissipates the interest groups’ rents, and socially optimal policies result. Similar intuitions underlie Wittman (1995) and Cowen (1994).

Many observers doubt that such optimistic models depict the reality of American politics. Olson (1971) and others had argued that disorganized or diffuse interests would be unequipped to bargain effectively against concentrated interest groups. North (1990) argued that Becker's model is unrealistic because significant transaction costs of expressing political voice make political influence difficult and unevenly available (see also Mashaw 1997). Hahn (1989: 173-75) observed that, at least through 1989, Becker's optimism had not been borne out empirically in environmental regulation, where suboptimal instruments had often dominated the field (although since 1989, incentive-based instruments for environmental protection have become more widely adopted, see Stewart 2001, Oates & Portney 2003, and further discussion below).

Recently, Croley (2008) has argued that optimistic pluralist models do depict the reality of actual American policy making, provided that the instrument choice and policy process are adequately shielded from the interest group biases of concern to Olson, North, Mashaw, and others. For example, Croley finds that separating regulatory agencies from the Congress, and requiring administrative procedures to transparently and honestly execute benefit-cost analysis, can ensure that regulation achieves public interest benefits. From a different institutional vantage point, Demsetz (1967) argued that bottom-up norms and judicial protections (rather than top-down legislative and administrative safeguards) can ensure the emergence of optimal policy instruments. Demsetz depicted the evolution of property rights as a process driven by the quest for net benefits, in which open-access resources become regulated by the establishment of property rights to exclude, whenever the benefits of internalizing externalities rise to exceed the costs of establishing and enforcing property rights.

B. Suboptimal Choice

1. The Olson Model

A more popular, and less sanguine, view of the political process was crystallized by Mancur Olson’s classic, The Logic of Collective Action (1971), first published in 1965. In Olson’s framework, political voice itself is a collective good that is restrained by transaction costs and is thus subject to the difficulty of mobilizing collective action. In
In contrast to optimistic models such as Aidt (1998), in Olson’s framework, public preferences are incompletely represented by interest groups. Interest groups have little incentive to lobby for nonexcludable benefits that will be reaped by others, so they advocate only or at least mainly for excludable benefits they can appropriate to their members. Members and groups prefer to free ride on nonexcludable benefits generated by others, so they underinvest in promoting legislation that would yield diffuse general interest benefits (and they underinvest in resisting legislation that would impose diffuse costs). The result is that concentrated interest groups are able to obtain their own benefits and extract rents at a cost to diffusely held interests.

The public choice literature contains numerous applications of Olson’s interest group theory to explain legislation and agency policies that benefit industry at the expense of the general public. Early work in political science accused the railroads (Kolko 1960) and military suppliers (Gerschenkron 1945) of distorting the democratic process to obtain legislation serving their narrow interests. A key concern in the 1960s was the problem of agency capture by regulated industry. Earlier economic models of democracy had been developed by Bentley (1908) and Downs (1957). What Olson (1971) added was an account of incentives, appropriability (excludability) and free riding in the political domain that helped formalize and generalize the analysis, showing that political voice itself can have public goods characteristics and thus go underprovided unless it is paired with private rewards. Stigler (1971, 1975) and Peltzman (1976) took an analogous approach to explain economic regulation that favors concentrated interests over diffuse public interests, such as licensing regulations that restrict entry to markets and raise prices to consumers. Buchanan & Tullock (1975) applied this theory to explain the emergence of policy instruments that reflect incumbent industry members’ preference for command-and-control standards rather than taxes or auctioned allowances, because taxes or auctioned allowances force industry to pay for their residual uncontrolled emissions, whereas command technology standards do not; thus, they argued, the prevalence of command technology standards in 1970s-era environmental law can be attributed to industry’s concentrated interest and political voice.

Wilson (1984) and Eskridge (1988) emphasized that Olson’s theory predicts more than just overprovision of narrow rent-seeking legislation that provides concentrated benefits to a few while imposing diffuse costs on the public. It also predicts underprovision of general interest legislation that provides diffuse benefits while imposing costs on concentrated interests. Moreover, Olson’s theory forecasts ambiguous results in cases of diffuse benefits and diffuse costs, and delegation to another body (such as a regulatory agency) in cases of concentrated benefits and concentrated costs.

2. Theory Meets Evidence

One significant challenge to Olson’s interest group theory is the empirical reality that much legislation has been enacted that yields diffuse general benefits while imposing costs on industry, such as environmental law and antitrust law, and the simultaneous reality that some economic regulation protecting concentrated industry interests has been undone. There are at least two possibilities to explain this reality: either the Olson
theory fails to describe actual politics, or seemingly general-interest legislation is not as
general-benefit oriented as it seems (and is actually driven by parochial rent-seeking).

On the first possibility, many critics challenge the core of the Olson approach, arguing
that the theory is too simplistic to capture the reality of the emergence of general benefits
legislation. For example, Farber (1992: 60) observes: “The Olson paradigm appears to
have a straightforward implication for environmental legislation: there should not be
any.... [T]he two basic predictions [of Olson's model] are that environmental groups will
not organize effectively and that environmental statutes will not be passed.” Revesz
(1997: 561) says: “it is difficult to explain, in public choice terms, why there would be
any environmental regulation at all.” Schuck (1997: 566) considers the emergence of
modern environmental law to be a “major predictive error of the new public choice
theorists.” Oates & Portney (2003) comment: “In fact, from [the Olson] perspective,
what does seem surprising is the extent to which environmental advocacy groups have
mobilized their constituencies so effectively. The benefits from programs to improve air
quality on a national scale, for example, would appear to represent an Olsonian large-
group case, where it would be extremely difficult to organize environmental interests. But
in seeming contradiction to the prediction of the theory, environmental groups have
proved to be a very powerful force in the policy arena.” These criticisms join earlier
observations that environmental law was an especially difficult case to explain with
standard public choice theory (Posner 1974; Elliott, Ackerman & Millian 1985). The
strong form of the Olson theory thus seems unable to explain, or at least seems to
underpredict, the adoption of general-interest legislation such as environmental and
consumer protection. Perhaps the inference to draw is that the general interest legislation
that arises is the product of episodic abnormal politics, as we discuss below; and that
Olson’s theory of normal politics favoring concentrated interests implies that,
normatively, we would be even better off if we could have even more general interest
legislation than we get in reality.

Similarly, much of the history of US economic regulation and deregulation in the 1970s
cuts against the Olson theory. Whereas Olson, Wilson and Stigler predicted the lack of
general interest legislation and the prevalence of legislation favoring concentrated
interests, the 1970s exhibited the opposite: the rise of (ostensibly) general interest
environmental and consumer regulation and the nearly simultaneous deregulation of rules
shielding incumbent industry interests against competition. Horwitz (1989) recounts the
repeal or reform of the very economic regulations in banking, transportation, and
communications that the Olson/Wilson/Stigler theory predicted would be most dear to
industry because those rules had delivered benefits to concentrated industry and had
imposed diffuse costs on consumers. McCraw (1984) documents a similar political story
when certain deregulation efforts removed benefits from concentrated interests and
bestowed general benefits on consumers. To be sure, some other economic regulations
have remained in place, including government relief policies that may benefit risk-taking
banks over the general public (Akerlof & Romer 1994); and the partial deregulation of
financial services in the last decade may reflect intra-industry rivalry and regulatory
capture (Hardy 2006). But the juxtaposition of sustained environmental regulation and
significant economic deregulation still presents a strong challenge to the basic Olson model.

More generally, critics have doubted that focusing on industry’s parochial gains offers an adequate account of observed political outcomes, in part because so many other variables are also in play. This suggests that Olson’s model suffers from foundational theoretical problems. Attributing observed political choices solely to industry lobbying may be a fallacy of “post hoc, ergo propter hoc” (Mashaw 1997: 203). Breyer (1982: 388 n.38) remarks: “Interest group theories, as causal explanations of either the historical origins of regulation or the actions of regulators, suffer several drawbacks. Where they are limited to producers [i.e., industry], they are often inaccurate. They cannot fully explain environmental, health, [and] safety regulation…. If the theory is expanded beyond producers, it risks becoming nonpredictive and nonexplanatory. All regulatory rules and programs benefit some group or other.” (See also Breyer 1984: 282, quipping that interest group theory is uninteresting where it is true, and untrue where it is interesting.) Noll (1989: 1277) worries that “the evidence is still far from fully conclusive. . . . [There] is the lurking danger of tautology, i.e., of attributing causality to an inevitable consequence of any public policy action. It is impossible to imagine that regulation could be imposed without redistributing income. Hence, a look for winners in the process—and organizations that represent them—is virtually certain to succeed. Until fundamental measurement problems about stakes, power, and gains are overcome, analysts will not be able fully to predict and to explain the details of regulatory policy.”

Alternatively, there is substantial evidence of the second possible reconciliation of Olson with the evidence: that the seemingly general-interest legislation that has been enacted is actually the result of the rent-seeking pursuit of parochial interests. Many studies observe that economic interests routinely use purportedly general interest regulation to raise their rivals’ costs (on the general theory, see Salop & Scheffman (1983); Salop et al. (1984); Tollison (1991)). For example, Ackerman & Hassler (1981) found that the sulfur dioxide air pollution regulations in the 1977 Clean Air Act required command technology standards (scrubbers) that favored high-sulfur eastern coal over low-sulfur western coal. Incentive-based instruments such as taxes or allowance trading would have removed this parochial advantage. In their view, this distortion of instrument choice was a victory for eastern coal, not for clean air. Pashigian (1985) investigated the same statute and found that the “prevention of significant deterioration” (PSD) provision enacted in 1977, ostensibly to preserve clean air in areas with already good air quality, was adopted by the votes of northern rustbelt states with poor air quality over the dissenting votes of southern and southwestern sunbelt states with good air quality. He inferred that PSD was imposed (through majority vote) by the rustbelt states in order to suppress economic growth in, and industry relocation to, their cleaner rivals in the sunbelt. Along similar lines, Bartel & Thomas (1987) found evidence that OSHA and EPA regulations protect large firms and rust-belt firms against smaller firms and sun-belt firms. And in a different topical domain, McChesney & Shugart (1995) argue that antitrust law, while ostensibly promoting the general interest, has been captured to deliver special interest protections to industry segments.
Another intra-industry strategy is to seek regulations that protect existing incumbents against the threat of new entrants. This strategy arguably accounts in part for the “new/old” distinctions in many regulatory instruments. New entrants tend not to exist or are not yet well organized politically, so, following Olson, they lose out in the political process to well-organized incumbents with concentrated interests. For example, Gruenspecht (1982) explained the preference of the current automobile manufacturing industry for more stringent restrictions on new sources of vehicle emissions. A similar strategy by incumbents could help explain the tighter regulations imposed on new stationary sources of air pollution, including “new source review” and more stringent technology-based controls at new or modified sources. Such a pattern of incumbent protection might also explain allowance trading systems with “free allocation” to current users, obliging new entrants to pay to receive permits or to purchase them from incumbents.

Evidence of rent-seeking in the development and design of administrative regulation finds a parallel in studies that view the development of property rights as a rent-seeking strategy. Whereas Demsetz (1967), as noted above, proposed an optimistic story of maximizing net benefits to explain the evolution of property rights to regulate resources, other scholars see a more coercive element of interest group politics in the evolutionary story. In these versions, the definition and allocation of property rights is subject to political pressures. Levmore (2002) tells a second evolutionary story that focuses on politics and rent-seeking, including examples in the development of intellectual property rights. Wyman (2005) offers a similar account of politics and state power regarding the evolution of property rights in fisheries. Libecap (2007) observes that the first-possession or first-appropriation methods of assigning property rights that are often applied to oil and gas, water, radio spectrum, pollutant emissions, and fishery ITQs, – rather than auctions or uniform allocations, which often are more efficient – tend to arise to protect incumbents. These political explanations of supposedly general-interest property rules are consistent with the Olson theory.

Thus, once competing intra-industry interests are appreciated, the evidence for and against the Olson theory of interest group politics presents a mixed picture: underprediction of general interest legislation (much enacted), overprediction of parochial economic regulation (much deregulated), and yet potentially accurate prediction of intra-industry rivalry embedded in both sets of policies.

3. More Complex Models

The models of positive choice discussed thus far have focused on the demand for regulation and particular regulatory instruments, and they have focused especially on organized private interest groups as political actors that demand regulation. More complex models of positive politics look at a wider set of actors on the demand side, such as advocacy groups and political entrepreneurs, and also give attention to the supply side of legislation, such as legislators and regulators.
Accounting for the role of advocacy groups that lobby for public interest regulation is not straightforward. On Olson’s account, such groups should have difficulty organizing effectively because they generate diffuse nonexcludable benefits. The inference might be that to the degree we observe advocacy from such groups, that effort is underprovided (compared to the social optimum) because free riding limits contributions to such groups. Moreover, much of the relevant legislation was enacted when advocacy groups favoring such legislation were not yet powerful; in the case of environmental legislation enacted in 1969-72, many of the major environmental advocacy groups we know today did not yet exist or were fledglings (Farber 1992).

Additionally, advocacy groups that purport to mobilize on behalf of the public interest regularly pursue policy objectives that depart from utilitarian normative criteria. Oates & Portney (2003), for example, observe that environmental groups historically opposed incentive-based regulation partly on the ground that these instruments might create “hotspots” that concentrated the adverse affects of pollution (as discussed above), and partly because they distrusted markets and favored bureaucratic control. These advocates instead supported command technology standards rather than taxes or auctioned allowances, even though most economists viewed command technology standards to be inferior to incentive-based regulations. Thus many advocacy groups joined with industry interests who similarly sought to avoid paying for inframarginal emissions (Buchanan & Tullock, 1975), and the combined result was the positive political choice to adopt suboptimal instruments. Oates & Portney also note, however, that in recent years, incentive-based instruments have been more widely adopted, raising the questions whether these interest groups have either lost influence or have changed their preferences regarding instrument choice.

An intriguing hypothesis by Yandle (1989) is that many organized groups achieve their policy objectives only after joining forces with other organized groups, including those with whom they share little in common or even oppose. Based on the story of the odd political alliance between Baptists and bootleggers to impose Sunday closing laws on liquor stores – in which Baptists sought to ban (at least official) liquor sales on the Sabbath, and bootleggers sought to shut down their (legal) rivals for one day a week – Yandle suggests that both environmental advocacy groups and industry achieve their desired policy outcomes when they ally together. Such cooperation thus generates laws that (purport to) limit pollution and while also favoring one industry segment over another. This kind of strategy may explain the air pollution laws discussed above that favored one industry region over another, the choice of command technology standards rather than taxes or (auctioned) allowances, and the new/old distinction favoring incumbents. And it further suggests that while advocacy groups may be influential in raising issues on the political agenda to spur adoption of laws, their influence is dependent on joining forces with industry which in turn works to torque the detailed content of those laws to favor incumbent industry interests (Wiener 1999b).

In addition to advocacy groups, the policymaking process also includes what are commonly called “policy entrepreneurs” (or political entrepreneurs) who capitalize on diffuse and even inchoate public preferences to propose and develop support for new
policy designs. Such actors were not emphasized in Olson (1971), perhaps explaining some of his failure to predict the coming wave of consumer and environmental legislation and of economic deregulation, but the role of policy entrepreneurs was highlighted by Wilson (1984: 370-71), Eskridge (1988: 285), Arnold (1990), and Schroeder (1998). Policy entrepreneurs may serve some or all of at least five kinds of functions. First, they lower the transaction costs of political awareness and voice through campaigns to collect and disseminate information. Similarly, policy entrepreneurs monitor policy adoption and implementation by institutions—monitoring that interest groups are not sufficiently incentivized to do themselves—by publicizing results and checking lapses in performance. Second, they construct a narrative or vision of social ills (often including blame) and of the good society (including the payoffs of action) that help overcome free riding and mobilize diffuse—or rival—interest groups to advance policy objectives. Third, rather than merely reflecting existing preferences, policy entrepreneurs anticipate as-yet unexpressed or unformed future preferences of key constituencies—what Arnold (1990: 10) calls the “potential preferences” of voters—and appeal to those forward-moving preferences. They may go further and shape new preferences through their narrative and vision. Fourth, they help devise new instruments that reduce costs and improve performance. This function is illustrated in the role of Environmental Defense Fund in designing and advocating market-based allowance trading systems to control air pollution (Keohane et al. 1998: 354; Oates & Portney 2003), and by the roles of regulatory reformers such as Alfred Kahn (at the Civil Aeronautics Board) and Stephen Breyer (then chief counsel of the Senate Judiciary Committee) in the 1970s in deregulating industries such as aviation, banking, and communications (McCraw 1984). In this capacity, the policy entrepreneur is more innovative and less political than the interest group competing to reduce costs envisioned by Becker (1983, 1985). The entrepreneur is not just extracting rents or pursuing policy outcomes but is instead a creative modernizer who invents a new policy instrument, something akin to a Schumpeterian innovation in the public sector (Wiener 2001: 1350-52). Policy entrepreneurship can arise from interest groups on the demand side of regulation, or from politicians and agencies on the supply side of regulation.

More complicated models give greater attention to the supply side of producing regulatory policy. These models suggest that Congress or its members may be “trolling” for interest groups with votes and campaign dollars, in order to “sell” regulatory legislation to these “buyers” by legislating benefits. Congress can also sell benefits by exercising power over regulatory agencies to favor targeted interests. Sometimes this transaction is complicated by rival bidders. Wilson (1984), as noted above, hypothesized that Congressional delegation to an agency is most likely when the issue is in conflict between concentrated interest groups on both the benefits and costs side; he suggested that Congress would likely write a statute expressing lofty ambitions but leaving the difficult tradeoffs to the agency. Fiorina (1983) developed a model that predicts when Congress will delegate to courts versus agencies, and Schwartz, Spiller & Urbiztondo (1994) similarly suggested when Congress might delegate policymaking to courts through articulating statutory intent.
These delegation decisions—and the models that predict delegation—became especially significant following the Supreme Court’s decision in *Chevron v. NRDC*, 467 U.S. 837 (1984), which not only upheld the agency’s authority to supply a reasonable interpretation of an ambiguous statute, but in particular, upheld EPA’s interpretation that the word “source” in the Clean Air Act authorized a form of emissions allowance trading (EPA’s bubble and netting policy) when Congress had left the statute’s wording unclear. This case illustrates the role of the agency as policy entrepreneur, and suggests a role for the courts in interpreting statutes (or deferring to agency interpretations) that promotes normatively optimal instrument choices and checks the distortionary effects of positive politics on instrument choice.

A related literature speaks to how Congress designs agencies in crafting and securing its preferred policy outcomes. Both Bawn (1995) and Epstein & Ohalloran (1994) suggest that congressional politics can explain certain structural elements of agency design. McNollGast (1987) offers a model explaining how Congress, through the Administrative Procedure Act, established a rulemaking procedure to ensure that organized interest groups maintain control over agency policymaking and can alert Congress to agency actions that threaten the interest groups’ deal with Congress. And McCubbins & Schwartz (1984) and Spiller (1990) illustrate how Congress monitors agencies after delegating to them substantial policymaking responsibility. This literature explains how Congress can respond to political demands—and thus reap the benefits from catering to organized interest groups—while still delegating policymaking responsibilities to agencies. On the other hand, Seidenfeld (1992) and Spence & Cross (2000) argue that delegation to agencies can reduce the public choice distortions of Congressional legislation, and Croley (2008) argues that once such power is delegated, the APA and White House oversight help shield agencies from Congressional micromanagement—and thus from interest group demands—while ensuring both transparency and attention to social net benefits in agency rulemaking.

A synthesis of these diverse approaches is developed by Keohane, Revesz & Stavins (KRS) (1998), which proposes a model with many of the complicating factors discussed above. KRS offer a supply and demand model to explain both environmental regulatory instrument choices and when those choices depart from instruments that are normatively preferred by economists. In their model, Congress is the supplier of regulation to a set of demanders that includes both the voting public and interest groups. On the demand side, KRS account for transactions costs of political voice and free riding (in line with Olson 1971 and North 1990, and unlike Aidt 1998) so that interest groups do not fully represent public interests but rather promote parochial special interests. Like Yandle (1989), they allow for multiple interest groups which may compete or enter into coalition alliances. Like Buchanan & Tullock (1975), they observe that industry prefers command technology standards rather than taxes or auctioned allowances (because the latter force firms to pay for residual uncontrolled emissions) and that incumbents prefer tighter controls on new entrants. And like Oates & Portney (2003), they observe that environmental groups have also historically preferred command technology standards, although KRS suggest that niche-seeking competition among environmental groups may
diversify these views toward support for incentive-based instruments (while calling for more empirical research on what advocacy groups actually seek).

KRS then add attention to the supply side: the provision of legislation by Congress. In their model, legislators seek re-election, which depends on both votes and interest group support, but because of a rising marginal cost curve for legislation, interest group demands cannot be fully met and instead there will be an equilibrium of demand and supply in the market for legislation. One consequence, they suggest, is that Congress may prefer command technology standards when it is more costly to learn about newer incentive-based instruments rather than continue employing familiar traditional instruments. KRS observe that Congress might also prefer command standards if they are more effective at dictating the distributional impacts of regulation on constituencies. (Command standards may more easily dictate the distribution of environmental impacts than do taxes or allowance trading [which afford “where” locational flexibility in abatement by sources, and hence reduce costs but may raise concerns about hotspots], but allowance trading systems may more easily dictate the distribution of compliance costs through the allocation of allowances among sources.) But like Becker (1983, 1985), KRS also find that interest groups have incentives to compete over time to offer Congress less-costly mechanisms to achieve the groups’ desired goals, because Congress cares about reducing costs in order to attract votes from other affected constituencies. Consequently, when there is no embedded interest group deal securing existing instruments, incentive-based instruments are more likely to be adopted to address new issues.

KRS argue that these factors combine to help explain the observed shift over time toward adoption of cap-and-trade allowance systems with allowances allocated for free to historical users (“grandfathering,” preferred by industry), tight caps (preferred by environmental groups), fewer total social costs to voters compared to alternative instruments, and learning over time by legislators about the success of these instruments. Incentive instruments such as cap-and-trade will also gain a more receptive audience among legislators when distributional impacts are favorable: when concerns about hotspots are attenuated (such as for greenhouse gases, which mix globally in the atmosphere) and when allowance allocations and/or revenues from taxes and allowance sales can be used to soften the distributional burden of the regulatory policy (DeShazo & Freeman 2007).

One illustration of this shift in instrument choice can be seen in the policies adopted to control acid rain from sulfur dioxide emissions in 1977 and 1990. Whereas the 1977 Clean Air Act effectively mandated command technology standards (scrubbers) at coal-fired power plants, a choice that favored eastern high-sulfur coal over western low-sulfur coal (Ackerman & Hassler 1981), the 1990 Clean Air Act amendments imposed quantity trading (tradable allowances or cap-and-trade), a choice that encouraged coal-fired power plants to switch from high-sulfur to low-sulfur coal. What explains this shift in instrument choice, and the reversal of eastern coal’s earlier interest group victory? Joskow & Schmalensee (1998) use a hybrid political economy explanation to account for the change in regulatory regime. First, consistent with Becker (1983), they argue that the high social cost of the 1977 command technology instrument led the regulated parties and
Congress to seek a less-costly alternative system in 1990. Additional factors were also at work. Second, the physical features of the acid rain problem—which prompted the 1990 Amendments—also lent themselves to a cap-and-trade instrument. The emissions of the targeted pollutant (sulfur dioxide) were spread regionally (diminishing hotspot concerns, though the issue was still raised), generated by widespread sources with varying costs of control, and capable of being monitored inexpensively at each smokestack. Third, on the demand side, acid rain was perceived to be a growing social problem, so the electorate associated the new law with increasing social benefits. Fourth, on the supply side, changes in political leadership in the late 1980s helped prompt the policy change by strengthening the hand of leaders from states receiving acid deposition: in the Senate, Robert Byrd, from West Virginia, a coal producing state, gave up his post as Senate Majority Leader to George Mitchell, from Maine, a victim of acid rain; and Ronald Reagan, from California, was succeeded in the White House by George H.W. Bush, from New England. The switch to cap and trade in 1990 was also facilitated by a bipartisan learning process called “Project ’88,” led by Senators Tim Wirth (Democrat) and John Heinz (Republican) and staffed by Robert Stavins, which explained the success of prior allowance trading systems (such as in phasing out lead in gasoline) and thus helped legislators learn about these instruments and their net benefits. And policy entrepreneurs became active in promoting a policy change. The Environmental Defense Fund – an environmental group carving out its niche as an advocate of incentive-based instruments - helped design the 1990 Acid Rain Trading Program for the new Bush administration (KRS 1998: 354).

Although the Joskow & Schmalensee hybrid model and the additional factors we have noted here represent more a mosaic description than a formal model, this analysis reveals that at least one substantial policy instrument change – on acid rain -- is best understood by moving beyond the narrow focus on industry interests emphasized by the Olson model and its progeny, and by instead including a wider array of interests and political forces on both the demand and supply side. Further analysis could address the use of a cap and trade instrument design in subsequent air pollution regulation, such as the major Clean Air Interstate Rule (CAIR) promulgated by the Bush administration in 2005 to further reduce SO2 and NOx emissions.

C. Public Attitudes

The approaches discussed so far focus mainly, though not exclusively, on the role of organized interest groups in shaping instrument choice. But can the views of unorganized voters – public attitudes – also influence instrument choice? This subsection addresses that question. The answer is yes, but the implications for instrument choice are mixed.

Several models of democratic governance have incorporated voter sentiments into policy outcomes. Downs (1957) developed a simple but influential model in which voters, of varying ideological preferences, acted directly and individually without organized interest groups. In that model, legislators seeking re-election thus adopt laws that satisfy the preferences of the median voter. Even in a world of interest groups, public attitudes can
still matter. Denzau & Munger (1986) model the views of the unorganized general public as the slope of the hill up which interest groups must push their lobbying agenda toward adoption. That is, the more widely or intensely general public holds a viewpoint on some issue, the more difficult it is for interest groups to overcome that viewpoint. This idea is similar to Becker (1983, 1985), in which interest groups compete to offer legislators policy designs that reduce costs and thereby please the general public (voters), and more directly entails the approach in KRS (1998), which depicts an upward sloping supply curve for legislation. Even some of the most formal economic models explicitly incorporate the influence of popular, dispersed, and unorganized voters (see, e.g., Peltzman 1976).

A more ambitious theory of public attitudes goes beyond, or overcomes, interest group politics, at least in unusual episodes. This is the approach of adherents of “republican moments” in which mass movements depart from “normal” politics to achieve transformative political change. Pope (1990) and Morone (1990) discuss the phenomenon in politics, and Ackerman (1998) argues that such republican moments have transformed American constitutional law and its understanding at key intervals such as the New Deal. Farber (1992) sees the wave of environmental legislation enacted in the early 1970s as a republican moment that overcame interest group politics (rather than reflected a new interest group alignment). Elliott et al. (1985) explain the early 1970s wave of environmental law as partly a strategy by national industry to replace a patchwork of state laws with uniform national laws, and partly a race to the top (to be more “green”) between presidential candidates Nixon and Muskie, but they do not address (and appear to assume) the underlying change in public attitudes that evidently motivated both the enactment of the state laws and the candidates’ race to the top. Schroeder (1998) argues that the republican moment that spawned the major environmental laws in the early 1970s was plausibly a result of rational voting and rational response by elected leaders. Speth (2008) and Purdy (2008) see – or hope for – a coming transformation in public attitudes to address global climate change.

The theory of republican moments may explain the origin of major laws, but its implications are unclear for the details of instrument choice. For example, the 1960s/70s republican moments may have overcome opposition to environmental and consumer protection, but it is less obvious that the mass movement influenced the choice between command technology standards, taxes, and allowance trading. It is possible that what are perceived to be republican moments are actually coalitions of Baptists and bootleggers (or are coopted by such coalitions), which would help explain why the surge of democratic mobilization in the 1970s lead to suboptimal instruments that protect incumbents (Yandle 1989).

Alternatively, the republican moment may influence instrument choice through framing effects. Perhaps the narrative of the mass movement, rooted in its critique of markets and economic growth, also swept aside economic incentive instruments in its zeal for direct government control, questioning, in effect, whether market-based instruments could be the solution when market failure is the cause for environmental ills. Depicting markets as the problem and pollution as a sin requiring expiation may have tilted public support
toward absolutist government edicts (Margolis 1996: 25). Framing market-based incentive instruments as “licensing the right to pollute” and “commodification” may have tarnished those instruments in the public eye, even if the reality was otherwise (Nash 2006). It may have taken two or three decades, including experience with successful market-based incentives and the failure of central planning in the Soviet Union, to change that public attitude (and legislators’ understanding) toward acceptance of the idea that market-based incentive instruments can repair and reconstitute markets (Stewart 1986). It took time to reframe the narrative, saying, in effect: it’s not that the environment is too important to leave out of markets (Wiener 1999a: 724).

Framing of instrument choice may be catalyzed by crisis events. Public attitudes and perceptions, including republican moments, appear to be influenced by and responsive to cognitive heuristics and biases that translate into political mobilization. The availability heuristic (Sunstein & Kuran 1999) focuses attention on recent visible unusual events, leading people to overestimate low-probability risks and underestimate routine familiar highly likely risks (Kahneman, Slovic & Tversky 1982; Krier & Noll 1990). Thus, crisis events may be crucial in spurring political responses (Percival 1998; Dominic & Madin 2008). The history of regulation reveals numerous examples, from the Cuyahoga River catching fire, the death of Lake Erie, and toxic waste under Love Canal in the 1970s, to the Exxon Valdez oil spill in 1989, the Enron and Worldcom scandals, and the 9/11/2001 terrorist attacks. Percival (1998: 20-22) argues that a “trigger” event like these is needed to spur new legislation. A longer memory includes the Triangle Shirt factory fire, which spurred workplace safety regulation, although its political impact arguably faded over time (Stein 2007: 787-88). These theories on the intersection of public attitudes and cognitive heuristics might explain why certain events trigger republican moments. Still, the relationship is not deterministic. Some regulation occurs without an immediately preceding crisis event, and not every crisis event spurs regulation (Kahn 2007). Processes of social amplification are only triggered in some cases.

Cognitive focus on “available” crisis events also has a crowding-out effect on other motivations for regulatory policy, as it arguably leads to public neglect of both routine familiar risks and also remote catastrophic risks. Catastrophic risks, such as asteroid collisions, abrupt climate change, pandemics (Posner 2004; Sunstein 2007) and financial collapse (Taleb 2006), threaten huge damages. But these kinds of extreme ultra-low probability catastrophic risks are neglected (even as low-probability but “available” visible crises evoke strong responses) because the catastrophic event occurs so rarely that it is not within the memorable experience that gives rise to the availability heuristic (Slovic & Weber 2002). And public attitudes may be subject to “mass numbing,” responding intensely to the plight of a single individual but remaining unmoved by the plight of thousands or millions (Slovic 2007).

If so, the upshot is that the public responds to recent visible crises (more strongly than do experts, who calculate probability times impact), but the public neglects (compared to experts) both routine widespread risks and rare catastrophic risks. Breyer (1993) and Posner (2004) bemoaned the influence of these heuristics on public attitudes, and in turn
on Congressional priority-setting, as a suboptimal distortion of regulatory priorities that Breyer (1993) termed the “vicious circle.” Thaler & Sunstein (2008) propose opportunities to use framing effects proactively to shape heuristics and improve regulatory policy.

The influence on instrument choice of these heuristic public attitudes requires additional inquiry. The crisis-response dynamic and its framing narrative may tend to favor regulatory designs that employ strict government edicts (“never again” or “zero tolerance”), such as command technology standards in environmental policy, stringent regulation of financial institutions, and aggressive homeland security policies that impinge on liberty and privacy. On the other hand, financial crises have often been met with adoption of strong information disclosure instruments (both in the New Deal securities laws, and in the Sarbanes-Oxley law following the Enron and Worldcom scandals). Likewise, the Bhopal chemical plant disaster was followed by the information disclosure requirements in the Toxics Release Inventory (enacted in 1986) and the risk management plans under Clean Air Act section 112(r) (enacted in 1990). These examples suggest a role for incentive-based instruments in some cases of crisis response. And to forestall what some call a looming climate crisis, the current political enthusiasm for cap-and-trade instruments (as opposed to other instruments) to reduce greenhouse gas emissions appears to derive from several factors: the expected cost savings, learning over time about instruments, the desire for effective caps on emissions, and the distributional attractiveness of deploying some allowance allocations to mollify critics and recycling allowance auction revenues to cut other taxes, all with little or no risk of environmental hotspots (DeShazo & Freeman 2007; Wiener 2008). Perhaps actually experiencing a major climate crisis (worse than Hurricane Katrina—say, Greenland ice melting and flooding lower Manhattan and Florida)—would spur adoption of instruments that are not so incentive-based, such as geoengineering strategies to cool the planet directly.

In sum, public attitudes appear to affect instrument choice, through unorganized voters’ preferences, episodic republican moments, and heuristic framing effects especially after available crisis events. But the direction and degree of amplification of these effects on instrument choice is mixed or unclear. Further empirical research could seek to disentangle these effects and test their relative influence.

D. Positive Choice at the International and Supranational Level

The theories and evidence assessed thus far have all examined positive politics in the United States, at the federal level. That system has particular constitutional features, including voting and election systems, campaign finance laws, legislative majority rule, a bicameral national legislature, separation of powers with the possibility of a presidential veto, federalism with the possibility of subnational regulation, and judicial review. It also has a distinct political culture. These features are obviously different in other political systems. Many countries employ parliamentary legislative systems, in which the prime minister necessarily has a majority in the legislature. Some impose restrictions on
political expenditures by interest groups, and some have more deferential doctrines of judicial review.

These and other kinds of variation in political institutions seem bound to affect the positive politics of mechanism choice. Below we suggest further research on this question.

A comparative analysis of the positive politics of instrument choice across countries is a worthy endeavor but beyond the scope of this paper. Here we offer two examples of positive instrument choice in other systems that may help put the US system and experience in context: international environmental treaties, and the European Union. Both illustrate the importance of the voting rule for adoption of law (see Buchanan & Tullock 1962).

1. International environmental regulation: CFCs and GHGs

There is no strong positive theory of global regulation. Standard public choice theory predicts even less regulation at the global level than at the national level because global beneficiaries are even more diffuse, industry cost-bearers are still concentrated, and – crucially –the voting rule for international treaties holds that countries must consent to be bound and hence must perceive national net benefits (broadly construed) to decide to join (unlike national majority rule or fiat, which can impose costs on dissenters) (Wiener 1999a, 1999b). As Farber (1997: 1314) reports, “The basic principle of international law, after all, is that it binds states only with their own consent.”

Moreover, rent-seeking (and the bootlegger side of Baptist-bootlegger coalitions) might seem more difficult under the consent voting rule at the international level than under the majority or fiat voting rule at the national level. Rent-seeking depends on the coercive power of the state to shift gains from losers to winners (Wiener 1999b: 769-771; Posner 1971; Posner 1974: 344; Peltzman 1976). Indeed, Mueller (1976: 401-03 & n.9) opined that for this reason, rent-seeking in international agreements is “out of the question,” though adding a caveat that there may be distributional battles over the joint gains from cooperation. Moe (1990: 221, 222 n.9) argued that “the unique thing about public authority is that whoever gets to exercise it has the right to tell everyone else what to do, whether they want to do it or not .... Public authority gives [the winners] the right to make themselves better off at [the loser’s] expense. Their decisions are legitimate and binding. ... This kind of outcome-- redistribution that makes some people better off and some people worse off-- is alien to the economic world of voluntary exchange.” According to this view, one would expect little, if any, rent-seeking regulation in the international arena of voluntary agreements.

Yet as an empirical matter, the reality is that there is significant regulation at the global level, more than would be expected under standard public choice theory. And there are some attempts at rent-seeking under international treaty regimes as well (Wiener 1999b).
Countries are of complex plural entities, not monoliths, so rent-seeking can occur within each country governed by majority rule or more coercive institutions; but regulation and rent-seeking are more surprising under the consent rule of international treaties. Seeking to explain this puzzle, Keohane (1983) and Ostrom & Keohane (1996) compare global treaty law to local systems of group property and argue that international regimes arise to facilitate the benefits of cooperation. Such regimes can succeed in situations of low transaction costs, repeat playing (long-term relationships), strong social norms, direct benefits and side payments to cooperate, and monitoring and sanctions for violations. Wiener (1999a, 1999b) hypothesizes that consent-based political systems (such as international law) are (i) less likely to adopt law than are more coercive political systems such as fiat or majority rule; (ii) less likely to adopt laws that impose costs or adverse side effects on parties than are more coercive systems such as national law; and, when they do adopt regulatory law, consent-based political systems such as international treaties are (iii) more likely to adopt allowance trading than to adopt either command technology standards or taxes. The first two hypotheses are consistent with Buchanan & Tullock (1962) who show that as the voting rule moves along the spectrum from fiat to majority to unanimity (i.e., from coercive to non-coercive systems), the costs of adopting law rise but the costs imposed on dissenters decline. The third hypothesis on the pattern of instrument choice follows, Wiener (1999a) suggests, from the lack of coercion and the need for side payments at the international level, which mean that allowance trading is better able (i.e. at less cost than alternative instruments) to attract participation, through the allocation of allowances.

For example, in the treaty negotiations over whether to phase out CFCs to protect the stratospheric ozone layer (a global public good), cooperation among countries required extensive negotiations and, ultimately, side payments to China and India through the Montreal Protocol Fund. In selecting the regulatory instrument, the 1987 Montreal Protocol regime set quantity limits on production and authorized a limited form of allowance trading called “industrial rationalization.” Meanwhile, rent-seeking was significant: US manufacturers of CFCs, who had initially opposed international controls on CFCs, switched positions in 1986 to press for an aggressive phaseout of CFCs in the Montreal Protocol, in part because they perceived profit opportunities in the new regime. Since US manufacturers were farther ahead in the production of CFC substitutes than were their competitors, a rapid CFC phaseout, although it would cost them some, would hurt their rivals far more (Wiener 1999b: 772-73; Litfin 1994: 108; Hammit & Thompson 1997: 43; McInnis 1992: 129). Although the US manufacturers no doubt would claim credit for making socially valuable investments in new products, their cooperation in the Montreal Protocol is clearly also consistent with an alignment of bootleggers and Baptists, in which a rent-seeking industry sought to burden its rivals while environmental advocates sought to protect the planet. The Montreal Protocol’s success also shows that such rent-seeking could occur in the international arena within a consent framework that constituted a Pareto improvement to all parties (countries). Indeed, a curiosity here is that the rival CFC producers (such as in Europe) went along – or perhaps were obliged to do so within their own national political systems.
The negotiations over the international climate change treaties also illustrate the importance of the voting rule. Progress has been slow because countries must consent to be bound, and several key countries have demurred. The US never ratified and then withdrew from the 1997 Kyoto Protocol, in part because the US Senate had voted 95-0 in 1997 not to ratify a treaty which lacked parallel commitments by major developing countries. China, India, Brazil, Indonesia, and other major developing countries have also resisted agreeing to limits on their future greenhouse gas (GHG) emissions. No doubt rent-seeking efforts are rife as different industries and energy sources seek to raise rivals’ costs through the restrictions imposed and subsidies doled out under a climate treaty regime. Meanwhile, as to instrument choice, when the US advocated allowance trading in the 1990s, the idea encountered initial opposition from many quarters, notably the European Union; but after considerable effort by the US to explain the benefits of allowance trading for greenhouse gases, that instrument was authorized informally in the 1992 Framework Convention on Climate Change and more formally in the 1997 Kyoto Protocol (Wiener 2001). A cap-and-trade system is likely to be the instrument adopted in future US and international climate policy as well because it enables cost savings, a cap on emissions, and a distribution of burden and of headroom allowances that best attracts participation, without inducing hotspots (Wiener 1999a; DeShazo & Freeman 2007).

In the absence of formal international governance regimes, there is still room for “private governance,” including self-regulation by multinational corporations, codes of conduct promulgated by international organizations, and consumer standards monitored by NGOs (Gereffi & Mayer 2006). Although private governance arrangements often arise in response to calls by consumer and political activists (Spar & La Mure 2003), they are also regularly products of rent-seeking and are sustained by industry initiatives not unlike those exercised in domestic politics (Wiener 1999b). Private coalitions that drive the formation of international agreements often resemble the purest forms of Baptist-bootlegger coalitions, because political sentiment is so thin and political monitoring so sparse in the international arena. Examples of coalitions with surprising bedfellows include the Kimberley Process, in which NGOs paired with DeBeers to remove conflict diamonds from the consumer market and to remove alluvial competitors from global supply (Richman 2009).

2. European Union environmental regulation: GHGs

Historically, the EU has preferred environmental taxes over allowance trading (Andersen 1994; Svendsen 1998; Wallart 1999; Harrington et al. 2004). In the 1990s, the EU proposed a carbon tax and denounced the US proposal of allowance trading for GHGs. But no EU carbon tax was adopted, and starting in 2001 the EU instead adopted allowance trading for GHGs. It has now implemented the EU Emissions Trading System (ETS), the largest cap-and-trade system in the world. Why the switch in instrument choice?

Several factors help explain this shift that do not relate to the European Union’s supranational institutions. When the EU finally turned to implementing its Kyoto commitments, it realized that the ETS would impose lower costs than would command
standards. (But a tax would arguably have imposed even lower costs.) Moreover, by 2001 the EU had been able to learn from the successful US experience with SO2 allowance trading under the 1990 Clean Air Act, as well as from several other cap-and-trade systems in the US, Canada, New Zealand, and elsewhere, and came to realize its benefits. (For discussions of Canada’s and New Zealand’s successful tradable permit systems to curb overfishing, see Wyman 2005 and Newell et al. 2005.) The successes of allowance trading in the US and the benefits of this instrument for GHG control were presented to EU leaders in numerous fora, including an early seminar on market-based environmental policies organized by US government officials (including Richard Stewart, Richard Schmalensee, and Jonathan Wiener) in February 1990 for officials from the EU and other countries during an IPCC session. This conversation continued through the 1990s, and the great success of the US acid rain trading program put to rest many concerns about cap and trade (see Svendsen 1998). Relatedly, policy entrepreneurs brought the theory and evidence of allowance trading systems from the US to the European Commission (Peter Zapfel, who studied with Rob Stavins and Richard Newell at the Kennedy School, is now one of the key staffers at the European Commission running the EU ETS). And although the EU denounced allowance trading in the 1990s in part for reasons of symbolic politics – in order to shame the US and seek the upper hand in post-Cold War international relations -- after George W. Bush withdrew the US from the Kyoto Protocol in 2001, the EU could easily shame the US without disparaging allowance trading.

Thus, by 2001, allowance trading had become a more politically and functionally appealing instrument for the EU and its member states. But a tax might still have looked more attractive to EU decision makers. To understand the EU’s shift from taxes to allowance trading, it helps to appreciate that the EU voting rules also played a critical role. EU law required unanimity or consent among EU member states to adopt an EU tax, but poorer EU member states objected to a uniform carbon tax and blocked adoption. A cap-and-trade system, by contrast, attracted consent because extra allowances could be distributed to poorer member states. Allowance trading enabled the initial distribution of burden to be decoupled from the question of the stringency of the cap and the market price for an allowance, whereas a tax could not be adjusted as easily to accommodate the comparative needs of the poorer member states. The EU voting rules requiring unanimity significantly affected the ultimate mechanism chosen to control GHGs, favoring trading over taxes, as hypothesized by Wiener (1999a, 1999b).

These two cases – international treaties and the supranational EU system – indicate the importance of voting rules and political institutions in the positive choice of policy instruments beyond the US context. They also illustrate the fruitfulness of examining how cross-national variation affects instrument choice, and implicate directly some of our suggestions for future research.

IV. Future Research
The literatures on positive and normative mechanism choice are already extensive, and we do not pretend to envision the entire future of this field of research. We instead offer here a few suggestions for future research, with a focus on the theory and evidence of positive politics – that is, on the public choice of mechanism choice.

A. Voting Rules. Since at least Buchanan & Tullock (1962), scholars have investigated how different voting rules affect politics. We have sketched here some variation in regulatory outcomes that can be explained by differences between, for example, the American versus European political systems. We encourage future research to examine more systematically how positive instrument choice is affected by alternative voting rules. Voting rules to be compared include fiat (e.g. dictatorships, Presidential executive orders, or edicts from corporate CEOs), majority rule (the standard case of rule by N/2 + 1), supermajority rule (e.g. 60 votes out of 100 needed to end debate in the US Senate, or 67 votes out of 100 needed for the US Senate to ratify a treaty), bicameralism (votes required in both the House and Senate, see Stearns (1994)), separation of powers (requiring both legislative vote and executive signature), consent (e.g. nations’ consent to an international treaty, neighbors’ consent to restrictive covenants, or firms’ consent to binding contracts), and unanimity (the extreme case of universal veto). This research would easily translate into normative conclusions about the optimal instrument under each voting rule. And we encourage inquiry into how options for side payments – mechanisms to compensate losers and thereby attract additional votes or consent – can be incorporated into alternative political systems.

B. Scale. The size of the polity, communications among members, and vertical (federal) relations may also affect positive choice among instruments. For example, the role of interest groups (factions), public attitudes, and regulators may differ at each scale (Revesz 2001). Madison worried in Federalist 10 that small polities would be more vulnerable to capture by factions, but others may worry that large-scale polities are more vulnerable to interest group capture because the transaction costs are higher for unorganized voters to learn about and monitor political decisions at higher scales. Innovations such as the internet may affect those costs. Research should pursue these questions of scale, information costs, administrative institutions, monitoring, enforcement, and culture, to investigate how they may influence instrument choice at the local, state, national, and international levels.

C. Constitutions. In addition to, or perhaps combining, the above issues of voting rules and scale, research could compare positive instrument choice in different constitutional systems of government, such as republics (with separation of powers), parliamentary, authoritarian, and other systems. For example, some have argued that public goods are protected better by democracies than by dictatorships (Congleton 2001: ch. 12), while others have alleged the converse (Ophuls 1977). Among democracies, one might investigate whether parliamentary systems are more likely to select certain policy instruments than are separation of powers systems like the US Congress in which opposing parties may control the legislature and the executive (Krehbiel 1998); and within each legislature, whether different committees (King 1997) or legislative procedures (Krehbiel 1992) affect instrument choice. Of particular interest might be how
constitutional designs enshrine current regulatory systems and make reform either more or less difficult. Examining political feedback mechanisms to understand how initial mechanism choice becomes an explanatory variable for future mechanism selections, would contribute to our understanding of policy outcomes and change. To be sure, path dependency is a common explanation for why policy instruments continue beyond their intended life, and mimicry might also explain parallels in mechanism selection across policy domains. But political institutions also shape debate, perceptions, and resource allocations that can affect the political process. Understanding the endogeneity of such mechanism selections might require more sophisticated models than those discussed here.

D. Topical domains. We also encourage additional inquiry into how instrument choice varies across topical domains, such as environmental protection, energy, communications, transportation, workplace safety, food safety, financial markets, and homeland security. In this paper, several models (e.g. Wilson 1984, McNollGast 1987, and Keohane, Revesz & Stavins 1998) discuss how different kinds of institutions might be prone to different instrument choices. Future research might focus on how the particular attributes of different topical areas might call for different types of instruments (both normatively and in positive politics). For example, environmental protection might be better addressed with instruments that differ from those best used in financial regulation, communications, homeland security, or food safety.

E. Behavioral sciences. New research in behavioral psychology, economics and neuroscience also offer significant returns to understanding mechanism choice. The burgeoning fields of behavioral law and economics, neuroscience, and political science collectively offer opportunities to understand the sources and consequences of phenomena such as framing, crises and availability heuristics, remote catastrophic risks, implicit trust or distrust in social institutions, and related factors. For example, as discussed above, some analyses suggest that market-based instruments often arise in response to mid-level crises, whereas major crises that seize the public’s attention often result in command-and-control regulation, especially if a narrative emerges from the crisis that demands expiation of sin and stringent edicts. But we also noted counterexamples of market-based incentives adopted in response to major crises in financial and environmental settings. This area deserves additional attention, and a satisfying inquiry will likely require multiple methodologies and theories in the social sciences.

F. Political Institutions and Instruments. Finally, though we are wary of the search for a single grand theory, we encourage the pursuit of models that integrate demand, supply, entrepreneurship, public attitudes, and institutions, in order to shed more light on how the spectrum of different political systems intermeshes with the spectrum of alternative regulatory instruments. Although significant work has examined cross-national variation in political design (see, e.g., Berger & Dore 1996), we lack a dynamic understanding of how political templates correspond to templates of regulatory instruments. And although there is a broad literature on the diffusion or transplantation of judicial doctrines across countries (Watson 1993), we are still groping toward a solid understanding of the
diffusion or borrowing of approaches to administrative regulation across political systems (Lazer 2005; Wiener 2001, 2006).

Regulatory policy and mechanism design research have advanced a long way from the traditional template that called for command-and-control regulation to correct market failures; the menu of instrument choices in the toolbox continues to grow, and the selections actually made by legislators and regulators appear to be following somewhat more closely the normative recommendations of analysts. Positive political research that had based predictions of parochial rent-seeking regulation mainly on the behavior of organized interest groups has puzzled over contrary evidence, such as the growth of general-interest legislation, the deregulation of economic regulation, and the shift toward adoption of incentive-based allowance trading systems in the US and EU. And yet we also observe rent-seeking within ostensibly general interest legislation and international treaty regimes. In response, research has sought to incorporate a broader array of explanatory factors. These are examples of partial moves toward a more comprehensive understanding of mechanism choice.
References


Buchanan, James M. & Gordon Tullock. 1975. Polluters' Profits and Political Response: Direct Controls Versus Taxes, 65 Am. Econ. Rev. 139-.


Dales, John H. 1968. Pollution, Property, and Prices.


Gerschenkron, Alexander, 1945. Bread and Democracy in Germany.


Goulder, Lawrence H., et al., 1997. Revenue-Raising Versus Other Approaches to Environmental Protection: The Critical Significance of Preexisting Tax Distortions, 28 Rand J. Econ. 708-.


Hardin, Garrett. 1968. The Tragedy of the Commons. 162 Science 1243-.


Keohane, Robert O. 1983. The Demand for International Regimes, in International Regimes 141, 146-47 (Stephen D. Krasner ed.).


Lowi, Theodore. 1979. The End Of Liberalism: The Second Republic Of The United States (2nd ed.).


Ostrom, Elinor 1990. Governing the Commons.


Peltzman, Sam. 1976. Toward a More General Theory of Regulation, 19 J.L. & Econ. 211.


Renda, Andrea. 2006. Impact Assessment in the EU (Center for European Policy Studies)


Taleb, Nicholas Nassim. 2006. The Black Swan.


Tollison, Robert D. 1991. Regulation and Interest Groups, in Regulation: Economic Theory and History 59, 64-66, 72 (Jack High ed.).


