Accountability and Principal-Agent Models*

Sean Gailmard†

August 2012

Abstract

Principal-agent theory encapsulates a tradition of rational choice modeling, in which some actor(s) (the principal(s)) uses whatever actions are available, to provide incentives for some other actor(s) (the agent(s)) to make decisions that the principal most prefers. Because principal-agent theory focuses on the responsiveness of the agents decisions to the principal's goals, and how this responsiveness is mediated by actions available to each actor as well as institutional settings in which they interact, it is a natural framework to study accountability in political institutions. This essay gives a basic overview of principal-agent theory and briefly reviews its application in two domains of political science: bureaucratic accountability to higher-level political actors, and electoral accountability of representatives to constituents. The essay emphasizes that principal-agent theory is in fact a highly flexible family of models, rather than an overarching set of assumptions and results.

*Chapter prepared for the *Oxford Handbook of Public Accountability*; forthcoming, Oxford University Press.

†Associate Professor, Charles & Louise Travers Department of Political Science, University of California, Berkeley. Email: *gailmard@berkeley.edu*

Public accountability is a function of the capabilities of principals to judge the performance of their agents (Achen and Bartels 2002; Healy and Malhotra 2010; Lenz 2012; Lupia and McCubbins 1998). But it is also in part a function of institutions themselves. To make sense of the relationship between accountability and institutional structure, it is useful to make use of a theoretical framework that can express widely varying institutional details and express their consequences for accountability in commensurate terms. Principal-agent theory has become a widely used paradigm for analyzing public accountability. This is because it provides a flexible framework for modeling innumerable variations in institutional arrangements, and comparing their potential for inducing desirable behavior by agents.

Applications of principal-agent theory in the study of public accountability have become sufficiently common and widespread that it is hopeless to attempt a comprehensive review of the literature in so short a space, and this chapter will not attempt such a review. Instead it attempts to give a sense of the conceptual flavor of principal agent analysis, as well as insights gleaned from two of the original and still more common substantive areas of its application in political science – control of politicians through electoral institutions; and control of bureaucracies by legislative, executive, and/or judicial actors.[1] Thus this chapter is organized as follows. Section 1 briefly reviews the basic structure and central tenets shared by principal agent models. Section 2 discusses applications of principal agent models to electoral politics. Section 3 discusses insights gleaned from applications to bureaucratic accountability. Section 4 concludes.

[1]For reasons of space this essay therefore does not consider the large empirical literatures on public accountability inspired by principal-agent theory. I am not aware of any published review of this empirical literature as a whole, but Besley and Case (2003) review significant portions, particularly those related to elections and selection of policy makers. Miller (2005) presents an insightful review of both theory and empirical research, focused particularly on the bureaucracy.

1 Principal-Agent Theory: An Overview

Analysis and evaluation of public accountability requires a specification of who is (or is supposed to be) accountable to whom. This is a core ingredient of principal-agent theory. In principal-agent models, some actor (or group of actors) called an agent undertakes an action on behalf of another actor (or group of actors) called a principal. The principal, for its part, can make decisions that affect the incentives of the agent to take any of its various possible actions. This process of structuring incentives for the agent is the central focus of principal agent theory. The decisions made by the principal that structure the agent's incentives to take various actions constitute a contract, in the language of principal-agent theory, and principal-agent theory is often taken as a specific area of contract theory more generally (Bolton and Dewatripont 2004).

Perhaps the most elemental point about principal agent theory is that it is not in fact a single overarching theory with a specific set of assumptions or conclusions. Principal-agent theory is more accurately described as a family of formal models addressing related concerns with similar styles of analysis. It is not much of a stretch to suppose that for any given actors labeled "principal" and "agent," and any pattern of interaction between the two, a principal-agent model can be written down with that pattern as an equilibrium outcome – and modelers might consider it a parlor game of sorts to do it. Given that, one must be wary of claims to "test" principal-agent theory empirically in any broad sense. By the same token, it is always possible to defend any status quo interaction between a "principal" and an "agent" as reflecting the greatest degree of accountability to which the agent can be held by the principal, given various informational asymmetries and commitment problems.

Specifying a member of the principal agent family of models requires specifying (1) what the agent(s) can do and how this affects the principal(s), (2) what the principal(s) can do and how this affect the agent(s), and (3) who the principal(s) and the agent(s) are. In other words, principal agent models specify a set of actors, possible actions they can take, and how they evaluate consequences of those actions. In this respect a principal agent model is necessarily a game in the formal sense, and correspondingly principal agent models in contemporary literature are almost exclusively analyzed with the tools of noncooperative game theory.

While this chapter stresses the flexibility of principal-agent theory, this background suggests some limitations of it. First, it inherits the limitations of game theory as a tool for explaining behavior. Principal-agent analysis is also inappropriate for analyzing accountability of some actor to another, when the second is unaware of its dependence on the first and/or can do literally nothing to affect the behavior of the first. Finally, its very flexibility is also a sort of limitation. Within its domain of application there does not seem to be any pattern of behavior that a principal-agent model cannot explain. While any particular model in this family may have empirical content, it is not clear what content the family as a whole has.

A fundamental distinction in types of principal agent models is between those dealing with moral hazard or hidden actions, and those dealing with adverse selection or hidden information. In moral hazard problems, the agent takes one of several possible actions that affect the principal's utility, the principal and agent have different preferences over the possible actions the agent can take, and the principal cannot directly control the agent's action. However, to make the problem interesting the principal observes some information affected by or correlated with the agent's action, and administers a reward or punishment (e.g., a bonus payment, reelection) based on that information.[2]

In adverse selection problems, the agent is privy to some information that the principal needs to make a decision in her own interest, but the

[2]It is perfectly possible for a moral hazard problem to lack this feature, but then the agent simply takes its own preferred action irrespective of the interests of the principal.

agent prefers that the information be used differently. In pure adverse selection models, the principal may be able to specify costs the agent must incur to take various decisions, or even outright prohibit certain decisions by the agent or require a particular one. All of these types of direct control over the agent's action are unavailable in pure moral hazard models. In adverse selection, the problem for the principal is that it does not know how to use this control; it does not know which action to direct the agent to take.

Key to the analysis of both moral hazard and adverse selection problems is incentive compatibility. If the principal is unrestricted in the types of contracts it can offer the agent – e.g., in moral hazard, bonus pay for results that are correlated with the principal's preferred action – it is often possible for the principal to induce the agent to take the principal's own most preferred action. Incentive compatibility simply means that the principal must make it worth the agent's while to behave in this way.

In general, incentive compatibility imposes constraints on the principal: the principal must trade off the benefits of an improved decision from its own point of view, against the costs of inducing the agent to take that particular decision. If the agent also has an option to exit the relationship with the principal (an individual rationality or participation constraint), or if the agent's liability for poor performance is limited in some way, then in general incentive compatibility constraints imply that the principal will generally not induce the agent to pursue the principal's own most preferred course of action, even if the principal is able to do so in theory. The difference in the action the principal prefers to induce, given incentive compatibility (and participation) constraints, and the action the principal would take itself if it could (and had the information and capabilities of the agent) gives rise to agency loss. Principals must trade off agency loss against the cost of satisfying incentive compatibility. When these two costs to the principal are in direct conflict, as they typically are in principal-agent models, the principal generally does not wish to eliminate one cost at the expense of the other. Agency loss is what occurs when the agent gets a bit of extra slack from the principal to pursue its own interests rather than the principal's.

Overall, then, principal-agent models suggest that agents need not in fact be perfect agents of their principals. Agency loss is not necessarily inevitable, in the sense that principals may (in some models) be able to eliminate it if they for some reason prefer to do so. But even in these models, principals typically would find it prohibitively costly to eliminate agency loss completely, so they choose to tolerate some. Therefore, the mere existence of agency loss does not imply that accountability of the agent is suboptimal or defective from the principal's standpoint.

2 Electoral Agency

The plenary power of the state raises serious threats to the welfare of citizens. What is to prevent the political officials holding the reins of state power from behaving opportunistically in policy making, elevating their own interests over those of voters? The promise of electoral democracy is that voters can hold political officials accountable for their policy choices, and thereby ensure a close connection between public will and public policy. Setting aside imperfections in voters themselves (their attentiveness, their rationality, etc.), are electoral institutions themselves up to this task? That is, even in a world of perfect voter attention and rationality, do electoral institutions themselves limit the accountability of policy to voters interests?

This issue is exactly the one explored in principal-agent analyses of elections. Voters are the principal(s) and politicians are the agents.[3] One

[3]Dixit, Grossman, and Helpman (1997) present an interesting twist on this, in which politicians are agents but (multiple, competing) interest groups are the principals. This

of the classics of the genre is due to Ferejohn (1986), who developed a pure moral hazard model of electoral accountability that taps into key concerns about elections as instruments of control.[4] The concern of political opportunism and mitigating agency loss with elections is inherently dynamic, so Ferejohn's model involves multiple periods an infinite horizon to be exact. In Ferejohn's model an incumbent politician can exert effort on behalf of (homogenous, in the baseline case) citizens, who prefer more effort to less. But the value of effort changes randomly from one period to the next. This value is observed by the incumbent politician, but not by the voters, who only observe their overall utility in each period. For politicians effort is costly but holding office has some value irrespective of how much effort the politician supplies. In equilibrium voters use a simple retrospective voting rule: they reelect the incumbent for another period if their period t utility exceeds a specific threshold. But voters incur significant agency loss to politicians. If the value of effort to voters is too small, the incumbent shirks completely, knowing even very high effort will not be enough to put voters' utility over its retrospective threshold. If the value of effort to voters is large enough, politicians exert some effort but only the minimal amount necessary to be reelected. This effort gets smaller when the value of effort gets larger, because the incumbent can work less hard and still satisfy the voters' retrospective threshold.

Of course, the magnitude of agency loss in some environment can only be judged relative to an alternative institutional arrangement. In an important paper, Persson, Roland, and Tabellini (1997) show that separation of powers institutions combined with checks and balances can mitigate the agency loss identified by Ferejohn. In essence, these authors argue, separation of powers when political actors have inherent preference conflicts

illustrates the positive use of principal agent analysis without any normative connection to public accountability.

[4]See also Barro (1973) for a complete information, finite horizon moral hazard model tapping into a similar concern as Ferejohn's.

(as between a legislature and executive) can lead to better information for voters about the possible value of policy making, and therefore allows voters to better tailor their voting threshold to the situation at hand. Persson, Roland, and Tabellini also show that these beneficial effects of separation of powers depend on checks and balances – in their case, political actors must come together for the purpose of determining policy.[5]

Until the last decade or so, principal-agent analyses of the electoral connection focused on the moral hazard dimension. The focus of this literature was on elections as instruments of control over politicians' behavior, given their preferences and capabilities. Banks and Sundaram (1993, 1999) initiated the formal analysis of elections with both moral hazard and adverse selection. In an influential essay, Fearon (1999) noted that elections not only provide an instrument with which to sanction wayward politicians, but also an instrument for selecting politicians with desirable traits or preferences in the first place.

At first blush, the possibility that elections can serve two purposes, sanctioning and selection, may suggest that they are more effective at mitigating agency problems than previously thought based on moral hazard models alone. However, this turns out not to be the case. Note that in the classic Ferejohn moral hazard model, the voter has no trouble implementing the evaluation standard that most tightly constrains the politician. This is because the politician and her possible replacement are identical from the voter's standpoint neither is better at generating favorable results than the other. If the voter were to learn that one of the pool of politicians had a more desirable "type" on the dimension of hidden information, and expected that "type" to generate better results in the future, then the voter would have no rational course of action but to select that type regardless of their past performance. This is the case, for instance, in Canes-Wrone,

[5]Austen-Smith and Banks (1998) offer another extension dealing with incumbency advantage.

Herron, and Shotts (2001). Though voters face both adverse selection (selecting competent politicians) and moral hazard problems (inducing the politician's choice of the voter's preferred policy), the adverse selection dimension trumps the moral hazard one in equilibrium. The key point of their model is that rational behavior by voters, combined with their information asymmetry and limited instruments of control over politicians, can lead to "pandering" or the decision by politicians to neglect their own private information about socially desirable policies and pursue a course of action that the less-informed electorate considers beneficial.[6]

This is illustrative of a common theme in the electoral agency literature: demonstrating that perfectly rational behavior under the right (or wrong) constellation of information asymmetries and limitations on precommitment power can lead to specific kinds of agency loss. The focus of this sort of analysis is on electoral institutions themselves. Principal-agent models make the point that public accountability in electoral democracies is inherently limited, regardless of any imperfections voters may exhibit in their decision making process.[7] This also illustrates a useful feature of the rationality postulate in principal-agent models, quite apart from its verisimilitude or lack thereof: it is useful to know the outer limits of public accountability that can be induced by elections, because this reveals that the limitations of voters can only be blamed so much for accountability pathologies in democracies.

[6] Majumdar and Mukand (2004) offer a related analysis of the incentives of politicians to invest in socially desirable experimentation with new policies.

[7] See e.g. Fox and Shotts (2009), who present a model in which voters sometimes reelect politicians known not to share the voters' preferences; or Ashworth and Bueno de Mesquita (2009), who present a model of elections in an adverse selection problem in which incumbency advantage arises in equilibrium.

3 Bureaucratic Agency

One of the earliest, and still most robust, principal-agent literatures in political science takes bureaucrats as agents of some constellation of political principals – most often Congress, the president or executive actors, and/or courts. The 1970s saw profound scholarly disillusionment in political science, law, and economics with the bureaucratic policymaking apparatus created mostly since the Great Depression. In political science, scholars such as Theodore Lowi and Hugh Heclo argued that the bureaucratic institutions created by ostensibly well-meaning Congresses had become ungovernable, and that Congress had reached a point of abdicating its responsibilities to govern by creating one new bureaucracy after another while constraining them with little more than an directive to "Go forth and do good." Given this abdication, bureaucrats may or may not "do good," but they do so at best according to their own ideals and conception of what that means – not according to any vision that is guided by, or accountable to, Congress. Whatever doubts one may express about the democratic pedigree or legitimacy of the U.S. Congress, it certainly occupies a seat closer to the governed than most bureaucratic policy makers do, and so this interpretation of bureaucratic architecture throws into doubt the public accountability of much of modern social and economic policy.

William Niskanen (1971) articulated similar concerns. Niskanen advanced several interrelated postulates in an economic model of bureaucratic service delivery: that bureaucrats seek to maximize their budgets; that bureaucrats know the value that legislators attach to the services they provide; that bureaucrats know more than those legislators about the cost of services they provide; and that bureaucrats essentially make take it or leave it budget offers to the legislature. The implication of this set of assumptions, Niskanen showed, is that bureaucrats extract rents from the legislature. Bureaucrats are willing to do so, owing to the assumption about the bureaucrat's objective; they are able to do so, owing to the asymmetric information and proposal rights they enjoy (or are assumed to enjoy) over the legislature. Niskanen's model gave rise to a school of thought explaining the growth in government spending as well as the ostensible inefficiency of bureaucratic service provision. The implications for public accountability are obvious and serious: service delivery by specific bureaucracies, and government budgets as a whole, are not meaningfully held in check by Congress, and the actual government at the ground level is not accountable in any serious way to others than those making the decisions and spending the money.

The first wave of self-aware applications of principal-agent theory in the analysis of bureaucratic policy making took issue with the interrelated concerns raised by Lowi, Heclo, Niskanen, and others. The common theme these scholars articulated is that the U.S. Congress does not seem to be doing well by bureaucrats. The simple observation of Morris Fiorina (1979) was that it is difficult to square the assumption that Congress is good at minding its own interests (Mayhew 1974) with the argument that bureaucrats systematically undermine those interests. The reason is that, whatever the institutional arrangements by which bureaucrats interact with Congress, Congress designed them. It would be bizarre for Congress to actively design institutions that lead to the undermining of Congress's own interests, and yet that is what the disillusioned scholarship on bureaucracy through the 1970s seemed to suggest.

The germ of insight in Fiorina's argument stems directly from a principal-agent perspective on the situation. One should understand Congress as a principal and various bureaucrats as its agents. Therefore one should interpret bureaucratic institutions and legislative-bureaucratic interaction, inasmuch as these institutions and interactions are designed or affected by legislative principals, as promoting the interests of the principal to the greatest extent possible. This is the central premise of thought on bureaucratic institutions based on principal-agent theory. Left for debate are the precise meanings of the terms in that central premise most conspicuously, what exactly Congress can do to design or affect bureaucratic institutions, and what exactly is the meaning of "to the greatest extent possible."

A simple but also stark illustration of this reasoning comes from a reinterpretation of a long-observed regularity in legislative-bureaucratic interaction in the U.S.: that when agencies submit budget requests to the Office of Management and Budget in the Executive Office of the President, and the President in turn submits a budget proposal to Congress, Congress for its part often makes trivial modifications (or none at all) to the line items for particular agencies. Before the steam of principal-agent thought on the topic took hold in the literature, this pattern could be interpreted as providing grounds for Niskanen's assumption that bureaucrats have nearly monopoly proposal power over budgets. After all, Congress often leaves the proposed amounts in tact, and that is exactly what one would expect when an agenda setter has monopoly proposal rights and certain knowledge of the preferences of Congress (Romer and Rosenthal 1978). But it is also exactly what one would expect if bureaucrats are completely under the control of Congress and know that they cannot get away with budgeting one cent more than Congress wants them to spend – with full knowledge of the costs and benefits of their expenditures (Kiewiet and McCubbins 1991). Nevermind that Congress often, in fact, does modify executive budget proposals; even the supposed stylized fact that it does not is entirely consistent with a principal-agent perspective on bureaucratic accountability.

The school of thought on bureaucratic accountability known as "Congressional dominance" takes principal-agent-inspired thought to its outermost possible reaches. This specific perspective is most closely associated with McCubbins, Noll, and Weingast (1987, 1989), who elaborated an entire theory of administrative structure and procedure as a set of devices for

Congress to mitigate its information asymmetries with respect to agencies, and achieve the best possible outcomes in the bureaucratic policy process from its own point of view.[8] McCubbins, Noll, and Weingast offered a particularly insightful interpretation of the range of instruments available to Congress to influence choices made by bureaucratic agencies. For example, they pointed out that Congress can influence rights of standing with respect to the decisions of bureaucratic agencies (the right to sue agencies in court due to harms inflicted by their actions), as well as the jurisdiction of federal courts that review agency actions when they are sued (as they always are in important policy matters). In this way Congress can make it easier for favored interest groups to keep a close eye on and potentially help to rein in wayward agencies. In other words, Congress designs these administrative procedures not primarily out of concern with due process and such-like legal niceties, but rather out of a desire to mitigate agency loss due to moral hazard problems in the bureaucracy.

An important point about principal-agent perspectives on bureaucratic accountability, and one sometimes lost in translation from formal principal-agent models to the substantive issue of bureaucratic policy making (cf. Moe 1987), is that nothing in these models implies that Congress "gets its way" in interactions with bureaucrats. Indeed, the ubiquity of agency loss in formal principal-agent models, as noted in section 1 above, should be sufficient to indicate that this is not the case. In other words, principal-agent scholars of bureaucratic accountability must concede that in general some accountability of policy and administrative decisions to the interests of principals is lost in delegation. The formal concept of agency loss and the substantive concern of sacrificed accountability are inextricably linked.

[8] Also see Horn and Shepsle (1989), Horn (1995), and DeFigueiredo, Spiller, and Urbiztando (1999). Banks and Weingast (1992) present a formal model in which agency budget requests serve as signals of agency costs. While signaling games are not necessarily applications of a principal-agent perspective, Banks and Weingast's model was explicitly inspired by Niskanen's assumptions.

The best case that can be made for delegation on principal-agent grounds is not that there is no compromise in accountability; it is that this compromise is better for the principal than the alternatives. This is perhaps best clarified in a canonical formal model of delegation due to Holmström (1984), and subsequently elaborated by Epstein and O'Halloran (1994, 1999), Gailmard (2002), Huber and Shipan (2002), Volden (2002), and Bendor and Meirowitz (2004). The agent is presumed to have some expertise the principal does not have, represented as information the principal needs to know to choose the best decision. And the principal and agent have different "ideologies," specifically, different conceptions of the optimal result of the policy decision. The principal's instrument to influence the agent is to specify a set of policy decisions the agent is permitted to make, and allow the agent to pick freely from them. In the most basic variant of this model, once the agent so chooses, the game is over. When this model is represented formally as a game, its equilibrium involves the agent picking its own most preferred policy (the one that, given the agent's expert information, leads to its own most desired end result) whenever it is in the set of policies delegated by the principal. Since the principal and the agent, by assumption, have different ideal outcomes, this necessarily implies that the principal does not (ever, in equilibrium) obtain its most preferred outcome – but for at least some possible states of its information, the agent does.

Given this assumed information asymmetry and this device (controlling policy discretion) available to the principal to mitigate agency loss from that asymmetry, it is not possible for the principal to better hold the agent accountable to the principal's goals than the equilibrium specifies. If it were, then by assumption since the principal is assumed to be the best possible steward of its own interests, the "equilibrium" would not in fact be an equilibrium. Now the principal in this model certainly has the power to specify once and for all the decision the agent must make it can restrict the agent to choose one particular policy. In this way the principal can attain complete control over the decision made by the agent. The principal does not (generally) behave this way in the equilibrium of model, because it is counterproductive from the principal's own interests. Control per se over the agent is not the point; the point is for the principal to get as close as it can to its most preferred outcome, and by extension, to induce the agent to deploy its expertise as much as possible in the principal's interest. The principal sacrifices control because doing so gives the agent greater incentive to use its information because by doing so the agent is allowed to pursue its own interest. The degree of accountability attained by the principal in equilibrium may not be the maximal degree of control possible in any circumstance, and if effort or information is sought from the agent, is always constrained by the agent's incentive compatibility conditions.

Policy discretion in the Holmström model is only one way for a principal to address its hidden information problem with respect to its agent. Another possible approach is for the principal simply to ask the agent for a direct report of its information, which the principal could then use itself in crafting policy. For instance, rather than delegating policy authority to a regulatory agency that is putatively an expert on the policy area in question, Congress might hold hearings in which it elicits this expertise, and writes statutes around it that are better tailored to Congress's own interests and ideological stance than are the regulations promulgated by agencies pursuant to their delegated policy discretion. Obviously, this would quickly become an unwieldy and impracticable approach: for Congress to ask the Environmental Protection Agency to report in a hearing everything it knows about mitigating the impact of water pollutants in rivers, and reflect that digested knowledge in a statute, would likely not be a fruitful exercise. Setting aside this eminently sensible point for a moment reveals an interesting aspect of the character of principal agent models.

These two approaches to eliciting an agent's information, delegating policy discretion versus obtaining a direct report, can be directly and formally compared in terms of "how much" of the agent's information gets applied in policy making, and the linkage between policy decisions and the principal's goals (the key point in analysis of accountability). This is the subject of an insightful paper by Dessein (2002) and elaborations and applications by Gailmard (2009) and Gailmard and Patty (2013). In particular, the principal is better off by simply delegating authority to the agent and completely relinquishing control over the decision (within the agent's range of discretion), than by requesting information from the agent and making a decision based on the report. The reason is that in the latter arrangement, the principal knows that she cannot fully trust the agent to be truthful; for if she did, the agent would mislead her about the true state of the information so that the principal would act in the interests of the agent rather than the principal herself. But when the agent receives delegated authority from the principal, the agent can use all available information in pursuit of his own interests. So more information gets applied under the delegation arrangement, to the improvement of the principal's (and the agent's) utility.

What's more, the exact same logic reveals that, when delegation is used by a principal to elicit an agent's information, the principal may actually be better off if it cannot oversee or review the agent's decision in any way. For if the agent believes that the principal can revise its decision after the fact, the agent will not consider its decision to be an act of policy making with any finality, but at best an advisory opinion that the principal may set aside. The principal may of course be interested in the agent's initial choice to the extent that it conveys information that the agent knows better than the principal. But then a game of delegation with revision by the principal devolves into nothing more than a game in which the agent sends a report of its information to the principal, and the principal decides the actual policy accordingly. And we have already sketched the logic demonstrating that this arrangement is inferior from the principal's point of view than a (firm, irrevocable) delegation of policy authority to the agent.[9]

This toy example illustrates an important point about principal agent models. Depending on the nature of the contracting environment, information asymmetries, and incentive problems facing the principal with respect to the agent, the optimal behavior of the principal may not comport with commonsense notions of exerting control. In the above example, oversight and review by the principal is counterproductive. This does not explain how an actual political principal would go about tying its hands to keep them off the decisions of a bureaucratic agent, but it does hint that in some situations it might want to tie its hands if it can. Therefore, to observe a principal limiting its own ability ostensibly to control an agent does not in itself suggest a challenge to principal agent theory. A principal-agent model need only show that the putatively puzzling behavior of the principal is an equilibrium of some game to show that the principal is doing the best it can to guard its interests vis-à-vis the agent in the problem postulated in that model.

Situations in which agents must not only be induced to apply their information faithfully in the interests of the principal, but to acquire that information in the first place, are particularly ripe for this sort of twist. The reason is that if information is costly for agents to obtain, they must obtain some rents from information in order to be willing to invest in it. Apparently strange behavior by principals, such as conceding seemingly irrational degrees of autonomy over policy decisions agents care about,

[9] Bueno de Mesquita and Stephenson (2007) present an explicit model making a similar but deeper point about the value of limiting review of agency actions. However, their model turns on hidden actions by bureaucrats, rather than hidden information. Their model shows that when some agency effort is observable and some is not, ex post review may induce the agent to substitute observable effort for unobservable, to the detriment of the principal in some cases. See also the multitask moral hazard model of Holmström and Milgrom (1987) for a similar point in economic contracting.

can sometimes be interpreted as a de facto method for principals to guarantee that agents will obtain the necessary rents, and therefore make it worth the agent's while to invest in information (Gailmard and Patty 2007, 2013). As is apparent by now, principal-agent theory is an extremely flexible theoretical framework (rather than a theory in and of itself) that is more about how one interprets observed interactions, than about predictions of specific patterns of interaction that may or may not occur.

Recent work in political science has shown that principal-agent theory is even consistent with substantially less control of bureaucrats by political principals, and substantially less accountability of bureaucrats to those principals, than has previously been considered in this stream of thought. Two substantively quite different contributions (Carpenter 2001, Moe 2006) share an interesting theoretical similarity on this point. Carpenter's book is not explicitly steeped in the language or concepts of principal-agent theory (though see Gailmard and Patty 2007 for a formal principal-agent model that is similar in several ways), while Moe's paper is. But both scholars explore the implications for bureaucratic accountability of the fact that bureaucrats may assemble electoral coalitions to influence the identity and ideology of who their legislative principals are in the first place. That is, scholars must consider that interaction between legislative principals and bureaucratic agents is embedded in a larger political system in which legislators are themselves agents. If bureaucrats occupy a dedicated enough share of the electorate to whom legislators answer, they can induce titular legislative principals to treat favorably the policy concerns and favored agencies of those bureaucrats. If the dedicated bureaucrats-in-electorate willing to work for their cause are in fact an intense minority, then this arrangement can skew the ultimate accountability of policy choices to the will of voters in general. Carpenter (2001) contends and adduces historical evidence that bureaucrats can act as opinion leaders in constellations of interest groups allied around related concerns and thereby induce those groups to lobby Congress in support of the very bureaucrats that Congress is nominally directing. In this way, Carpenter argues, bureaucrats can carve out some autonomy not mere discretion from political principals to pursue policy initiatives of their own choosing on their own terms. Whether the ends pursued by such bureaucrats in particular cases are salutary from the standpoint of public welfare or not, the point remains that this inversion of the canonical legislative principal-bureaucratic agent relationship reflects a failure of political institutions to engender accountability of bureaucrats to the will of Congress.

Some degree of confusion has arisen from the subsidiary observation that bureaucrats in fact have multiple principals. For one thing, Congress is a "they," not an "it," and furthermore that bureaucrats have other principals besides the multiple principals in Congress itself. These points are valid and should be considered in principal-agent thought on the topic (and have been many times over; see e.g. Moe 1985; Bendor, Taylor, and Van Gaalen 1987; Calvert, McCubbins, and Weingast 1989; McCarty 2002; Whitford 2006; Gailmard 2009). However, the multiple principals (or "common agency," as it is known in the principal-agent literature; cf. Bernheim and Whinston 1986; Laffont and Tirole 1993; Laffont and Martimort 2001) point does not negate the principal-agent-based reaction to concerns over bureaucratic accountability. Instead, it simply means that one should interpret the system on the assumption that each principal is doing its best, with whatever tools it has, to induce a favorable reaction from bureaucratic agents, taking as given that each other principal is trying to do so as well. This may lead to a more complicated model than single-principal variants, but does not resuscitate interpretations based on bureaucratic dominance. It is of course possible to postulate a model in which the attempts at influence by multiple principals cancel each other out completely, leaving the bureaucrat in a position of complete autonomy with respect to its principals; accounts of bureaucratic drift due to stalemate among multiple principals (e.g. McCubbins, Noll, and Weingast 1987) have some of this flavor on a small scale. But this would appear to be a relatively specialized model rather than one encompassing some general thrust of the implications of common agency for bureaucratic accountability.

In summary, the principal-agent perspective has fundamentally reshaped how political scientists view bureaucratic arrangements (see Gailmard and Patty 2012 for a more detailed review). But this cannot be said to arise from incontrovertible evidence that principal-agent models in which political principals tightly control agents, and only those models, are consistent with observed events. Rather, it is in some part because an assumption that political principals pursue desirable outcomes from their agents, however they understand that, in the most effective way possible given the problems and constraints that they face. The flexibility of principal-agent theory has made it possible to reconcile this interpretation with a great variety of seemingly puzzling or pathological bureaucratic institutions (see e.g. Miller 1993; Prendergast 2003, 2007; Gailmard and Patty 2007; Ting 2009). But that flexibility is not so great that the core idea of political principals exercising direction over policy as best they can is fundamentally compromised.

4 Conclusion

Principal-agent theory has proven to be a flexible and useful approach for interpreting the effects of institutional arrangements on accountability of public decision makers and public policy. Continuing development of this approach is the subject of a rich and ongoing literature. Particularly active areas at present include intrinsic motivation of agents (Prendergast 2007, Besley 2007, Gailmard and Patty 2007); the power of legal precedent in multi-tiered court systems (Carrubba and Clark n.d.); and delegation in parliamentary systems (Huber and Shipan 2002; Dewan and Myatt 2010).

The "theory" is better understood as a family of models with a related perspective, than as a single encompassing theory with a specific set of assumptions and conclusions. It is hard to imagine how the theory as a whole, rather than a specific model, could be falsified. Principal-agent theorists in political science have been reasonably adept at postulating configurations of information asymmetries, incentive problems, and contracting limitations to rationalize, as the best that a principal can expect to attain under the circumstances, a wide variety of seemingly puzzling and pathological institutions and behaviors. By the same token, when used in this particular way, principal-agent theory does not help us to be more demanding of our public officials. Instead this approach to principal-agent analyses of institutions asks how it could be that these pathologies might nevertheless result from principals doing the best job they can at holding their agents accountable. In this way it reveals the fundamental limitations on accountability imposed by political institutions, apart from the capabilities of the principals that operate within them.

References

Achen, Christopher and Larry Bartels (2002). Blind Retrospection: Electoral Responses to Droughts, Flu, and Shark Attacks. Presented at the Annual Meeting of the American Political Science Association, Boston, MA.

Ashworth, Scott and Ethan Bueno de Mesquita (2009). Electoral Selection, Stratetic Challenger Entry, and the Incumbency Advantage. Journal of Politics 70: 1006-1025.

Austen-Smith, David and Jeffrey Banks (1998). Electoral accountability and incumbency. In Models of Strategic Choice in Politics (P. Ordeshook, Ed.). Ann Arbor, MI: University of Michigan Press.

Banks, Jeffrey and Rangarajan Sundaram (1993). Moral hazard and adverse selection in a model of repeated elections. In Political Economy:

Institutions, Information Competition, and Representation (W. Barnett, et al., Eds.). New York: Cambridge University Press.

Banks, Jeffrey and Rangarajan Sundaram (1998). Optimal Retention in Agency Problems. Journal of Economic Theory 82: 293-323.

Banks, Jeffrey and Barry Weingast (1992). The Political Control of Bureaucracies under Asymmetri Information. American Journal of Political Science 36: 509-524.

Barro, Robert (1973). The Control of Politicians: An Economic Model. Public Choice 14: 19-42.

Bendor, Jonathan and Adam Meirowitz (2004). Spatial Models of Delegation. American Political Science Review 98: 293-310.

Bendor, Jonathan, Serge Taylor, and Roland Van Gaalen (1987). Politicians, Bureaucrats, and Asymmetric Information. American Journal of Political Science 31: 796-828.

Bernheim, Douglas and Michael Whinston (1986). Common Agency. Econometrica 54: 923-942.

Besley, Timothy and Anne Case (2003). Political Institutions and Policy Choices: Evidence from the United States. Journal of Economic Literature 41: 7-73.

Besley, Timothy (2007). Principled Agents? The Political Economy of Good Government. New York: Oxford University Press.

Bolton, Patrick and Mathias Dewatripont (2004). Contract Theory. Cambridge, MA: MIT Press.

Bueno de Mesquita, Ethan and Matthew Stephenson (2007). Regulatory Quality with Impefect Oversight. American Political Science Review 101: 605-620.

Calvert, Randall, Mathew McCubbins, and Barry Weingast (1989). A Theory of Political Control and Agency Discretion. American Journal of Political Science 33: 588-611.

Canes-Wrone, Brandice, Michael Herron, and Kenneth Shotts (2001). Leadership and Pandering: A Theory of Executive Policymaking. American Journal of Political Science 45: 532-550.

Carpenter, Daniel (2001). The Forging of Bureaucratic Autonomy. Princeton, NJ: Princeton University Press.

Carrubba, Clifford and Tom Clark (n.d.). Rule Creation in a Political Hierarchy. American Political Science Review, forthcoming.

DeFigueiredo, Rui J.P., Pablo Spiller, and Santiago Urbiztando (1999). An Informational Perspective on Administrative Procedures. Journal of Law, Economics, and Organization 15: 283-305.

Dessein, Wouter (2002). Authority and Communication in Organizations. Review of Economic Studies 69: 811-838.

Dewan, Torun and David Myatt (2010). The Declining Talent Pool of Government. American Journal of Political Science 54: 267-286.

Dixit, Avinash, Gene Grossman, and Elhanan Helpman (1997). Common Agency and Coordination: General Theory and Application to Government Policy Making. Journal of Political Economy 105: 752-769.

Epstein, David and Sharyn O'Halloran (1994). Administrative Procedures, Information, and Agency Discretion: Slack vs. Flexibility. American Journal of Political Science 38: 697-722.

Epstein, David and Sharyn O'Halloran (1999). Delegating Powers. New York: Cambridge University Press.

Fearon, James (1999). Electoral Accountability and the Control of Politicians. In Adam Przeworski, Bernard Manin, and Susan Stokes (eds.), Democracy, Accountability, and Representation. New York: Cambridge University Press.

Ferejohn, John (1986). Incumbent Performance and Electoral Control. Public Choice 50: 5-25.

Fiorina, Morris (1979). Control of the Bureaucracy: A Mismatch of Incentives and Capabilities. In The Presidency and the Congress: A Shifting Balance of Powers? William Livingston, Lawrence Dodd and Richard Schott, eds. Austin, Texas: Lyndon B. Johnson School of Public Affairs, Lyndon Baines Johnson Library.

Fox, Justin and Kenneth Shotts (2009). Delegates or Trustees? A Theory of Political Accountability. Journal of Politics 71: 1225-1237.

Gailmard, Sean (2002). Expertise, Subversion, and Bureaucratic Discretion. Journal of Law, Economics, and Organization 18: 536-555.

Gailmard, Sean (2009). Multiple Principals and Oversight of Bureaucratic Policy Making. Journal of Theoretical Politics 21: 161-186.

Gailmard, Sean and John Patty (2007). Slackers and Zealots: Civil Service, Bureaucratic Discretion, and Policy Expertise. American Journal of Political Science 51: 873-889.

Gailmard, Sean and John Patty (2012). Formal Models of Bureaucracy. Annual Review of Political Science 15: 353-377.

Gailmard, Sean and John Patty (2013). Learning While Governing: Expertise and Accountability in the Executive Branch. Chicago: University of Chicago Press.

Healy, Andrew and Neil Malhotra (2010). Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence. Quarterly Journal of Political Science 5: 193-208.

Heclo, Hugh (1977). A Government of Strangers: Executive Politics in Washington. Washington, DC: Brookings Institution Press.

Holmström, Bengt (1984). On the Theory of Delegation. In Marcel Boyer and Richard Khilstrom, editors, Bayesian Models in Economic Theory. New York: North Holland.

Holmström, Bengt and Paul Milgrom (1991). Multitask Principal Agent Analyses. Journal of Law, Economics, and Organization 7: 24-52.

Horn, Murray (1995). The Political Economy of Public Administration. New York: Cambridge University Press.

Horn, Murray and Kenneth Shepsle (1989). Administrative Process and Organizational Form as Legislative Responses to Agency Costs. Virginia Law Review 75: 499508.

Huber, John and Charles Shipan (2002). Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy. New York: Cambridge University Press.

Kiewiet, D. Roderick and Mathew McCubbins (1991). The Logic of Delegation. Chicago: University of Chicago Press.

Laffont, Jean Jacques and Jean Tirole (1993). A Theory of Incentives in Procurement and Regulation. Cambridge, MA: MIT Press.

Laffont, Jean Jacques and David Martimort (2001). The Theory of Incentives: The Principal-Agent Model. Princeton, NJ: Princeton University Press.

Lenz, Gabriel (2012). Follow the Leader: How Voters Respond to Politicians' Performance and Policies. Chicago: University of Chicago Press.

Lowi, Theodore (1979). The End of Liberalism (2nd ed.). New York: Norton.

Lupia, Arthur and Mathew McCubbins (1998). The Democratic Dilemma: Can Citizens Learn What They Need to Know? New York: Cambridge University Press.

Majmundar, Sumon and Sharun Mukand. Policy Gambles. American Economic Review 94: 1207-1222.

McCarty, Nolan (2002). The Appointments Dilemma. American Journal of Political Science 48: 413-428.

McCubbins, Mathew, Roger Noll, and Barry Weingast (1987). Administrative Procedures as Instruments of Political Control. Journal of Law, Economics, and Organization 3: 243-277.

McCubbins, Mathew, Roger Noll, and Barry Weingast (1989). Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies. Virginia Law Review 75: 431489.

Miller, Gary (1993). Managerial Dilemmas: The Political Economy of Hierarchy. New York: Cambridge University Press.

Miller, Gary (2005). The Political Evolution of Principal-Agent Models. Annual Review of Political Science 8: 203-225.

Moe, Terry (1985). Control and Feedback in Economic Regulation: The Case of the NLRB. American Political Science Review 79: 1094-1116.

Moe, Terry (2006). Political Control and the Power of the Agent. Journal of Law, Economics, and Organization.

Niskanen, William (1971). Bureaucracy and Representative Government. Chicago: Aldine.

Persson, Torsten, Gerard Roland, and Guido Tabellini (1997). Separation of Powers and Political Accountability. Quarterly Journal of Economics 112.

Prendergast, Canice (2003). The Limits of Bureaucratic Efficiency. Journal of Political Economy 111: 929-958.

Prendergast, Canice (2007). The Motivation and Bias of Bureaucrats. American Economic Review 97: 180-196.

Romer, Thomas and Howard Rosenthal (1978). Political Resource Allocation, Controlled Agendas, and the Status Quo. Public Choice 33: 27-43.

Ting, Michael (2009). Whistleblowing. American Political Science Review 102: 249-267.

Volden, Craig (2002). A Formal Model of the Politics of Delegation in a Separation of Powers System. American Journal of Political Science 46: 111-133.

Weingast, Barry and Mark Moran (1983). Bureaucratic Discretion and Political Control. Journal of Political Economy 91: 765-800.

Whitford, Andrew (2006). The Pursuit of Political Control by Multiple Principals. Journal of Politics 67: 28-49.