Electoral Accountability: Recent Theoretical and Empirical Work

Scott Ashworth

Harris School of Public Policy Studies, University of Chicago, Chicago, Illinois 60637;
email: sashwort@uchicago.edu

Keywords

Abstract

Competitive elections create a relationship of formal accountability between
policy makers and citizens. Recent theoretical work suggests that there are
limits on how well this formal accountability links policy decisions to citizen preferences. In particular, incumbents' incentives are driven not by the
voters' evaluation of the normative desirability of outcomes but by the outcome's information about the incumbent's type (e.g., competence or ideology). This review surveys both this body of theory and the robust empirical
literature it has spawned. It concludes with a short discussion of ongoing
work that attempts to integrate this theoretical perspective with a richer
view of policy-making institutions.

INTRODUCTION

Competitive elections create a relationship of formal accountability between policy makers and citizens—electoral rewards and punishments can be handed out on election day. Ideally, this formal accountability leads to better governance. How well does this work in practice?

A particularly useful perspective on this question is given by the theory of political agency. This theory uses tools from the economic theory of contracts to elucidate the incentives provided by elections. This theoretical literature has taught us two major lessons: incentives are driven by the incumbent's desire to impress the voters, and this desire often conflicts with the normative imperative to advance the voters' interests. Importantly, the incumbent's electoral incentives to act contrary to the voters' interests might arise even if all voters act perfectly rationally.

It might seem counterintuitive that rational voters can create such bad incentives. But consider, as an example, an incumbent who must decide whether to prioritize policy aimed at boosting the economy in the short run, or policy aimed at a long-run problem like global warming. Assume that work on global warming has greater impact on the voters' welfare. If the performance of the economy is more informative about the incumbent's competence or about her ideology, voters will nonetheless respond more to the economy. After all, voters can only affect what happens in the future, so rationality demands that the voter select the best candidate going forward. So voters are led by fully rational Bayesian information processing to respond to the more informative outcome, even though they know that the incumbent has ignored their best interests in pursuing that outcome.

This review surveys the literature, theoretical and empirical, developing this perspective. I first take a closer look at how equilibrium considerations jointly determine voter behavior and incumbent policy choices. Doing so elucidates the forces that link an incumbent's performance and her chances of reelection, helping to identify the conditions under which formal accountability does, or does not, lead to policy choices valued by voters. The theoretical examination also generates comparative static predictions, identifying several institutional and contextual features that make an incumbent's behavior more or less responsive to the incumbent's anticipation of voter responses. The analysis also identifies an important selection effect arising from the internal logic of the accountability relationship itself, something that will complicate empirical testing.

The second step is empirical, assessing the comparative static predictions in field data. I emphasize a recent trend toward quantitative case studies that exploit detailed knowledge of institutions and political context to produce highly credible research designs. These studies show quite robustly that responsiveness in practice conforms to the comparative static predictions of the theory. In particular, responsiveness is greater when the rewards to reelection are higher, and when voters get more precise information about the incumbent's performance.

I conclude by briefly discussing recent work that further develops the theoretical perspective by embedding the electoral accountability relationship in a richer institutional environment, shedding light on the separation of powers. These applications point to an underdeveloped strength of the political agency approach, namely, that it is a natural laboratory for integrating extant thinking about elite strategy and mass behavior.

THEORETICAL FRAMEWORK

A theoretical account of electoral accountability must contain at least two components: an electorate that decides whether or not to retain an incumbent, at least potentially on the basis of her performance, and an incumbent who has the opportunity to respond to her anticipations of the electorate's decision. Most of the applied theory literature adopts a simple framework with

two policy-making periods as the simplest way to embody both components. In the first period, an incumbent is in office, and she must choose an action. This action affects some performance measure observed by a representative voter, who then decides to reelect the incumbent or to replace her with a challenger. Whichever candidate wins this election then chooses an action in the second policy-making period. The two policy-making environments are the same in terms of the actions available to the politician. But they differ in that there is no election after the second period—instead, the game ends with the politician's action choice.

We will see below that this framework accommodates a rich variety of both policy problems and assumptions about what voters do and do not know about politicians. But some determinants of incentives are common to all of these models. In every case, the incumbent gets some benefit, $B > 0$, if she is reelected. This benefit incorporates things like the salary and perks of office, the value of improving policy while continuing in office, gains from potential corruption open to winners, or direct psychic benefits to holding office. Incentives for the incumbent, then, come from a desire to adjust her behavior in such a way that her probability of reelection, and thus of getting benefit B, is greater. To formally capture this incentive, assume that there are only two possible actions, $a = 0$ and $a = 1$. Absent incentives, the incumbent prefers to choose $a = 0$ (which gives her a nonelectoral benefit worth $b > 0$), but she knows that she is more likely to get reelected if she chooses $a = 1$ (which gives her nonelectoral benefit 0). Write Pr(reelect|a) for the probability that she is reelected if she chooses action a. Then her payoff when she chooses $a = 0$ is

$$B\text{Pr(reelect} \mid 0) + b,$$

and her payoff when she chooses $a = 1$ is

$$B\,\text{Pr(reelect} \mid 1).$$

The incumbent's action choice is responsive to the voter's sanction if

$$B\,\text{Pr(reelect} \mid 1) \geq B\,\text{Pr(reelect} \mid 0) + b,$$

which we can rewrite as

$$B[\text{Pr(reelect} \mid 1) - \text{Pr(reelect} \mid 0)] \geq b. \qquad\qquad 1.$$

Already we can see the first empirical implication of this class of models: The incumbent's action choice is more likely to be responsive to the voter's sanction when the benefit to being reelected is greater. This predicts, for example, that responsiveness will increase when the salary of the winner increases. Taken to the extreme, this condition implies that term limits completely remove the possibility of responsiveness; if the incumbent cannot be reelected, then both probabilities on the left-hand-side are 0, and the condition cannot be satisfied.

Another implication comes from the cost side: The incumbent's action choice is less likely to be responsive to the voter's sanction when the value of b, i.e., the direct benefit from action $a = 0$, is greater. For example, stronger incentives from parties or interest groups to choose some particular policy will reduce responsiveness to voters.

To make further progress, we must be more specific about how the probabilities of reelection are determined. This is done by analysis of the voter's reelection rule, which reelects the incumbent only if performance meets some standard. What determines this standard?

In most of the contemporary literature, the standard is determined in equilibrium by the voter's desire to select better-performing politicians in the future. This contrasts with the seminal works of Barro (1973), Ferejohn (1986), and Austen-Smith & Banks (1989). In their models, all politicians are identical, so there is no scope for selection—the voter is indifferent between the two candidates at the time of the election. Thus any reelection rule is sequentially rational, and any reelection

rule combined with an action choice by the incumbent that is a best response to that rule make up an equilibrium.

A determinate prediction in these pure moral hazard models comes from selecting the equilibrium that maximizes the voter's payoff. This equilibrium selection is appealing on several grounds. First, it is the equilibrium that voters would select, perhaps through pre-play communication, if they could determine which equilibrium would be played. It is thus a natural normative benchmark for the agency relationship. Second, this selection allows for analytical separation between the welfare losses from playing a bad equilibrium as opposed to losses that are inherent in the agency relationship.

Although the pure moral hazard model has continued to be useful in applications (Seabright 1996, Persson et al. 1997, Shi & Svensson 2006, Bueno de Mesquita 2007, Fearon 2011), the consensus in the literature is that it is not fully satisfactory from a theoretical point of view. In an important article, Fearon (1999) shows that the set of equilibria of the pure moral hazard model is often not robust to allowing candidate heterogeneity. Differences between candidates that are relevant to the voter's future payoffs, even small ones, can radically change the predictions. The intuition is simple and compelling: if the candidates differ in any way that is relevant to the voter's future payoff, he will not be indifferent between the two candidates. With that, the procedure used to solve pure moral hazard models falls apart. Instead, the voter's reelection decision is typically pinned down by rational learning and forward-looking optimization inside of an equilibrium.

The literature has explored several substantive assumptions about the candidate heterogeneity. Here are some specific examples:

- **Spatial policy making (based on Fearon 1999).** The incumbent must choose a policy in a one-dimensional policy space. Both the voter and the incumbent have ideal points in this space and want policy to be close to their ideal points. The voter is uncertain about the incumbent's ideal point—it might be close to his, or it might be more extreme.

- **Corruption in program implementation (based on de Janvry et al. 2010).** The incumbent is a mayor who must oversee the implementation of a new program out of funds from a grant from the central government. The success of the program depends on whether the incumbent chooses the cost-effective implementation strategy or an inefficient strategy that creates scope for graft. The voter wants the cost-effective strategy used but is uncertain about whether the incumbent shares that preference or if she prefers to reduce program efficiency in exchange for rents. The incumbent, though, does know her own preference.

- **Constituency service (based on Ashworth & Bueno de Mesquita 2006).** The incumbent is a legislator who must decide how to allocate her effort between constituency service and other activities. The voter observes only the quality of his interactions with various agencies, a quantity that depends on both the oversight effort and the competence of the incumbent, as well as factors completely independent of the incumbent ("luck"). The voter knows that the incumbent would prefer to devote less effort to the oversight task. But he does not know how competent the incumbent is.

These examples do not exhaust the richness of the political agency approach (see Persson & Tabellini 2000, Canes-Wrone et al. 2001, and Besley 2006 for other interesting specifications). Rather than being discouraged by all of this diversity, an outside observer should be gratified that all of these models tell roughly similar stories about incentives and comparative statics, suggesting that the insights transcend the specific models in which they were initially developed. The rest of this section explores the contributions these insights have for traditional thinking about accountability, whereas the following section describes the comparative statics and their confrontations with data.

Retrospective Voting?

In each of the examples, the candidates are differentiated by type: moderate or extreme policy preference in the first case, interest in corruption in the second case, and competence in the third case. Call the types θ_I for the incumbent and θ_C for the challenger, and assume that the types are ordered so that the voter prefers higher types. The voter observes some public signal s—e.g., success of the program or the value of constituency service—and then forms expectations about the type of each candidate, $\mathbb{E}(\theta_I \mid s)$ and $\mathbb{E}(\theta_C)$. These expectations are crucial for rational voting—the voter knows that the winner of the election will choose an action that affects his welfare, and the higher is the type, the more likely is that action to advance voter welfare. This could be because the action is done more competently or because it is guided by preferences more like his own.

The voter chooses whom to elect by comparing the difference in these expectations to a threshold v:

$$\text{reelect the incumbent if and only if } \mathbb{E}(\theta_I \mid s) - \mathbb{E}(\theta_C) \geq v.$$

This threshold does not depend on the incumbent's action in any way—the action matters only through its effect on the voter's beliefs about the incumbent's type.[1] Instead, the threshold captures any factors that affect vote choice that the incumbent cannot affect. An exogenous source of incumbency advantage is captured, for example, by assuming $v < 0$. The threshold can be random, to capture partisan swings between policy making and the election. Indeed, the random component of the threshold could even arise from nonrational factors, like depressed affect toward incumbents as the result of a recent loss by the local football team (Healy et al. 2010).

This theoretical perspective perhaps seems more familiar than it really is. Voters look to the past to decide how to vote, but they do so from purely prospective motives. And precisely because they are behaving rationally, these prospective voters might well elect an incumbent who they know took the wrong actions from the perspective of maximizing the citizens' own welfare.

It's useful to contrast this modern perspective with the older literature based around so-called retrospective voting (Key 1966, Fiorina 1981). The idea is that voters apply a backward-looking standard to evaluate incumbents (or incumbent parties), throwing out candidates who did not meet the standard. That much is entirely consistent with the models discussed above. But the traditional literature went further, and contrasted this retrospection with a prospective notion of behavior. V.O. Key (1966, p. 61), for example, writes:

> [The electorate] judges retrospectively; it commands prospectively only insofar as it expresses either approval or disapproval of that which has happened before. Voters may reject what they have known; or they may approve what they have known. They are not likely to be attracted in great numbers by promises of the novel or unknown.

This is not at all a good description of what's going on in modern accountability models. A rational voter chooses whom to vote for based on the expected payoffs he can expect from the incumbent and from the challenger; approval or disapproval of something that has already happened is neither here nor there. A rational voter will absolutely look to the past for information,

---

[1]Several researchers consider infinite horizon political agency games with heterogeneity (Banks & Sundaram 1993, Duggan 2000, Banks & Duggan 2006, Meirowitz 2007, Snyder & Ting 2008, Schwabe 2009). It is no longer true in these models that voters care only about type. Instead, they must select the candidate who they believe will be best for them in the future, taking into account both type and predicted equilibrium play.

but he does so only insofar as it helps to make a prospective evaluation of the candidates in the election at hand. As Besley (2006, p. 106) writes:

> One key implication of [the accountability model's] approach is that there really is no meaningful distinction between prospective and retrospective voting. It is precisely because there is information content in past actions about future behavior that retrospective voting is rational.

So far, this discussion can be thought of as a simple recapitulation of the Fearon (1999) critique of models in the tradition of Ferejohn (1986). But we can go further—fully taking on board the informational rationale for allowing the past to influence the vote also calls into question the idea that the reelection standard reflects "approval or disapproval of that which has happened before" (Key 1966, p. 61). Forward-looking, rational voters will vote based on how well the incumbent advanced their interests only when advancing voter interests also convinces the voter that the incumbent is a high type. Unfortunately, these two concepts come apart in quite natural settings.

This point is perhaps easiest to see in the context studied by Daley & Snowberg (2011). They consider an incumbent who must divide her time between working on improving public policy and working on her campaign. Under their assumptions, campaigning is a lower cost way to influence voter beliefs about competence than is policy work. This means that incumbents will devote effort to campaigning, even though voters do not directly value that effort. Voters are not fooled; indeed, they may complain loudly that incumbents spend too much time on their campaigns. Nonetheless, the campaign activity does effectively signal competence, and voters, being rational and forward-looking, respond by rewarding incumbents for the very campaigning that they do not like. As a result, the incumbent will not choose the best action profile for the voter. And all because she is being responsive to the electoral incentives provided by rational voters!

It is important to note that the discrepancy between actions that are best for the voter and actions that best attract votes is not based on some discrepancy between what voters want and what's "really" good for them. The utility function that guides voter choices in Daley & Snowberg's model is precisely the utility function that describes voters' true welfare. The point, rather, is that the voter's choice in the election is not guided by his evaluation of the actions the incumbent took. It is instead based on the information that all of his observations contain about the incumbent's type.

When Will Incentives Be Misaligned?

Daley & Snowberg's (2011) model of campaign finance is an example of a "multitask model," in the spirit of the classic article by Holmström & Milgrom (1991). Holmström & Milgrom's model has also been adapted for elections by Lohmann (1998), Ashworth (2005), Ashworth & Bueno de Mesquita (2006), and Gelbach (2007). Related normative conclusions arise in the "pandering model," which derives from the reputational herding literature initiated by Scharfstein & Stein (1990) and is adopted for elections by Coate & Morris (1995), Canes-Wrone et al. (2001), Maskin & Tirole (2004), and Besley (2006).

The multitask model starts with the reasonable assumption that the incumbent takes multiple actions, each in a different area. Which will she work hardest on? One factor that contributes to the answer, clearly, is the incumbent's own preferences. To highlight the divergence between voter preferences and electoral incentives at issue in this section, assume that the incumbent has exactly the same preferences over actions as the voter. Thus, without incentives from the election, the incumbent would take exactly the best actions for the voter.

But the incumbent also wants to get reelected. And she knows that the probability of that happening is increasing in the voter's expectation of her type, conditional on the outcomes that

the voter observes. As a result, the incumbent wants, all else being equal, to take actions that have the maximal possible impact on the voter's beliefs. And if there are two actions, one of which has a bigger impact on beliefs than the other, then the incumbent has an extra reason to increase the high-impact action. This creates a distortion away from the best action for the voter.

The root cause of inefficiency in pandering models, on the other hand, is ex ante uncertainty about which of two policies is best from the perspective of voters. The voters have some beliefs about which policy is best, while the incumbent has better information. But she cannot share her information credibly with the voters. So what should she do when her private information indicates that the voters' beliefs are incorrect?

The problem is easy when voters know with certainty both that the incumbent shares their preferences and that the incumbent is at least as competent as the challenger. Then there is an equilibrium in which the incumbent always does what she thinks is best, and the voters always reelect her. This is because, given the voter's certainty about the incumbent's type, an unexpected policy choice does not serve as a signal of anything bad about her. Thus there is no risk to her in doing what she thinks is the best policy.

Matters are different when the voter is uncertain enough about the incumbent's type to think that there is some chance the incumbent is systematically more likely than the challenger to like the wrong policy, either because of misaligned preferences over outcomes or because of incompetence in processing information. [Fox & Shotts (2009) study a model in which both worries are important.] Then choosing a policy different than the voter's best guess about what is in his interest does carry electoral risk for the incumbent. To see this, consider a voter who believes both that a particular policy is likely to be best for him and also that the incumbent is likely to share his preferences over policy. If he sees an unexpected policy choice, he learns that one of his assumptions is less likely to be true. Rational Bayesian updating implies that he reduces the probability he attributes to each of the potentially suspect assumptions; in particular, this means that the voter's assessment of the quality of the incumbent goes down, and that makes the incumbent vulnerable in the election.

As a result, the incumbent has an incentive to choose a policy that is ex-ante popular with voters, even though she knows it is the wrong decision, given all of the information. This is what the literature means by pandering. The key to controlling this behavior is to prevent voters from drawing adverse inferences about the incumbent when policy is not what they think that they want. This will be the case when, for example, the voter's uncertainty about which policy is best resolves before the election. Substantively, this might mean that the election is far away, so there is time to learn that the incumbent's choice was actually good. Canes-Wrone & Shotts (2004) implement this idea empirically and find that presidents make budget proposals more in line with public opinion when the election is soon and when presidential popularity is near its mean.

Potential institutional designs that respond to the distortions created by multitask problems or pandering involve weakening the link between the tempting actions and reelection. For example, Prat (2005) and Fox & Van Weelden (2012) show that backing away from transparency can help: if voters cannot observe that the popular action was taken, they cannot reward it. Ashworth (2005) argues that term limits can improve incentives when there is a multitasking problem. Both policy responses work by taking away the possibility of using the action choice to impress the voters. This is the electoral version of a common theme in the contract theory literature: low-powered incentives can be best when performance measures are not properly aligned with the principal's (here, the voter's) interests (Holmström & Milgrom 1991). These policies unambiguously improve the action choice of the incumbent when she shares the voter's preferences. When she does not, the effect on the quality of action choice is ambiguous—in that case, getting her to pay any attention to what impresses the voter might also lead to higher voter welfare. An additional problem with

term limits or hiding the action choice is that each reduces the voter's ability to select good types for reelection. This creates a trade-off between getting better policy in the first period versus a better type in the second period.

What Evidence Is Relevant for These Models?

The predictions of modern political agency models flow from the interaction of the details of the policy problem faced by incumbents and the information available to voters. As such, the family of models is quite rich in terms of the behaviors it can accommodate. Unfortunately, this richness is often ignored when researchers interpret empirical results on voter behavior.

To get a sense of the interpretive difficulties, consider the work of Healy & Malhotra (2009). They examine, at the county level, policy toward natural disasters, in terms of both preparedness spending and relief spending. They argue that preparedness spending is effective at preventing disaster but is not rewarded electorally, and that, as a result, incumbents underinvest in preparedness spending, at substantial welfare costs to voters.

They base this conclusion on regressions that include fixed effects for counties and years, plus spending measures for preparedness and relief, and the actual level of damage. Unfortunately, it is unlikely that this procedure recovers the causal effect of preparedness spending. It is quite likely that the only observation voters have on the level of preparedness spending is the level of realized disaster. So finding that, holding disaster fixed, prevention spending has no impact on votes does not mean that preparedness spending, in fact, has no causal effect on votes.

Leaving this objection aside, Healy & Malhotra (2009) seem to interpret their results as evidence of irrationality on the part of voters. They write (p. 402):

> [Voters] are myopic in the sense that they are unwilling to spend on natural disasters before the disasters have occurred. An ounce of prevention would be far more efficient than a pound of cure, but voters seem only interested in the cure.

Furthermore, "[t]he resulting inconsistencies in democratic accountability reduce public welfare by discouraging reelection-minded politicians from investing in protection" (p. 402). A closer look at the modern theory of accountability, however, shows that neither of these conclusions follows from their evidence. To see this, it is useful to look at two models in turn, one to assess each conclusion.

On the subject of voter unwillingness to spend on prevention, the interpretive difficulty is obvious from a multitask perspective. One task is hardening sites against disaster, whereas the second task is providing relief to areas that suffer damage. The second of these tasks is almost certainly a less noisy way to influence voter beliefs—voters are unlikely to observe detailed breakdowns of spending, whereas the level of damage and the level of relief activity are easy to observe. Relief activity is a straightforward indicator of relief-provision effort, but the level of damage is affected by many factors outside of the incumbent's control. In this kind of environment, fully rational voters will respond more strongly to relief than to prevention. This has nothing to do with how willing voters are to have public funds spent on prevention; rather, it is purely a reflection of which outcomes are most informative about the future performance of the incumbent. Indeed, it is easy to read Healy & Malhotra's (2009) results as a confirmation of the multitask model with fully rational voters.

A different model calls into question the claim that the voting rule Healy & Malhotra identify reduces voter welfare. Bueno de Mesquita (2007) studies a model with a very similar structure to that of the preparedness/relief problem, with three salient differences. First, and least important for our purposes, his disasters are caused by strategic terrorist groups. Second, there are three tasks:

observable prevention, unobservable prevention, and unobservable corruption. Third, he allows the voter to commit to the ex-ante optimal reelection rule. He finds that this optimal rule leads to a deviation from the allocation recommended by a standard cost-benefit analysis, devoting more effort to observable prevention than to unobservable prevention. Intuitively, biasing prevention effort in favor of observable activities makes it easier to control corruption. Applied to the disaster case, this suggests that a bias toward relief spending, for rational or irrational reasons, might enhance voter welfare if incumbents have the opportunity to engage in corruption while funding infrastructure projects.

There is a subtle point here about the relationship between causal inference and the evaluation of the theory. Many sources of random variation, e.g., natural disasters, help to identify causal effects that are only minimally relevant to testing political agency models. What is needed for testing is variation that identifies the causal effects corresponding to the comparative statics of the theory. The following section develops some of these comparative statics, and looks in depth at some of the empirical studies that confront them directly with data.

COMPARATIVE STATICS AND EVIDENCE

The basic theory of accountability developed above has many implications for how institutional variation affects behavior. A recent empirical literature has explored several of these implications in detailed micro studies. These quantitative case studies allow researchers to exploit knowledge about institutional and contextual features to identify plausibly exogenous variation in the features that the theory identifies as important determinants of responsiveness. Any one such study, of course, has sacrificed external validity for the internal validity that comes with credible identification of exogenous variation. But the accumulation of many such studies from different countries, levels of government, and policy problems, all broadly consistent with the theory, creates a strong case that the theory has latched on to important causal mechanisms in the world.

This section focuses on predictions about how responsiveness should vary with changes in the informational environment and in the noncompetence factors of reelection. The comparative statics on informational environments are particularly important for the normative perspective developed above. Rational voters are led to adopt reelection rules that create normatively undesirable incentives because those rules are best responses to particular informational environments. We should have more faith in the analysis if variation in actual informational environments leads to changes in behavior that track the theoretically predicted comparative statics.

Most models considered in the literature have similar, and very sharp, predictions for such informational changes. The voter observes a noisy measure of the incumbent's action. And the likelihood of responsiveness is decreasing in the amount of noise. Substantively, this says that responsiveness will be higher when incumbents have more influence (relative to other factors) on outcomes, and when third parties, perhaps the media or challengers, provide more information about hard-to-observe policy outcomes. Similar predictions come from considering the threshold v. Increases in the absolute value of the mean of this advantage reduce the likelihood of a responsive equilibrium. Increases in the variance of the threshold have more subtle effects: if the mean is close to zero, responsiveness is decreasing in the variance, whereas if the mean is very large, responsiveness is actually increasing in the variance.

These effects on responsiveness of variances, both of the noise and of the threshold, have an intuition familiar from models of targeted redistribution. When deciding whether or not to take the costly action $a = 1$, the incumbent thinks about the probability that this decision will be pivotal in whether or not she wins the election. Looked at in this light, it is not surprising that the effect of increasing the variance depends on the size of the expected advantage. When the

advantage is large, the action will be pivotal only when the random components of the vote lead to a relatively large swing. That means that incentives to take action $a = 1$ increase in the likelihood of these large swings, namely, when the variance is large. On the other hand, when the advantage is small, large swings make the action less likely to be pivotal, and so reduce the incentive to be responsive.

Empirical Evidence on Variation in Information

The models of accountability link exogenous changes in the informational environment to changes in incumbent behavior through several steps. Changes in information affect how responsive the voter is to what he observes, and those changes in turn change the incentives of incumbents. Several interesting studies use field experiments to directly examine the voter response part of this chain (see Pande 2011 for a survey of such work in developing countries). An interesting study with naturally occurring variation is carried out by Berry & Howell (2007), who exploit contextual knowledge about South Carolina school board politics to examine this comparative static prediction. Starting before the 2000 election cycle, South Carolina instituted a statewide school accountability system, as part of which school-level standardized test scores are made public. During the 2000 election cycle, schools reported raw scores to parents, but before the 2002 election cycle, they largely changed over to reporting the school's standing in broad categories, which conveyed significantly less information than the raw scores. In addition, between 2000 and 2002, the tone of media coverage about the tests changed from neutral to negative, emphasizing problems in implementation of the accountability program and concerns about "teaching to the test." Overall, it is reasonable to think that voters had less useful information about school board–relevant performance in 2002 and 2004 than they did in 2000.

Consistent with the theoretical prediction given these changes, voters were more responsive to test scores in 2000 than later. In 2000, moving from the 25th to the 75th percentile of increase in test scores was associated with 3 percentage points more votes for the incumbent. In 2002 and 2004, on the other hand, the effects were negative and imprecisely estimated. There is strong evidence that incumbents responded strategically to these different electoral environments. In 2000, but not 2002 or 2004, changes in test scores strongly predicted both standing for reelection and whether the incumbent ran unopposed.

Snyder & Strömberg (2010) exploit mismatches between the geography of congressional districts and the geography of media markets in the United States to trace the entire process by which more or less media coverage leads to more or less responsiveness to reelection concerns. To see the idea behind their research design, consider a large city and its suburbs. It's likely that many of the suburban voters get their local news from the big-city newspaper, even though they live in congressional districts that do not include the city. If the newspaper preferentially covers representatives of the city district where it is located, then city residents get more media coverage of their representative than do residents of the suburbs. If the accountability models are correct, this will lead to greater responsiveness from the city representative than from the suburban representative.

Snyder & Strömberg's first step is to build a comprehensive measure of congruence between newspaper markets and congressional districts. They then trace the process that links media coverage to policy choices. First, they show that more congruence leads to more coverage of the incumbent. This increase is consequential; more congruence is associated with greater knowledge about the incumbent. Importantly, more congruence does not lead to more knowledge about senators, suggesting that congruence is not just proxying for political knowledge in general. Congruence also leads citizens to be more likely to turn out to vote. Turning to incumbent behavior, Snyder & Strömberg find that incumbents in more congruent districts are more active in

committee hearings and vote on roll calls in a less partisan way. Finally, congruence seems to feed all the way through to policy: federal spending in a district is greater when there is more congruence.

Taken together, Snyder & Strömberg's results constitute a powerful case for an accountability view of elections. Not only is the reduced form relationship between congruence and the policy as the model predicts, each component of the theoretical story shows up in the data. One additional fact strengthens this conclusion. The selection mechanism implicit in political agency models predicts an incumbency advantage. Ashworth & Bueno de Mesquita (2008) show that, assuming incumbent elections and open seat elections get comparable coverage, this incumbency advantage is larger when media coverage is stronger. And, indeed, Snyder & Strömberg find that the incumbency advantage is larger in more congruent districts.

Lim et al. (2010) perform a similar exercise for judges on state courts. They calculate the congruence between newspaper markets and districts for state criminal courts. They find that greater congruence is associated with greater coverage of judges, and with a greater match between the severity of sentences and citizen preferences about how severe sentences should be (measured by votes on ballot propositions). Although they do not have the same data on intermediate steps of the process, such as voter knowledge, that Snyder & Strömberg (2010) have for congressional districts, Lim et al. do provide powerful evidence that reelection is important for the effects they find. In districts where judges are appointed, there is no relationship between congruence and either the harshness of sentencing or the relationship between voter preferences and harshness.

Another important channel through which voters get information on incumbent performance is challengers in elections. Models of this relationship have considered the direct revelation of information by challengers (Ashworth & Shotts 2011) and an indirect channel, in which the very fact of entry is a costly signal that informs voters that the incumbent has underperformed (Gordon et al. 2007).

Combining the idea of challengers as monitors with a political agency model predicts that incumbents who will stand for reelection against a challenger will be more responsive than incumbents who face voters without a challenger. Testing this proposition is not easy. In most cases, an incumbent who runs unopposed does so because no one decided to challenge her. It strains credulity to think that those incumbents are comparable to those who face stiff challenges. Judicial elections provide a unique laboratory for examining this issue, as many judges face retention elections, i.e., elections in which voters get to vote on whether or not to keep the incumbent without getting to choose the replacement. (If the incumbent loses, the governor appoints a replacement.) Kansas is particularly useful, as its 31 judicial districts are split between those that use retention elections and those that have partisan competitive elections.

Gordon & Huber (2007) argue that the enhanced electoral incentives have a clear behavioral prediction for judges making sentencing decisions in criminal cases: Be more punitive. This incentive arises from an asymmetry in risks borne by a judge who makes a sentencing decision. Incorrect decisions to imprison are bad, but their badness is largely hidden from voters. Incorrect decisions not to imprison, on the other hand, create the possibility of recidivism, an event that might be seized upon by a challenger to argue that the incumbent is soft on crime. As a result, Gordon & Huber predict that judges in competitive districts will be more likely to send convicts to prison than judges in retention districts.

They test this prediction with case-level sentencing data. Kansas has a guideline-based system for minimum sentences, although judges have discretion to impose harsher sentences. This allowed Gordon & Huber to control both for the guideline sentence and for a rich set of additional case-level covariates. Their estimates suggest that switching from retention elections to competitive elections raises the probability of incarceration by about 3 percentage points, from a baseline of 18%.

Lim (2011) also looks at judges in Kansas. Unlike Gordon & Huber, she focuses on partisan differences in the response to the enhanced electoral incentives created by competitive elections. In addition to showing that competitive-election judges are different than retention-election judges in terms of sentencing, she also documents differences within the set of competitive-election judges based on what district they represent. In particular, judges in the five liberal-leaning, competitive-election districts sentence more leniently than judges who face only retention. In the nine conservative-leaning districts, she replicates the Gordon & Huber (2007) finding of greater harshness for competitive-election judges. These election effects are largest for the incumbents who Lim (2011) estimates to have the highest returns to reelection.

The example of judges in Kansas is also important for demonstrating that the increased responsiveness might be normatively problematic. Gordon & Huber's theoretical expectation that judges will be more punitive when they face competitive elections already leans in this direction—there is no reason to expect that imprisonments are normatively desirable when they are driven by the fear of Willie Horton–style attack ads. Park (2012) provides some evidence that we should be concerned about the increase in punitiveness. He finds that virtually the entire burden of increased punishment in Gordon & Huber's data is from enhanced punishment of African American defendants.

Empirical Evidence on Variation in Partisanship

The informational comparative statics are driven by changes in the likelihood that the incumbent's action is pivotal in her reelection bid. Ashworth & Bueno de Mesquita (2006) explore a similar intuition about partisanship in a multitask framework. They show theoretically that incumbents in less partisan legislative districts, who expect to face more competitive elections, will substitute away from policy making in favor of constituency service, which is more easily observed by voters and thus a more efficient way to affect the voter's expectations about competence.

Dropp & Peskowitz (2012) empirically examine these predictions. Doing so is tricky, as standard measures of closeness lead to measurement error. Specifically, they show that if the model is correct, then controlling for partisanship with lagged vote share attenuates the effect of contemporaneous constituency service. To get around this, they collect data at the Texas legislative district level on partisanship in judicial and regulatory commission elections to use as proxies for partisanship in the legislative elections. Representatives of less partisan districts, according to this measure, were more likely to respond to experimentally generated requests for constituency service, as the theory predicts. Furthermore, Dropp & Peskowitz go on to show that more expected closeness does lead to reduced authorship of bills. Even the null results in this last analysis are what the theory, combined with a plausible view of legislative work, implies: Authoring a bill requires (member and staff) effort in way that, say, cosponsorship does not, so the predicted substitution is much greater in one case than the other.

TERM LIMITS

The most elementary prediction of the accountability models is that a term-limited incumbent, who cannot derive any benefit from impressing the voter, will not be responsive to voters. So a first thought about testing these models is to compare candidates who face binding term limits to those who do not. But the prediction is an all-else-equal one, and actually holding all else equal in empirical work is surprisingly challenging, for reasons that the model itself explains.

In particular, the models imply that winners of reelection are different, on average, than candidates in open seat elections. Consider, for example, a system that limits each politician to only two

terms. An incumbent who faces a binding term limit has been elected twice, whereas an incumbent who does not face a binding term limit has been elected only once. In the accountability model, the voter reelected the first incumbent after her first term because her expected type was above average. Thus the term-limited incumbent has a better type than the non-term-limited one, at least on average. As we normally think that type affects performance, this creates an important confound for simple comparisons designed to use term limits to capture the incentive effect.

In addition, there are standard concerns about omitted variable bias. For example, introducing term limits into a state might change the set of candidates who seek office. The question of whether performance is better in the first or second period thus comes down to the question of how the incentive effect compares to the electoral selection effect, holding the pool of potential candidates fixed.

Empirical Evidence on Term Limits

Ferraz & Finan (2011) study term limits and incumbent behavior using an innovative data set derived from a program of anticorruption audits carried out by the Brazilian central government. The program randomly selected municipalities to audit, producing measures of total grants to the municipality and cash values for all corrupt practices found. Ferraz & Finan find that term-limited mayors are about 1.9 percentage points more corrupt than other mayors. They then do several exercises to explore omitted variables and selection effects. To rule out omitted municipality-level covariates, they do a regression discontinuity analysis, comparing municipalities where a first-term incumbent barely won the last election to those where the first-term incumbent barely lost the last election. In the first case, policy is made by a lame duck, and in the second, by an incumbent who has an incentive to impress the voters. They also address the problem of electoral selection by restricting the analysis to incumbents who in fact won reelection. Both of these exercises yield estimates roughly twice the size of the reduced form effect. Finally, they show that the effects are much smaller for hard-to-observe corruption. All of this is what we would expect under a model in which there is a mix of honest and dishonest incumbents, and when there are reelection incentives, some of the dishonest ones pool with the honest.

Ferraz & Finan (2008) make further use of this data set to examine the actual response of voters to corruption. In the model of selection for honesty, incentives come from the fact that voters will try to estimate corruption and retain the incumbent only if they believe it is unlikely that she engaged in corruption. It would be hard to buy into this mechanism if voters did not punish incumbents when they were presented with news that the incumbent had, in fact, been corrupt. To examine this, Ferraz & Finan compare municipalities audited before and after an election. The audit actually raised the chance of reelection for incumbents with no violations but lowered the chance of reelection for incumbents with two or more violations. (One violation was the mean for the incumbents in the municipalities audited after the election, so it is a reasonable baseline level for voters to use as neutral news.) More radio penetration seems to magnify the effects, although this is more speculative because radio penetration was not randomized.

The most widely cited study to use term limits to examine electoral accountability is Besley & Case (1995a). They look at governors in states with term limits and compare outcomes in terms that are and those that are not limited. They find that both taxes and spending are higher in terms when the incumbent cannot run for reelection. In a follow-up study that extended the sample of governors by an additional decade, Besley & Case (2003) find that the effect of term limits vanished in the more recent data.

Alt et al. (2011) reexamine the data on governors and economic performance in an attempt to explicitly disentangle the incentive and electoral selection effects identified by standard political

agency models. Their point of departure is that states differ in whether their term limit laws allow for two terms or only for one. This variation sets up two comparisons, each of which isolates a different effect. A first-term governor who can be reelected and a first-term governor who is term-limited have each won election once, and so faced the same selection pressures, but have different incentives to impress the voters. Comparing these governors isolates the incentive effect. A term-limited governor who has won election once and a term-limited governor who has won election twice both face the same incentives, but with different histories of selection. Comparing them isolates the electoral selection effect.

Alt et al. (2011) estimate these effects on a panel of states from 1950 to 2000, using as outcomes per capita spending and taxes (for comparison to Besley & Case), growth in income relative to other states, and also a measure of the state's borrowing cost, designed to capture financial markets' views of the state's economic position. They use within-state variation in term-limit regimes to identify the different effects of incentives and selection. Both effects are sizable in magnitude. Moreover, the effects are about the same size. This equality implies that first-term and second-term incumbents in a two-term-limit state will have similar levels of performance, so a simple comparison of term-limited and not-term-limited incumbents in a two-term-limit state will not find much difference. This provides a nice resolution of the Besley-Case puzzle of declining effect, as there has been a strong trend away from one-term limits to two-term limits in the states.

There is no reason to expect the two effects to always be similar in magnitude. de Janvry et al. (2010) exploit a Brazilian conditional cash transfer designed to keep kids in school. The program was funded by the central government but administered by local officials. There were typically not enough funds to enroll all eligible children, so de Janvry et al. are able to do a careful program evaluation at the municipal level. They find that the program was successful on average, decreasing dropout by about 8%, but with substantial heterogeneity across municipalities. They then turn to politics and find that mayors with reelection concerns have about 2 percentage points better performance than term-limited mayors, representing a 36% better performance. Thus it seems that, in this context, the incentive effect is significantly larger than the selection effect.

Other Variation in Rewards

The prediction that term-limited incumbents will not be responsive generalizes to the prediction that incumbents will be more responsive, the greater the reward to winning is. Two recent studies use regression discontinuity designs to test this proposition.

Ferraz & Finan (2011b) exploit a discontinuous function relating municipal population to maximum wages for Brazilian municipal legislatures. Wages, but not other municipal characteristics, jump up at the population thresholds that allow for higher wages. These jumps are associated with different performance in the form of more bills proposed, more bills approved, and more resources for municipal schools and health clinics. There are two plausible explanations for the change in performance. One, captured in the models considered above, is that higher wages imply a greater payoff to reelection and, thus, lead any given incumbent to work harder. The other, not captured in those models, is that higher wages attract more qualified individuals to seek office. The evidence points to both stories capturing parts of the improvement. The higher wages at the discontinuities are associated with more schooling and higher skill for politicians. These characteristics of incumbents do have predictive power for outcomes, but they do not attenuate the effect of the higher wages. This suggests that neither the characteristics nor any omitted variable correlated with them are the channel that links wages to performance, and that the incentive effect is important.

Gagliarducci & Nannicini (2011) exploit a similar population-based discontinuity in wages for politicians, this time for mayors in Italy. Discontinuous increases in wages are associated with

discontinuous increases in the education of both candidates and winners. In addition, there is a discontinuous drop in the size of government, one that seems driven by enhanced efficiency of service provision. Gagliarducci & Nannicini do not, however, find evidence for an incentive effect from reelection. Among the population of candidates who served two terms, there is no difference in the effect of higher wages between first-term and second-term performance. How to reconcile this finding with the other evidence about the incentive effects of elections is an interesting topic for further exploration.

INTERACTIONS WITH THIRD PARTIES

The recent empirical work on the basic accountability mechanism canvassed above has grown in parallel with further developments in the theory. A particularly important strand of the theoretical literature uses political agency models to study interactions among political elites, including media coverage of incumbents and the separation of powers. The focus of these studies is typically whether the interaction leads to better or worse policy choice, a pressing issue given that there is no presumption that incumbents optimally serve voter interests in the equilibrium of a political agency model.

Different assumptions about the authority over policy and the incentives of the additional elite actor give rise to a rich set of modeling possibilities. For example, Ashworth & Shotts (2010) model a media outlet as a third party that can criticize, but not veto, policy in a pandering model. In their model, a newspaper gets private information about which policy best serves the voter's interests, and it acts as a commentator, making statements about whether the incumbent chose the correct policy. The newspaper gets to observe the policy choice before commenting, so it may, when it gets relatively weak evidence that the incumbent chose the wrong policy, act as a yes man, herding on the incumbent's choice. The newspaper comments before the next election and is thus relevant for the voter's decision to retain or replace the incumbent, creating the prospect that media commentary might discipline an electorally motivated incumbent's policy choice.

This newspaper does not always eliminate pandering; introducing a fallible media can actually introduce pandering that would not occur without a media commentator. An additional surprising result is that a yes-man media is in many circumstances more effective than a truthful media at reducing pandering.

To understand this, notice that pandering arises because voters treat the incumbent's possible policy choices asymmetrically. They reelect an incumbent who chooses a popular action unless he is proved wrong, but they only reelect an incumbent who chooses an unpopular action if he is proved right. If, at the time of the next election, it is unclear which policy choice was correct, voters will reelect the incumbent if and only if he chose the initially popular policy. This asymmetric burden of proof creates the incentive for pandering—if the public is sufficiently unlikely to learn whether the incumbent's policy choice was correct, then choosing the action with a lower burden of proof is optimal, even when that action is unlikely to be correct.

The media eliminates pandering when it induces the voter to treat initially popular and initially unpopular actions symmetrically. For example, in some circumstances the incumbent is reelected unless his action is somehow shown to be incorrect, either by a clear public signal that he indeed chose the wrong policy or in the absence of such public information, by informative media commentary criticizing the incumbent's action. The two actions thus lead to the same burden of proof, and the incumbent has no incentive to pander.

Focusing on asymmetric voter responses also helps explain the potential benefit of a yes-man media, one unwilling to criticize the incumbent unless it observes overwhelmingly clear information that the incumbent chose the wrong policy. When the media is a yes man, its contradictory

reports are definitive; i.e., voters know that the media criticizes the incumbent only when it is sure that he chose the wrong policy. As a result, voters treat an incumbent criticized by the media after taking the popular action exactly as they treat an incumbent who is criticized after taking the unpopular action. This symmetry gives a yes-man newspaper a leg up on eliminating pandering. If the media does not act as a yes man, then its contradictory reports are not definitive, and the incumbent may have a prior reputation sufficiently superior to that of his challenger that he can win in the face of media criticism of the popular action, but not the unpopular one.

Fox & Stephenson (2011) model judicial review with a third party who can veto the incumbent's policy choice but, being appointed for life, has no career concerns of his own. They set up their model so that without judicial review, low-quality incumbents have an incentive to "antipander"— choosing the policy that, ex-ante, is less likely to be optimal for voters. This incentive arises because going against the prior in that way could be optimal for a policy maker with no reelection concerns only if she were actually high-quality. Thus antipandering is an attempt to mimic the high-quality type.

Adding judicial review to the game has an ambiguous effect on the likelihood of antipandering. On the one hand, low-quality incumbents face a lower policy cost of antipandering if courts might rescue them from their mistakes. This bailout effect works to increase democratic failure. On the other hand, having her policy overturned is bad news for the incumbent—voters increase their probability assessment that the incumbent is low-quality.

In addition to speaking to the welfare effects of judicial review, the analysis also has implications for debates by legal scholars about what they call "substantive judicial review." First, in equilibrium, courts will be unpredictable, sometimes because no pure strategy equilibrium exists and sometimes because there are multiple equilibria. Thus unpredictability, often discussed as a negative effect of short-termism, is actually a consequence of instrumentally rational and forward-looking judicial behavior. Second, the analysis sheds new light on doctrines of "justiciability," such as standing doctrines or the refusal of courts to consider political questions. In the model, judicial review can lower voter (and judge) welfare if the bailout effect is strong enough. In such a case, judges have an incentive to develop doctrine that limits judicial review.

This basic framework can be developed in many directions. Fox & Van Weelden (2010) consider a veto player who has an incentive to appear competent and, in addition, might have partisan preferences between the incumbent and the challenger. Fox & Jordan (2011) ask how the possibility of delegation to a more informed bureaucrat can allow the incumbent to relax her incentives from voters via plausible deniability. Other potentially fruitful areas for application include electoral incentives in federal systems and electoral incentives for decisions in the international arena.

LOOKING AHEAD

The past decade has seen an explosion of research on the accountability relationship in elections. Theoretical work has clarified the relationship between formal accountability, responsiveness of incumbents to voters, and voter welfare. And careful empirical work has found strong support for the theory in a wide variety of contexts, both American and comparative.

Much remains to be done, of course. To my mind, the most exciting future direction comes from something implicit in the work on third-party monitoring in pandering models. The substantive conclusions of these studies are certainly interesting, but they are perhaps even more interesting as a model of how to integrate, in a theoretically serious way, the study of elites and the mass of citizens. These models feature the interaction of elites whose payoffs depend on the expected responses of voters. And those responses are themselves endogenously determined in equilibrium. Thus we can see the prospect of models of the joint determination of policy making and electoral

outcomes in a world with opinion leaders, cue taking from elites, or other mechanisms discussed in the literature on political behavior.

Several areas call out for work in developing this perspective. One is the development of tractable models of policy making that go beyond simple veto models. Equally important is work exploring alternatives to the standard approach to voters, who do Bayesian updating according to the correct model. Empirical evidence is mixed on this approach: Voters seem to do a good job of relative performance evaluation on the economy (Besley & Case 1995b, Kayser & Peress 2011) but do less well with filtering out factors like oil price shocks (Wolfers 2007). Recent work in psychology and economics provides resources for modeling information processing in ways that deviate from the Bayesian rational expectations standard; incorporating these factors into political agency models, with and without third-party monitoring, will be an important challenge in future.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

Alt J, Bueno de Mesquita E, Rose S. 2011. Disentangling accountability and competence in elections: evidence from US term limits. *J. Polit.* 73(1):171–86

Ashworth S. 2005. Reputational dynamics and political careers. *J. Law Econ. Organ.* 21(2):441–66

Ashworth S, Bueno de Mesquita E. 2006. Delivering the goods: legislative particularism in different electoral and institutional settings. *J. Polit.* 68(1):169–79

Ashworth S, Bueno de Mesquita E. 2008. Electoral selection, strategic challenger entry, and the incumbency advantage. *J. Polit.* 70(4):1006–25

Ashworth S, Shotts KW. 2010. Does informative media commentary reduce politicians' incentives to pander? *J. Public Econ.* 94(11, 12):838–47

Ashworth S, Shotts KW. 2011. *Challengers, democratic contestation, and electoral accountability*. Unpublished manuscript

Austen-Smith D, Banks J. 1989. Electoral accountability and incumbency. In *Models of Strategic Choice in Politics*, ed. PC Ordeshook. Ann Arbor, MI: Univ. Mich. Press

Banks JS, Duggan J. 2008. A dynamic model of democratic elections in multidimensional policy spaces. *Q. J. Polit. Sci.* 3:269–99

Banks JS, Sundaram RK. 1993. Moral hazard and adverse selection in a model of repeated elections. In *Political Economy: Institutions, Information, Competition, and Representation*, ed. WA Barnett, MJ Hinich, NJ Schfield. Cambridge, UK: Cambridge Univ. Press

Barro R. 1973. The control of politicians: an economic model. *Public Choice* 14:19–42

Berry CR, Howell WG. 2007. Accountability and local elections: rethinking retrospective voting. *J. Polit.* 69(3):844–58

Besley T. 2006. *Principled Agents?: The Political Economy of Good Governement*. New York: Oxford Univ. Press

Besley T, Case A. 1995a. Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. *Q. J. Econ.* 110(3):769–98

Besley T, Case A. 1995b. Incumbent behavior: vote-seeking, tax-setting, and yardstick competition. *Am. Econ. Rev.* 85(1):25–45

Besley T, Case A. 2003. Political institutions and policy choices: evidence from the United States. *J. Econ. Lit.* 41(1):7–73

Bueno de Mesquita E. 2007. Politics and the suboptimal provision of counterterror. *Int. Organ.* 61:9–36

Canes-Wrone B, Herron MC, Shotts KW. 2001. Leadership and pandering: a theory of executive policymaking. *Am. J. Polit. Sci.* 45:532–50

Canes-Wrone B, Shotts KW. 2004. The conditional nature of presidential responsiveness to public opinion. *Am. J. Polit. Sci.* 48(4):690–706

Coate S, Morris S. 1995. On the form of transfers to special interests. *J. Polit. Econ.* 103(6):1210–35

Daley B, Snowberg E. 2011. Even if it is not bribery: the case for campaign finance reform. *J. Law Econ. Organ.* 27(2):301–23

de Janvry A, Finan F, Sadoulet E. 2010. *Local electoral accountability and decentralized program performance*. Unpublished manuscript

Dropp K, Peskowitz Z. 2012. Electoral security and the provision of constituency service. *J. Polit.* 74(1):220–34

Duggan J. 2000. Repeated elections with asymmetric information. *Econ. Polit.* 12(2):109–35

Fearon JD. 1999. Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance. In *Democracy, Accountability, and Representation*, ed. A Przeworski, S Stokes, B Manin. New York: Cambridge Univ. Press

Fearon JD. 2011. Self-enforcing democracy. *Q. J. Econ.* 126(4):1661–708

Ferejohn J. 1986. Incumbent performance and electoral control. *Public Choice* 50:5–26

Ferraz C, Finan F. 2008. Exposing corrupt politicians: the effect of Brazil's publicly released audits on electoral outcomes. *Q. J. Econ.* 123(2):703–45

Ferraz C, Finan F. 2011. Electoral accountability and corruption: evidence from the audits of local governments. *Am. Econ. Rev.* 101(4):1274–311

Ferraz C, Finan F. 2011b. *Motivating politicians: the impacts of monetary incentives on quality and performance*. Unpublished manuscript

Fiorina MP. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale Univ. Press

Fox J, Jordan SV. 2011. Delegation and accountability. *J. Polit.* 73(3):831–44

Fox J, Shotts KW. 2009. Delegates or trustees? A theory of political accountability. *J. Polit.* 71(4):1225–37

Fox J, Stephenson MC. 2011. Judicial review as a response to political posturing. *Am. Polit. Sci. Rev.* 105(2):397–414

Fox J, Van Weelden R. 2010. Partisanship and the effectiveness of oversight. *J. Public Econ.* 94(9, 10):684–97

Fox J, Van Weelden R. 2012. Costly transparency. *J. Public Econ.* 96:142–50

Gagliarducci S, Nannicini T. 2011. *Do better paid politicians perform better?: Disentangling incentives from selection*. Unpublished manuscript

Gelbach S. 2007. Electoral institutions and the national provision of local public goods. *Q. J. Polit. Sci.* 2(1):5–25

Gordon SC, Huber GA. 2007. The effect of electoral competitiveness on incumbent behavior. *Q. J. Polit. Sci.* 2(2):107–38

Gordon SC, Huber GA, Landa D. 2007. Challenger entry and voter learning. *Am. Polit. Sci. Rev.* 101(2):303–20

Healy A, Malhotra N. 2009. Myopic voters and natural disaster policy. *Am. Polit. Sci. Rev.* 103(3):387–406

Healy AJ, Malhotra N, Hyunjung C. 2010. Irrelevant events affect voters' evaluations of government performance. *Proc. Natl. Acad. Sci. USA* 107(29):12804–9

Holmström B, Milgrom P. 1991. Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *J. Law Econ. Organ.* 7:24–52

Kayser MA, Peress M. 2011. *Benchmarking across borders: electoral accountability and the necessity of comparison*. Unpublished manuscript

Key VO. 1966. *The Responsible Electorate: Rationality in Presidential Voting, 1936–1960*. Cambridge, MA: Belknap

Lim CSH. 2011. *Turnover and accountability of appointed and elected judges*. Unpublished manuscript

Lim CSH, Snyder JM, Strömberg D. 2010. *Measuring media influence on US state courts*. Unpublished manuscript

Lohmann S. 1998. An informational rationale for the power of special interests. *Am. Polit. Sci. Rev.* 92(4):809–27

Maskin E, Tirole J. 2004. The politician and the judge: accountability in government. *Am. Econ. Rev.* 94(4):1034–54

Meirowitz A. 2007. Probabilistic voting and accountability in elections with uncertain policy constraints. *J. Public Econ. Theory* 9(1):41–68

Pande R. 2011. Can informed voters enforce better governance? Experiments in low income democracies. *Annu. Rev. Econ.* 3:215–37

Park K. 2012. *Judicial prejudice or politics: which explains black-white gaps in incarceration rates?* Unpublished manuscript, Harris School of Public Policy

Persson T, Roland G, Tabellini G. 1997. Separation of powers and political accountability. *Q. J. Econ.* 112:1163–202

Persson T, Tabellini G. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press

Prat A. 2005. The wrong kind of transparency. *Am. Econ. Rev.* 95(3):862–77

Scharfstein DS, Stein JC. 1990. Herd behavior and investment. *Am. Econ. Rev.* 80(3):465–79

Schwabe R. 2009. *Reputation and accountability in repeated elections*. Princeton, NJ: Princeton Univ. Typescript

Seabright P. 1996. Accountability and decentralisation in government: an incomplete contracts model. *Eur. Econ. Rev.* 40:61–89

Shi M, Svensson J. 2006. Political budget cycles: Do they differ across countries and why? *J. Public Econ.* 90(8–9):1367–89

Snyder JM, Strömberg D. 2010. Press coverage and political accountability. *J. Polit. Econ.* 118(2):355–408

Snyder JM, Ting M. 2008. Interest groups and the electoral control of politicians. *J. Public Econ.* 92(3, 4):482–500

Wolfers J. 2007. *Are voters rational? Evidence from gubernatorial elections*. Unpublished manuscript

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Political Science* articles may be found
at http://polisci.annualreviews.org/