Adapting Agencies: Competition, Imitation, and Punishment in the Design of Bureaucratic Performance

Andrew B. Whitford

The U.S. Department of Veterans Affairs (VA) serves millions of veterans across the world on a daily basis. It is the single largest governmental direct provider of medical care in the United States, serving 4.2 million patients per year; it also provides compensation to over 2.7 million veterans and their survivors yearly and oversees 2.1 million insurance policies. Not surprisingly, a primary concern is the rapid provision of benefits and settlement of claims. As its Departmental Performance Plan for fiscal year (FY) 2000 notes, the problem is to refine its goals, objectives, and strategies; measure its performance; and critically evaluate its programs.

For benefits provision, the VA uses a mixture of scorecards, measurement strategies, and information management technologies to meet quality and timeliness standards such as the rapidity of the treatment of cases. One component is the enhanced managerial review of over fifty regional offices through better cost accounting, regional comparisons, and benchmarking (1997 Strategic Plan for Fiscal Years 1998–2003).

Many have analyzed the 1993 Government Performance and Results Act (GPRA), some laudatory and some not (National Academy of Public Administration 1999). The VA's experiences show a concern greater than just performance measurement and analysis. The ongoing experiments focus on organizational change, comparison, and redirection. They recognize that agencies adapt, that they can be compared and compete with one another, and that they can learn from their and others' experiences.

This essay investigates the design of bureaucratic performance through comparison, competition, imitation, and redirection. Studies of bureaucratic politics, organizational change, and policy implementation regularly assume that agencies respond to external and internal circumstances and that organizations can be redirected. Agencies adapt when they choose actions or use information based on past experiences. They may adapt when they compete with other agencies over resources, constituencies, or policy areas, perhaps causing them to perform differently than they would absent competition. They may perform differently because an interested overseer compares agencies and uses those comparisons to reward or punish. Just as states can serve as policy laboratories, agencies may also perform differently because their counterparts offer opportunities for imitation when they compare how their own actions and performances and others' relate. In fact, comparison is often how the media, consultants, and academics learn about agencies' performance and productivity.

The central assumption in this essay is that agencies adapt to their surroundings in an imprecise way. This assumption is strongly related to an older but continuing theme in the study of organizations. Herbert Simon's bounded rationality—and the corresponding idea of incremental decision making—has served as a motivating force in numerous studies of decision makers, organizations, and even public decision processes (e.g., Bendor 1995; Cyert and March 1963; Lindblom 1959; Newell and Simon 1972; Simon 1947, 1957, 1976b).

In this context, I concentrate on three levers of redirection: comparison, punishment, and imitation. I begin by creating a simulation model of a baseline agency with measurable attributes of performance, goals, and technology. The agency can take only one action: it can search for technology, or a means of implementing policies, by means of procedures, information services, human knowledge, or even mechanical methods. In this model, the agency establishes goals or performance targets, observes its performance, adopts a technology, and sets new goals.

This model agency provides a baseline for the comparison of three specific structural features. In the first extension, I introduce the comparison and punishment of agencies, or simple competition. In the second, I allow for the differential punishment of agencies in a competitive system. Third, I allow agencies to imitate one another; "learning by others' doing" is often the purpose of experiments, pilots, and laboratories.

In the last extension, I replace these deterministic systems with a naive stochastic alternative: random punishment.

These extensions illuminate the difficult trade-offs encountered in designing agencies: predictability versus change, search versus outcomes, stasis versus growth. These represent competing "goods" in attempts to alter agency direction. By encouraging change, we upset prediction; by encouraging innovation, we do not guarantee outcomes. Additionally, the exercise showcases a flexible and powerful modeling strategy for the evaluation of specific design combinations discussed throughout political science and public administration.

This essay proceeds as follows. In the next section, I review basic literatures on adaptation and competition in bureaucracy. I then offer a simple model of adaptive agencies to establish a baseline against which to compare the specific structural features of comparison, imitation, and random punishment. Last, I conclude with thoughts on this modeling strategy, the results, and directions for future research.

Adaptation and Competition in Bureaucracies

This essay combines two major initiatives in the study of public agencies as organizations: the modeling of agencies as adaptive systems and the normative value of competition as a structure for pressing agency performance. In this section, I first review adaptation in public bureaucracies and technological search as basic ways of visualizing agency actions. Then I discuss issues of comparison, competition, and information revelation in bureaucracies.

The models I build here have attributes of complex adaptive systems, a descendant of cybernetics (Ashby 1957). This approach allows the creation of realistic models of organizational behavior in which the emphasis is on process, benchmarking, and flexibility. Studies of adaptive public bureaucracies date in Lindblom's work on incrementalism in policy (1959; see also Bendor 1995) and Simon's studies of individual decision making, bounded rationality, and incrementalism in public agencies (Newell and Simon 1972; Simon 1957, 1976b). Many have offered dynamic models of adaptive behavior in the behavioral tradition (Axelrod 1976, 1997; Bendor and Moe 1985; Cohen 1981, 1984; Cohen and Axelrod 1984; Cohen, March, and Olsen 1972; Crecine 1969; Cyert and March 1963; Kollman, Miller, and Page 1992; March and Simon 1958; Padgett 1980; Simon 1947). Adaptive models have grown in popularity with the massive growth in computing power, as in computational organizational theory (Burton and Obel 1995; Carley and Prietula 1994). In economic history, adaptation is a feature of and a solution to problems in large, modern organizations. With the coevolution of technology and business institutions, new forms are chosen because they solve organizational problems (Chandler 1990). Technology and efficiency intertwine as mechanisms of adaptation and selection. As Arthur (1989) and David (1985) show, the technology in place at a point in time is truly cumulative technology; today's technical advances build from and improve upon status quo technologies.

The baseline model here is an agency searching for technology in order to change its performance. Search is a way to change the relationship between performance and action, but it is fraught with error. Technology is a technical method; it is how the agency applies knowledge for practical purposes. Public agencies make constant investments in finding and implementing new technologies. As an example, the VA has made massive investments in integrating information technologies. One "soft" technological search in which agencies invest is selecting and training personnel.

Like all agencies, this adaptive agency resides in a larger social system that is concerned about social efficiency, the output of government, and political responsiveness. Political overseers shape public agencies and what they do. Over the past four decades, competition has been offered as a way to fuse the public goods provision with the efficiency of markets (Tullock 1965; Downs 1967; Niskanen 1971). Essentially, the competition movement is an exercise in the politics of bureaucratic structure, shaping the incentives of agencies to redirect their behavior to new ends.

Niskanen argues that competition allows the comparison of relative prices of competing bureaus and so shifts power from bureaucrats to politicians. As Boyne (1998) notes, these models' core hypotheses are that under competition total spending on services will fall and technical efficiency will increase; they make no specific prediction about allocative efficiency under competition. Direct descendents are "contracting out" and compulsory competitive tendering (Niskanen 1968; Boyne 1998; Kettl 1993).

Refinements of Niskanen argue that the model inaccurately characterizes the interaction between bureaucrats and legislators (e.g., Blais and Dion 1991). Migue and Belanger (1974) note that bureaucrats may maximize goals other than the supply of public services. Conybeare (1984) reveals an implicit assumption of perfect price discrimination (see also Bendor, Taylor, and Van Gaalen 1985; and Breton and Wintrobe 1975). Miller and Moe (1983) show that modeling the legislature alters predictions about the power of privatization and competition. Empirically, there are mixed findings at best on competition and agency costs and output (Boyne 1998; Conybeare 1984; Higgins, Shughart, and Tollison 1987). Conybeare (1984) notes that even if multiple, competitive bureaus are competing for funding rather than producing equivalent goods (as in McGuire, Coiner, and Spancake 1979) there may be negative side effects such as high monitoring costs.

What, then, is the value of competition? Miller and Moe (1983) show that competition for the public supply of a good is valuable when it reveals information about actual supply costs and thus places monitors in better decisional positions and enhances their power. The advantage of competition is how it reveals information by allowing comparison. Another reason for multiple and comparable agencies is the value of redundancy or parallelism so as to reduce system-level performance errors (Bendor 1985; Landau 1969; Heimann 1993).

Agencies will respond to comparison, competition, and information revelation because of the real world implications of failure. While bankruptcy is not possible in the public sector, there is agency deletion (e.g., the Interstate Commerce Commission), reductions in budgets (e.g., the Reagan era Environmental Protection Agency), media coverage (e.g., the Bureau of Alcohol, Tobacco, and Firearms), external options for agency heads (e.g., the Department of the Treasury), the simple incentive just to do well in the organization (Edwards, Nalbandian, and Wedel 1981), and other nonmonetary incentives (Crewson 1995).

Five Models of Adaptive Agencies

Basic Agency Adaptation

I start with Levinthal and March's (1981) model of adaptive organizational search. In this model, organizations change through their adaptive search for new technology. The model shows that an organization's behavior and performance reflect the consequences of simple adaptation in an ambiguous environment. The approach is a generalized search model from an organizational learning perspective. In this essay, this model serves as a baseline against which to examine the cases of competition, imitation, and random punishment among competitors (327).

The basic assumption is that organizations change their performance and goal attainment behavior by searching across possible implementation technologies. Agencies search by innovating (Radner 1975; Knight 1967; Nelson and Winter 1978). The technology an agency discovers is exogenous, but the search process depends on its past performance, goals, and search investments.

Agencies perform, have goals (that may be attained), and make expenditures to determine the best way to implement policies. To start, the agency sets a performance goal based on its past experience. The agency has only one way to attain that goal: to search for a new implementation technology. When it begins the search process, the agency examines its past performance and expenditures to assess the effectiveness of different types of technology search. Doing so helps determine the agency's propensity to search, its search efficiency, and the amount of resources it allocates to searching.

The key is how many times the agency will undertake a refinement or innovation search. Refinement involves only a local search; innovation is far ranging (but not global). The agency compares the search processes' results and its status quo technology. Once implemented, the agency experiences a new performance level based on its technology, expenditures, and environmental variation. The agency then adapts its goals based on this performance experience and starts the process again.

Specifically, performance in a given time period, P_t, is a function of the technology it implements (T_t), search costs, and an exogenous and varying environmental variable (a_t). The agency makes expenditures for two types of search: refinement (R_t) and (I_t).

$$P_t = (1 + a_t)\ T_t - R_t - I_t.$$

At the beginning of the period, the agency sets a performance target or goal (G_t) before experiencing performance. It modifies that target given performance as

$$G_t = b_1 P_{t-1} + (1 - \mathrm{b}_1) G_{t-1}.$$

The term P_{t-1} is organizational performance in time $t - 1$.[1] Performance, goals, and technology are simple indices (or complex combinations that comprise complex phenomena) (Jones 2001). Organizations may pursue multiple goals simultaneously; this can affect search behavior in an organization and may even enhance organizational performance (Cohen 1984).

Technology largely determines the agency's performance, and the search process is what the agency controls. The agency never forgets its past technology, which forms a baseline against which all new technologies are compared. Technological search proceeds by means of refinement and innovation. The agency chooses the best of three technologies: that obtained from refinement (T_r) search, that obtained from innovation (T_i) search, or its technology from the past period.[2]

$$T_t = \max\,(T_r,\ T_i,\ T_{t-1}).$$

The key decision is for the agency to decide which resources it will commit to the search for a new technology. The agency compares its performance and the search resources it committed in the past period. The agency determines whether it met its goal: if it is met, it deems "successful" the type of search in which it engaged (innovation or refinement). Meeting a goal is the message the agency receives about the appropriateness of its past behavior. It is interested only in whether it meets its goal, not in how much it is exceeded.

The agency also assesses the value of engaging in any search at all. This *general propensity* for search ($S_{s,t}$) depends on whether search resources were made available and the agency met its performance target or goal ($Q_{s,t-1}$):

$$S_{s,t} = Q_{s,t-1}b_2 + S_{s,t-1}(1 - b_2).$$

Here $Q_{s,t-1}$ is a test that takes the value 1 in two cases: if search resources were expended and the performance goal was met or if search resources were not expended and the performance goal was not met. The term b_2 determines whether this propensity has memory; it would limit the agency's ability to make sudden changes, failure would not easily translate into lower search investments, and success would not necessarily cause more searching.

Second, the agency's propensity to undergo a particular kind of search ($S_{i,t}$ or $S_{r,t}$) depends on a test relating the size of the type of expenditure and past performance: innovation ($Q_{r,t-1}$) or refinement ($Q_{i,t-1}$).

$$S_{r,t} = Q_{r,t-1}b_3 + S_{r,t-1}(1 - b_3),$$

$$S_{i,t} = Q_{i,t-1}b_4 + S_{i,t-1}(1 - b_4).$$

The terms $S_{r,t}$ and $S_{i,t}$ do not necessarily sum to one because an agency may not search at all; their sum may exceed one given the differential allocation of resources between the two types.

While this model allows testing the differential effects of learning rates, I assume that $b_2 = b_3 = b_4 = 1$ so that the search propensities are functions of the past period's performance only. These specific learning rates make the change for a search propensity incremental in goal attainment; this is a restrictive assumption, but it is consistent with past work on learning, organizations, and policy formation (Lindblom 1959).

These search propensities drive both the amount of sampling and the search space through which agencies learn about alternate technologies. The agency's search resources in a time period are limited, and the amounts allocated for refinement and innovation search can differ. First, the agency determines the total amount of resources it can dedicate to search ($U_{s,t}$) given $S_{s,t}$.

$$U_{s,t} = S_{s,t}(P_{t-1} + R_{t-1} + I_{t-1}).$$

The restrictive assumption here is that the agency is not limited in search, performance, or goal attainment by a lack of resources.

Second, this total is allocated to the competing types. For public agencies, it is likely that the resources available for refinement are greater in lean times and that innovation resources are greater when the organization attains goals. If the performance goal was achieved, the refinement resources are greater than those for innovation.

$$U_{i,t} = U_{s,t},$$

$$U_{r,t} = (U_{s,t})^{1/h_r}.$$

Here h_r, the refinement slack coefficient, controls the shift from one type of search to another given performance achievement. If the performance goal was not achieved,

$$U_{i,t} = (U_{s,t})^{1/h_i},$$

$$U_{r,t} = U_{s,t},$$

where h_i is the innovation slack investment coefficient.

Not all available resources must be expended in this model. The resources expended on a search type depend on the propensity to undertake a kind of search and the amount of search resources available for that type. On one hand, agencies with limited resources may expend them all. On the other, agencies meeting performance goals may not expend any available resources when they reach a balanced state of performance and goal achievement.

The search propensities and resources available determine the innovation (I_t) and refinement (R_t) search resources expended.

$$I_t = S_{i,t}U_{i,t},$$

$$R_t = S_{r,t}U_{r,t}.$$

These search resources help determine the agency's sampling efforts for refinement ($K_{r,t}$) and innovation ($K_{i,t}$).

$$K_{r,t} = k_r R_t E_{r,t},$$

$$K_{i,t} = k_i I_t E_{i,t}.$$

The terms $E_{r,t}$ and $E_{i,t}$ are search efficiencies that reflect efficiency increases with increasing search but at a decreasing rate (Levinthal and March 1981, 313).

$$E_{r,t} = (E_{r,t-1} + K_{r,t-1})^{1/w_r},$$

$$E_{i,t} = (E_{i,t-1} + K_{i,t-1})^{1/w_i}.$$

Altogether, the agency has four interlocking constraints on its ability to generate changes in the technology it implements. It is limited by the existence of resources to commit to the search. Its propensity to search, regardless of resources, is limited by its past performance and expenditure experience. The agency's search efficiency is determined by its past search experience. Last, it is constrained by the weight applied to control the shift from one search type to the other. Most importantly, these are constraints on how the agency searches for technology—not on how it discovers it.

The technology an agency implements depends on the number of technologies it samples, the underlying distribution of possible (or undiscovered) technology, and its current technology. The returns to the search agencies undertake depend on their expenditures, the efficiency of the search, and the current opportunities for development. The likelihood of finding an appropriate change depends on the number of opportunities sampled.

Given $K_{r,t}$ and $K_{i,t}$, the search is drawn from distributions based on the technology implemented in the past time period. For refinement, in each time period, let $n = K_{r,t}$. A drawn procedure is a change from T_{t-1}, where

$$T_{r,n} \sim N(0, V_{r,t,n}),$$
$$V_{r,t,n} = c_1 T_{t-1} \quad \text{for } n = 1,$$
$$V_{r,t,n} = c_2 V_{r,t,n-1} \quad \text{for } n > 1.$$

In a time period, if search resources are great enough, the agency may make a number of draws from this distribution. As it does, the variance of the distribution falls over multiple draws. The term c_1 means that the variance in refinement depends on the past technology; c_2 means that that variance falls in the time period when an agency makes multiple refinement searches. The space of possible refinements becomes smaller over multiple searches.

For innovation, a drawn procedure is a change from T_{t-1}, where

$$T_i \sim LN(0, V_i),$$
$$V_i = c_3 T_{t-1}.$$

The term c_3 plays the same role as c_1.

It is possible to learn *not* to search due to estimation errors.

$$EE_{r,t} \sim N(0, V_{e,r,t}),$$
$$EE_{i,t} \sim N(0, V_{e,i,t}).$$

The term $V_{e,i,t}$ is initially greater than $V_{e,r,t}$, but each declines with the implementation of a given type of change in procedure.

The value of either a refinement or an innovation is known with certainty once it is implemented. It is known only with error when it is not implemented. Once drawn, T_r, T_i, and T_{t-1} are compared in each time period and the largest value is selected and implemented. The term $V_{e,i,t}$ shrinks if innovation is the final technology; $V_{e,r,t}$ shrinks if refinement is successful.

In this model, the agency chooses a technology, experiences performance, compares that level with its goal, and starts the search process

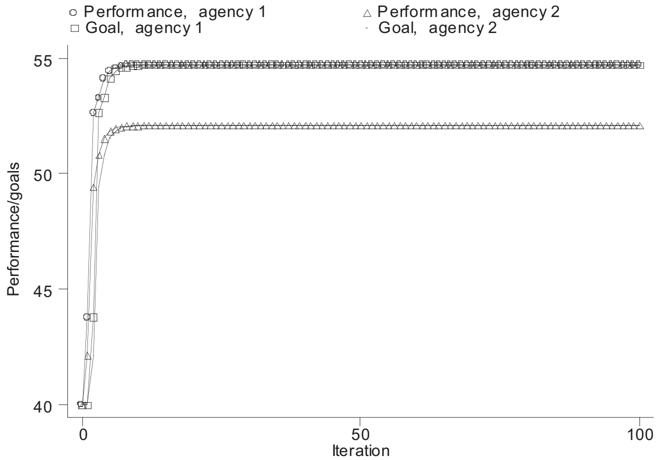

Fig. 1. Baseline: Performance and goals

again. I simulate this baseline model of simple adaptation for two agencies given the initial conditions and parameters in the appendix at the end of this essay.[3]

Figure 1 shows the performance and goal paths for two agencies where each agency evolves over one hundred time periods. Both agencies quickly settle down to equilibrium levels of performance. Under these initial conditions, agency performance and goals are in a steady state. By giving the agency a "shock" in the first time period (forcing it to make three technological searches), changes in technology spur changes in performance and goals. Both agencies explore new technologies and alter their performance and goal attainment; for each, the result is a higher performance level.

Figure 2 shows the time paths of one agency's propensities to undertake search for technology (general, innovation, and refinement); only one agency's results are shown because of their similar experiences. The general and innovation propensities track one another (due to the averaging process) and settle down over time. As this clearly shows, the equilibrium performance level occurs because the agency becomes unlikely to search after a period of time. As expected, refinement dominates innovation because this type of search is most closely related to an adaptive theory of organizational behavior. In fact, the adaptive model specified here should produce this result.

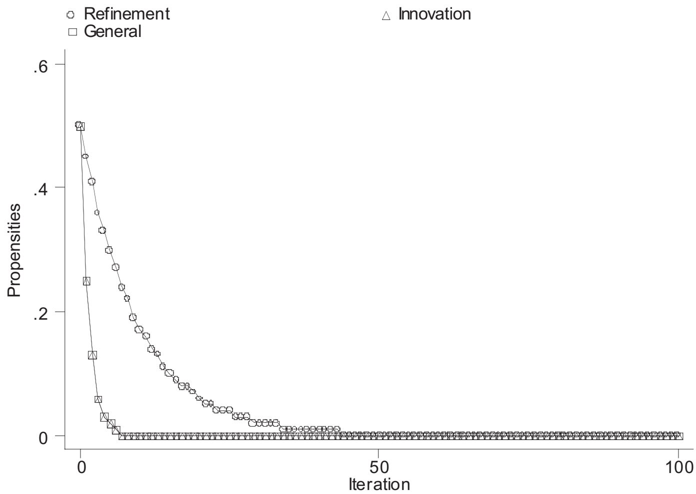

Fig. 2. Baseline: Propensities

Figure 3 shows the outcome of a search, or the likelihood of choosing an innovation, choosing a refinement, or retaining the status quo. On this scale, the agency is choosing innovations when the score is close to two, refinements when the score is close to one, and the status quo when the score is close to zero. The agency chooses innovation in the first period and then quickly rejects the search outcomes in favor of the status quo. This also shows that the likelihood of success (of performance exceeding or matching goals) goes to one as the likelihood of retaining the status quo goes to zero. Figure 4 shows how the number of searches approaches zero rapidly.

This baseline model demonstrates three basic propositions about agency learning and change. First, technology search, as defined here, produces changes in performance and goals, subject to constraints on search. It is the central mechanism by which the organization changes its operating behavior. Second, this system reaches a steady state of performance, goal attainment, and acceptance of the status quo technology. Without disturbances, the agency is unlikely to alter its technology, its goals, or its search for new technologies. Third, multiple indicators of agency activity are useful for examining how agencies use the search mechanism to change their performances. Here indicators go beyond simple performance attainment to include search propensities, success, and outcomes.

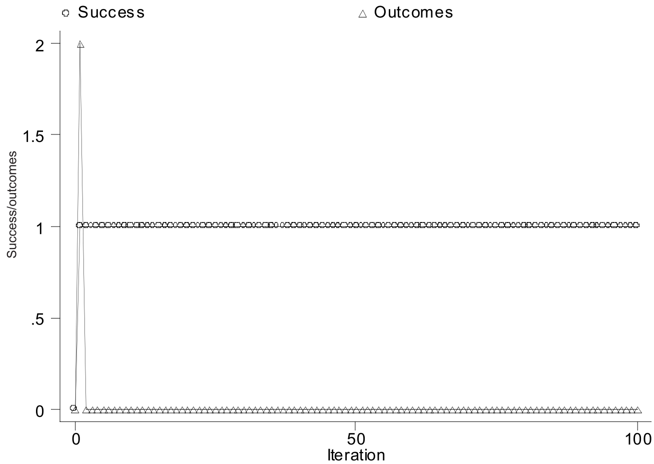

Fig. 3. Baseline: Success and outcomes

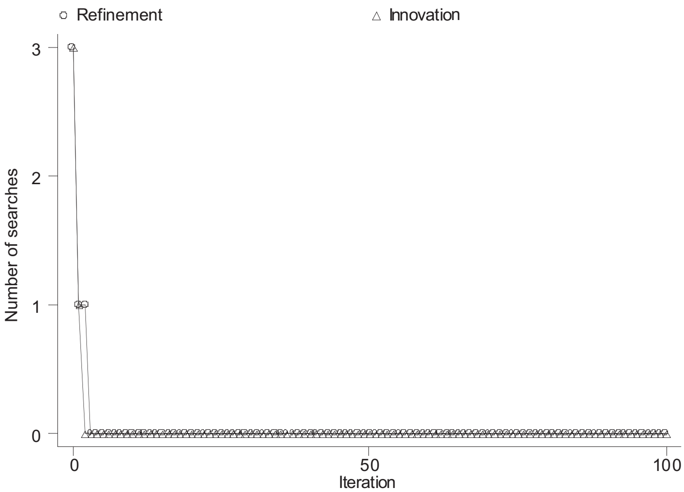

Fig. 4. Baseline: Number of searches

Competition among Adaptive Agencies

In the second model, two adaptive agencies compete on the basis of performance. The losing agency is punished, and agencies adapt their goals given their realized performances. A fundamental reason for competition is that comparison reveals information about what is possible for agency performance. The baseline model provides a point of comparison with

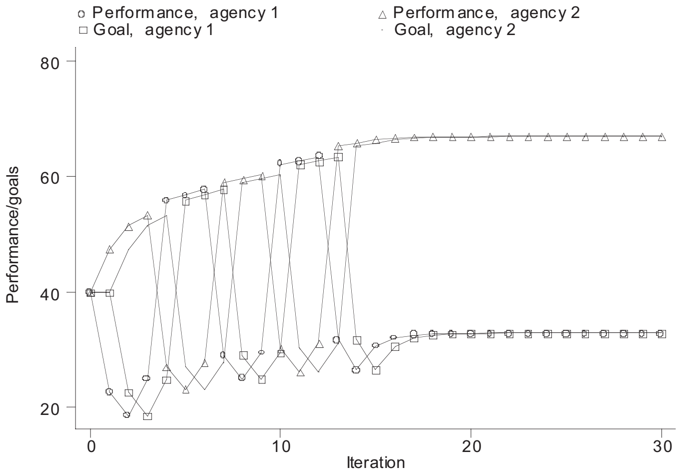

Fig. 5. Competition: Performance and goals

this second, institutional layer. This is competition over neither finite resources nor a pool of customers through price competition or rivalry. Perhaps the best metaphor is athletes competing in time trials.

Specifically, the agencies have independent search processes but their performances are compared at the end of each period. The lower-performing agency is punished and given only partial credit. Its observed performance is recalculated, and the losing agency's new goal now depends on this partial credit. In this application, two agencies are compared to one another. "Thin" competition allows me to isolate the effects of competition and may be realistic for public agencies.[4]

In the first model, the losing agency's performance measure is halved and punishment is symmetric (agencies are not distinguished). Figure 5 shows the time path for original initial conditions, where the model now includes a competition step. The paths in this figure are substantially different from those in Figure 1. Both display greater variation in their performances over the iterative process; the goal path of each agency tracks its performance path. The strong oscillation in the agencies' performance and goal paths is linked. Essentially, the "meeting-separating" pattern is a core result of this competitive process, as over time the aggressive movement of one agency to increase its performance leads to lagged increases in the other's performance. Eventually, both agencies settle into separate performance equilibria.

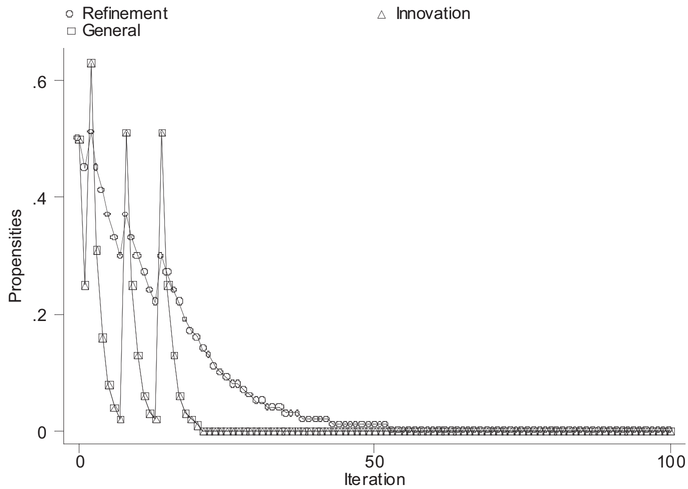

Fig. 6. Competition: Propensities

As noted earlier, the propensities help gauge how an agency's performance and goal find equilibrium. Figure 6 shows the time path for one agency for the three propensities. In contrast to figure 2, the search propensities do not settle down quickly. Moreover, innovation spikes follow punishment phases and refinement spikes follow innovation spikes: punishment induces innovation and local search follows innovation. Figure 7 visually confirms punishment's role in driving the agency's internal search process. The amount of search switches for the first periods of the simulation, but the number of searches falls over time.

When an agency is docked for underperformance, its key indicator is whether its credited performance does not meet its internal goal. This disparity generates any internal changes that the agency makes in response to competition and punishment. Specifically, these signals knock the agency off the performance-goal achievement steady state it naturally obtains in the underlying adaptive model. In this model, the central effect of competition and comparison is that agencies can find new levels of performance because performance inequities motivate their enhanced search for new technologies.

Competition with Differential Punishment

An alternate case is one in which competing agencies are punished differentially. In this case, the simulation involves a "90/10" punishment

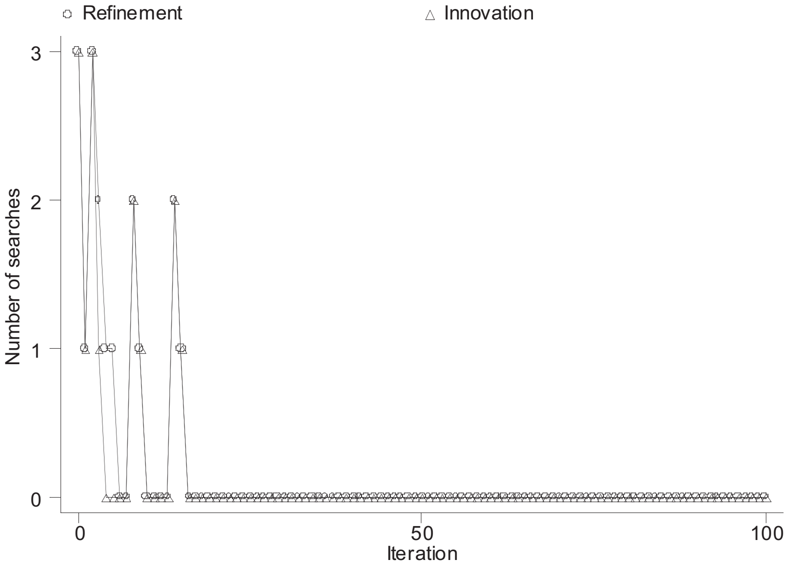

Fig. 7. Competition: Number of searches

scenario.[5] In this exercise, one underperforming agency receives 90 percent credit for its performance; the second receives 10 percent credit if it has a lower performance. The logic of this experiment is to uncover the potential differential impact of punishment on agency performance. Are lower punishments reflected in smaller agency increases in search, expenditures, and performance achievement? Figure 8 shows the paths for the performance and goals of the two agencies in one simulation.

This simulation highlights the differential responsiveness of an agency given the two punishment scenarios. The shallow valleys shown for agency 2 reflect the 90 percent credit rule; the deep valleys for agency 1 reflect the 10 percent rule. Note that agency 1—even given the strict 10 percent rule—finds a way to recover to globally higher performance levels. At a point in the model's evolution, agencies 1 and 2 diverge to different equilibrium performance levels.

What should we make of agency 1's higher equilibrium performance level given the strict 10 percent rule? Figure 9 shows pairs of equilibrium performance values for 100 simulations. For two agencies, 60 percent of the "10 percent" agency's equilibrium values are higher than that for the "90 percent" agency. However, the situation separates: for the other 40 percent, the 10 percent agency's values are substantially lower and the 90

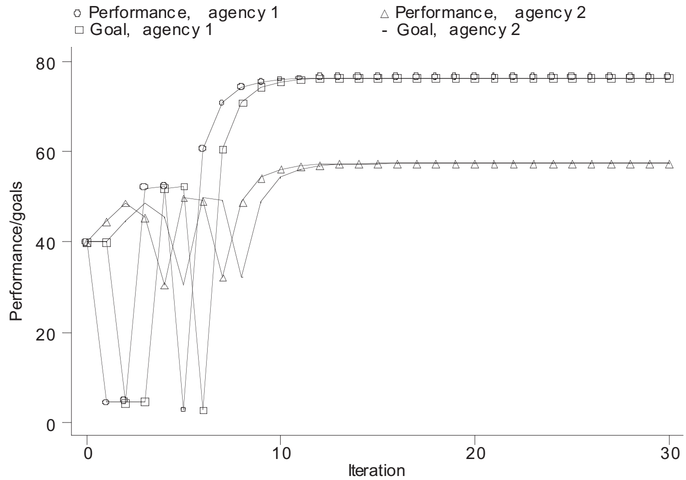

Fig. 8. Asymmetric competition: Performance and goals

percent agency's mean value is significantly higher. This difference is visually confirmed in figure 10 and is also confirmed statistically (unpaired t-test with an assumption of unequal variances; $t = 4.7902$; Satterthwaite's degrees of freedom, $df_t = 108.345$; difference significant at better than $p = 0.0001$).

The reason for this separation is in how the agencies search for technology. Punishment means a lower credited performance, making necessary the enhanced search for innovation, but one that is sustainable only for a short period of time. If the payoff to search is slow in coming, search stops and an agency falls into a position of lethargy, with low performance and no innovation.

Essentially, a valuable indicator is how punishments are implemented. The core test within the agency for determining propensities remains whether performance fails to meet expectations (given punishments). Within this model, systematic significant punishments can spur greater global performance than light punishments (the punishment magnitude affects only $U_{s,t}$).

An agency switches to a refinement search when its performance exceeds its goals. If it is in refinement mode and a significant punishment occurs, the agency's propensity for refinement falls. While the punishment may reduce $U_{s,t}$, it encourages innovation by discouraging refine-

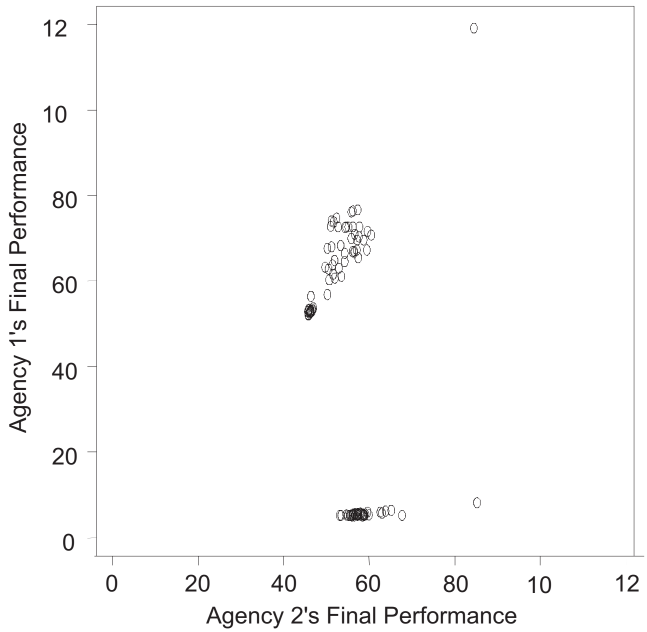

Fig. 9. Asymmetric competition: Performance of the two agencies

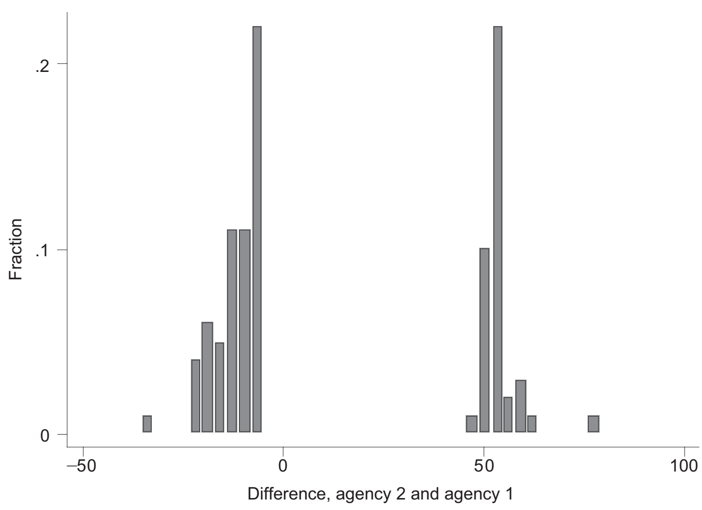

Fig. 10. Asymmetric competition: Relative performance

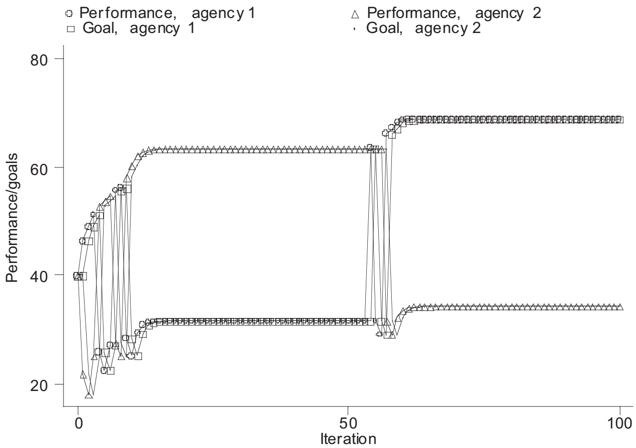

Fig. 11. Competition and imitation: Performance and goals

ment. Performance increases come from small levels of available funds because of the disjunction between credited performance and internal goals. The central result in this exercise is a counterintuitive one: on average, agencies that receive relatively heavy punishment (less credit for their performances) produce higher performance levels in the long run.

Imitation among Adaptive Agencies

Altering an agency's credited performance is just one way to alter its goals and search behavior. Instead, competing, adaptive agencies may imitate one another. Specifically, a losing agency may imperfectly imitate a dominant agency's technology. After it searches for a technology, experiences a level of performance, and competes with a counterpart, the losing agency imperfectly imitates a dominant agency's technology. Its technology is now a simple average of the two technologies in the system at that point in time.

Figure 11 shows the time path for the same initial conditions as in the adaptive system, except that the agencies now compete (with symmetric punishments set at 50 percent credit) and losing agencies imitate. The simulation initially proceeds similar to that under simple competition: initial oscillation is followed by a period of stability.

However, imitation produces a third dynamic. Here, a long period of stability is interrupted, as one agency's performance jumps once its tech-

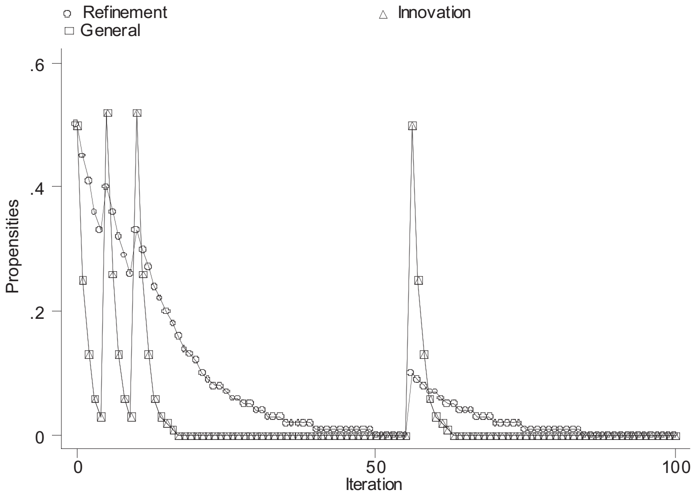

Fig. 12. Competition and imitation: Propensities for agency 1

nology evolves to a point where real performance changes occur. At that point, the formerly dominant agency is punished and responds with a short period of innovation. Notably, in this case the agencies actually switch positions and the dominant agency is replaced with the imitating one. As figure 12 shows, the dominant agency's search propensities signal the subordinate agency's intrusion in the last innovation spike. Agency 2 shows a similar propensity to search, generating increased search and causing a reversal of position.

In this model, imitation plays a peculiar role: it provides a basis for displacing dominant agencies with agencies that have been punished through competition. Competition creates the opportunity for the first agency to dominate. Yet imitation actually produces a globally higher level of performance *in the system* when it reaches equilibrium.

For agencies, the benefits of competition may be short-term benefits for the first agency. Under competition, imitation may produce a second set of benefits when the second agency matches and then replaces the first. Even so, imitation may have a downside when the dominating agency is beaten and then imitates the second agency's past, underperforming technology. In this model, this agency's response is to adopt the underperforming technology, resulting in a lower equilibrium performance path.

Naive Random Punishment

The last extension replaces competition and comparison with a system of random punishment. Random punishment alters the performance with which the agency is credited; this credited performance feeds back into the agency's adaptive process of goal setting and technology search. Neither competition nor imitation is allowed; the only shock to the system is random punishment.

Specifically, I draw a uniform random number and impose random punishment on each agency one-third of the time. If an agency is punished, it receives a 10 percent reduction in its performance credit. Two-thirds of the time, an agency will not be punished. In this exercise, agencies are never punished simultaneously. Figure 13 shows the performance and goals for two agencies in one simulation.

In this essay, I do not claim that these level increases are general. However, across many simulations the step-level increases in agency performance appear to be general. Again, the principal mechanism is search: once either agency has arrived at an equilibrium performance level, the agency no longer searches for new implementation technology. As figure 14 shows, the propensities to search for an agency are markedly different in this system. Search ebbs and wanes in direct relationship to the incidence of random punishment.

What is the relevant indicator of agency performance: its raw performance level or its search propensity? In each of these models, the baseline agency lacking external shocks fails to search for new technologies in order to enhance its performance once it is in equilibrium. Three basic means of external shocks are proposed: punishments based on simple competition; the imitation of known, winning technologies; and random punishments.

Each produces more complex behavior by this "model" agency; each generates internally consistent agency search for new technology. Naive random punishment produces the greatest amount of search and the most predictable pattern of increasing performance but at the expense of predictability for the agency. And under even random punishment there is significant nonlinear variation in the performances of agencies over time.

Discussion

This essay first provides a baseline model for understanding how agencies search for techniques and tools for bettering their performances. It then nests this baseline model in four institutional frameworks for redirecting agency behavior. The implications of these models are offered as

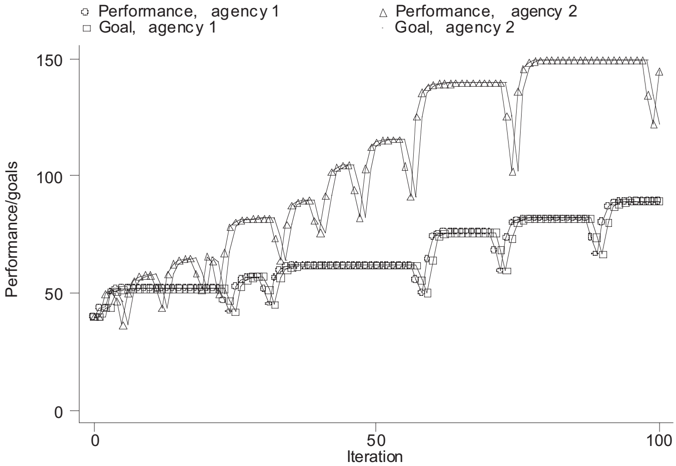

Fig. 13. Random punishment: Performance and goals

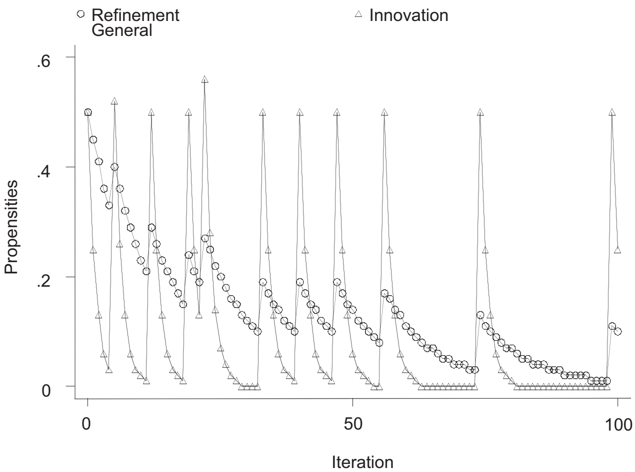

Fig. 14. Random punishment: Propensities for agency 2

exhibitions of how institutional frames alter internal agency processes in nonlinear, adaptive ways. They provide a fertile ground for extending past research on the nature of competition, comparison, imitation, and raw punishment in bureaucratic politics. In each case, the baseline agency is fully consistent with past work on adaptation and incrementalism in bureaucratic politics, public administration, and organizational theory.

Adaptation, competition, and imitation each can lead to radically different agency behaviors over time. Competition and comparison may better the performance of agencies and address concerns about how agencies produce outputs. In fact, the baseline agency may perform better under competition; in this model, there clearly is enhanced search for more productive technologies and procedures.

This achievement comes at two costs, though. First, greater performance may be the enemy of stability. Agencies perform better but with less stability in specific measurements. Specifically, agencies achieving a higher equilibrium level of performance (or long-term stability) may suffer oscillation getting there (or short-term instability). Second, agencies may separate into groups of winners and losers. In this case, a punishment strategy may ensure that one agency is a systematic underperformer.

In contrast, imitation produces a counterintuitive result: agencies that dominate other agencies produce less after being imitated, and agencies that imitate are forced to search. It appears that the amount of diversity in the system (the extent of search for new technologies) is enhanced by imitation but the system's stability is limited. Essentially, agencies rely on imitation, and search at points and thus are more likely to produce high performances. However, relative agency performance is less predictable.

The introduction of random punishment represents a quandary. The model suggests that unpredictable punishment leads to greater search opportunities—and increased performance—than under imitation with competition or even competition alone. It is important to point out that this exercise does not address substantial concerns like the costs of random intervention or the gaming of performance indicators. Yet the fact that naive punishment has this effect on a model agency signals the importance of knowing which system aspect a designer is attempting to control. If the intent is to create diversity of search and technological change, perhaps with the assumption that raw performance will increase, random punishment may be a first-best approach. If the point is predictability and stability, punishment strategies may be less than desirable. Some might even wonder if this is not a more realistic model of political intervention in agencies.

While this model's implications may not hold for thick competition (and current work in computational economics indicates that thick interactions are important for these social settings), thick interaction may not be empirically relevant for many agencies. Indeed, local interaction may be much more significant.

What matters for bureaucratic performance in this model? Clearly, goal setting, search and innovation, and the linkage of performance and search are key to increasing performance levels. Layering such an organization in a competitive process (even thin competition) suggests that while some will do better losers may do worse. Allowing collaboration—or at least imitation—may enhance any agency's learning, but it may also reduce the overall predictability of outcomes at both the system and agency levels. The trade-off is clear: innovation and predictability compete when punishment is present.

Appendix

TABLE A1. Variable Definitions and Initial Conditions

Variable Name	Initial Condition
Environmental variable (a)	0.05
Weight given to performance in changing goals (b_1)	1
Control of rate of change of refinement search propensity (b_2)	1
Control of rate of change of innovation search propensity (b_3)	1
Control of rate of change of general investment propensity (b_4)	1
Relation of variance in refinement to new technology (c_1)	0.01
Reduction of variance in refinement draws with experience (c_2)	0.98
Relation of variance in innovation to new technology (c_3)	0.01
Efficiency of refinement ($E_{r,t}$)	0
Efficiency of innovation ($E_{i,t}$)	0
Goal in current time period (G)	40
Refinement slack investment exponent (h_r)	1.1
Innovation nonslack investment exponent (h_i)	1.1
Innovation investment in current time period (I)	5
Number of refinement draws this period ($K_{r,t}$)	3
Number of innovation draws this period ($K_{i,t}$)	3
Constant for converting to refinement draws (k_r)	0.2
Constant for converting to innovation draws (k_i)	0.2
Performance in the current time period (P)	40
Refinement investment in the current time period (R)	5
Refinement search propensity in the current time period ($S_{r,t}$)	0.5
Innovation search propensity in the current time period ($S_{i,t}$)	0.5
Propensity to invest in this period ($S_{s,t}$)	0.5
Standard deviation of refinement estimation error ($V_{e,r,t}$)	0.5
Standard deviation of innovation estimation error ($V_{e,i,t}$)	0.75
Standard deviation of distribution of refinement opportunities ($V_{r,t}$)	$c_1 T_t$
Standard deviation of distribution of innovation opportunities ($V_{i,t}$)	$c_3 T_t$
Refinement efficiency exponent (w_r)	3.5
Innovation efficiency exponent (w_i)	3.5

Notes

This essay greatly benefited from the comments of George Krause, Ken Meier, Gary Miller, John Sprague, and Dan Wood. An early version of the essay was presented at the 1999 National Public Management Research Conference in College Station, Texas. All errors that remain are my own.

1. The term $b_1 > 1$ makes the target an exponentially weighted moving average of past performance; I assume $b_1 = 1$.

2. A general model would allow technology to change over time (to decay or improve). I assume constant technology in order to isolate the effects of organizational search on agency performance.

3. One way to present simulations is to generate a population of representative agencies, allow the population to evolve, and summarize the results for the population by either summary statistics or plotting the trajectory for the average member (for an example, see Cohen, Roilo, and Axelrod 1999). The figures for the simple adaptive agencies are representative of those obtained from multiple simulations. The figures for the extensions (except where indicated) are also representative. These simulations were written in Gauss.

4. Early work on adaptive systems shows that behavior becomes unpredictable when the number of interacting agents is large or as an agent's environment becomes more complex (Masuch and LaPotin 1989).

5. The use of light punishment has support in Ostrom's work on common pool resources (1990).