The Quest for Results: Opportunities for Using Performance Measurement in Chinese Public Administration[1]

Stuart Kasdin, George Washington University

Abstract: This paper examines the current situation of public management in China and the potential that management reforms might bring improvement. The primary goal of the paper is to examine the opportunities that incentivized performance measures have to enhance that agency management. The paper analyzes the conditions for how performance information can be fashioned into a metric, which is contractible. It then looks at the types of incentives that can be tied to the metric. It also considers the flexibility of the government agency, the central budget office, and the oversight entities, and the roles that each plays in ensuring successful implementation of a performance management system.

Good government doesn't simply happen. To encourage good governance, one hopes that the incentives built into the structure of the government, political system, and possibly of the community civil society network will be sufficient. This is not easy. In a 1980 speech Deng Xiaoping chronicled his concerns with Chinese bureaucracy, citing an over-concentration of power, paternalism, life tenure, and the existence of all forms of special privileges. The result, he indicated, was an abuse of power, inefficiency, overstaffing, organizational redundancies, corruption, and irresponsibility (Hamrin and Zhao, 1995; Zheng, 1997).

One of the ways that Chinese officials have worked to encourage greater efficiencies in state enterprises (mostly in the competitive sector) has been in using performance contracts (Shirley and Xu, 1997). Performance contracts function as a formal agreement between the manager of a state enterprise and the government, in which the manager commits to achieve specific targets in a certain time frame, while the government promises to award that accomplishment in some fashion (Shirley, 1998). However, Shirley and Xu (1998) showed that on average the performance contracts did not improve productivity in China's state enterprises and may even have reduced it. However, for the small portion (2.2%) of the total analyzed contracts that contained good elements–features that helped reduce information asymmetries, established effective incentives, and convinced managers that government promises in contracts were credible–productivity growth rates increased by 10 percent.

Ensuring efficient and effective organizational performance is difficult in all circumstances. However, government agencies are typically confronted by an even more complicated set of circumstances. The environment of government produces different management incentives from those of industry because the underlying conditions–clear incentives, transparent information, and unified principals–are rarely as simple. And establishing incentives is particularly difficult in government because their products are generally not easily measured, like profit or sales are for the private sector, and there are rarely external benchmarks or competition for comparing agency performance. As a result, it can be difficult to both evaluate program and managerial performance, as well as to motivate it with financial or administrative incentives.

For government to be effective, it needs to ensure that adequate incentives are present for program managers. Direct oversight may partially compensate in the short run for inefficient budget incentives, but ultimately, a lack of structural or institutional incentives will generate inefficient fund allocations and ineffective program administration. Thus, the design of a budget, like the design of any contract, will determine how the contract is fulfilled.

This paper describes how to develop an effective system, by which we can assess government performance, and thereby be in a position to offer rewards and penalties based on that performance. Using incentivized performance measures, program managers may be motivated to find the most effective ways to accomplish program goals, as well as increasing their operating efficiency.

The paper starts with a general discussion of some of the issues and concerns affecting Chinese bureaucracy. The paper then describes criteria for successful design of performance measures as metrics. To make use of performance measures for incentives to improve agency management, there must be consequences for good or bad program performance. Therefore the paper reviews the array of different budget and management tools that a central or regional planning office might use. Finally, the paper describes the various oversight entities that can review program or agency performance to ensure that the performance measures are used, and do not just become administrative make-work for the agencies.

Chinese Bureaucratic Performance and Public Management

Even as the Chinese private sector has grown enormously and displayed great advances in efficiency, questions about the public sector's performance remain. Unlike private sector entities, bureaucracies do not survive entirely on the basis of their profitability. Moreover, both the mission of a bureaucracy and its environment are quite different than that of the private sector. As such, the incentives for efficiency are not as strong in bureaucracies as in the private sector.

Chinese bureaucracies have faced a peculiar and challenging environment in recent years. Over the past decades, there has been an expanded decentralization of the bureaucracy, coupled with increased authority for agencies to raise their funds locally. Decentralization has meant fewer revenues from the center for provinces that had grown more affluent. This has, of course led both to the search by localities for increased sources of revenue, as well as decreased control by the center on how these funds were spent (Ma, 1996). Decentralization also meant decreased supervision of bureaucracy, both by establishing more *first-in-command* management with limited accountability, as well as making oversight entities more distant and less effective. There have been great consequences for how the bureaucracy behaves.

The process of decentralization and restricted appropriations made it imperative for agencies to seek out additional funds to maintain (and expand) operating capacity and to guard against swings in revenue. Also, another result has been that managerial autonomy grows as fund sources become less transparent to outside observers and more subject only to the agency manager's own decisions. Finally, in building businesses and adding resources, there are the bureaucratic rewards associated with managing a larger, more prominent enterprise. These incentives generally encourage bureaucratic expansionism, decreased outside scrutiny, and lower efficiency.

Thus, over the decades, Chinese government agencies have developed a general tendency toward seeking self-sufficiency, regardless of the size of the organization. Each government organization has sought to become *de facto* managerially independent of other government agencies by seeking control of as many resources as possible (Zheng, 1997). For instance, many government agencies now have their own dining halls, kindergartens, clinics, shops, transportation systems, construction companies, etc.

These problems are also seen in many of the state holding companies, for example. These companies suffer from ambiguous oversight and multiple lines of supervision, with limited financial and programmatic monitoring (Lo & Lee, 2001). In addition, the incentives of the top management are to grow larger and become more diversified – thus revenue enhancement and size matter most, along with ensuring good political relations with one's government superiors (Lo & Lee, 2001).

The loose controls on government funding are especially prominent at lower levels of government, in which the informal agency strategy of "raising some, borrowing some, moving around some, mortgaging some, accepting some" seems to dominate (Lü, 2000, p. 218). Bureaucratic agencies frequently will set up businesses as economic subsidiaries for the purpose of generating profits. The growth in business opportunities encouraged some bureaucracies to neglect their duties and focus instead on personal gain (Ma, 1996). The bureaucratic agencies' operation lies outside the formal state budgetary process and personnel management system, blurring the mission, goals, and expected function of the government agency. These government-sponsored businesses are also nested within a network of reciprocal behavior and potential insider favoritism.

One consequence is that some government agencies may participate in rent-seeking arrangements; that is the lobbying by state-owned, bureaucratic businesses, or private enterprises for special treatment. In such a case, government resources are diverted or regulatory treatment is altered to favor particular entities for non-programmatic reasons. Administrative rent-seeking has grown considerably in recent years, with more than 10 billion *yuan* worth of rents collected by officials and agencies endowed with regulatory power or monopoly, such as in the banking and finance sectors (Lü, 2000). Factors that are said to contribute to rent-seeking include insufficient budgetary appropriations for the government units, plus a continued expansionary pattern for the bureaucracy, coupled with the lack of a hard budgetary constraint for agencies (Lü, 2000).

Moreover, on the state and lower levels, agencies may seek additional funds and less bureaucratic dependence through the so-called *san luan*; that is, illicitly imposing fines or fees and apportioning forced donations. Without tight supervisions, agencies seeking out additional revenues outside the budget process can produce excessive and irregular charges on the local population (Lü, 2000). In 1995, it was estimated that extra-budgetary funds amounted to 38 billion *yuan*. Although the original intent of extra-budgetary funds may have been to encourage local governments to have an incentive to participate in economic development, some unintended consequences have emerged as the funds took on a life of their own. Certainly, although not all the extra-budgetary charges are illegitimate, nonetheless, they all tend to reduce the ability of an external agent to control funds expenditures and assess performance.

With respect to supervision, the *nomenklatura* system may not offer an ideal mechanism for managerial oversight. The system makes party committee "bosses" of all party and government officials, one level down. To the extent that the party officials have more knowledge in ideological issues than technical concerns, they are more likely to emphasize their interests. In addition, when technical information on program performance and management are poorly accessible and difficult to interpret, the incentive to focus more on non-performance concerns is heightened. In addition, the cadres have incentives to expand service organizations in order to augment the numbers of leading positions. When an agency creates a self-financing service organization, it is adding positions to the cadre management system (Lo & Lee, 2001). Promotion in a service organization is the same as promotion in an administrative agency.

As additional countervailing forces, traditionally, the Party stressed ideology and mass campaigns as incentives for the cadre and bureaucracy. Ideological indoctrination was viewed as the most important device in controlling the bureaucracy (Ma, 1996). A competing traditional influence is that of *guanxi*. As a potential response to uncertainty in Chinese politics, *guanxi* does not encourage system-wide efficiency. Given its reliance on social networks and gift exchange, the *guanxi* system tends to shift program performance away from its most efficient outputs to allow for the betterment of individual managers and their social connections.

These elements collectively help to create an environment in which Chinese bureaucratic organizations are less likely to achieve efficient outcomes for society. However, if this background on Chinese bureaucracy should sound slightly pessimistic, it shouldn't be surprising. Much of the world's nations have routinely come to somewhat negative conclusions about the efficiency of their own bureaucracies, albeit for different and specific reasons. For example, in the United States, the FY 2002 President's Management Agenda noted that, "Over the past three decades, reform initiatives have come and gone. Some genuine improvements have been made. But the record on the whole has been a disappointing one." (White House Office of Management and Budget [OMB], 2001, p. 8)

Management and budget reform initiatives for the public sector are often proposed, but to little success to date. This paper offers an approach to potentially enhance the efficiency of Chinese bureaucracy by using performance measures.

Using Measurement and Reward Systems to Improve Program Performance

Performance measurement is a tool to assess the quality of an agency's management and to motivate improved program administration. It is also associated with "results-oriented" budgeting (or "performance budgeting"), which aims to link budgetary decisions to changes in performance. Through the use of performance measures, a "principal" tries to encourage improved program management by focusing managerial and/or public attention on program outputs and outcomes, not just input levels. Using performance objectives could therefore generate improved program outcomes and enhanced accountability. The idea was to provide managers the flexibility and authority to produce appropriate goods and services while being held accountable for achieving the agreed-upon results (Rasappan, 2000).

In the United States, the FY 2002 President's Management Agenda (OMB, 2001) described the goals:

> Everyone agrees that scarce federal resources should be allocated to programs and managers that deliver results. Yet in practice, this is seldom done because agencies rarely offer convincing accounts of the results their allocations will purchase. There is little reward, in budgets or in compensation, for running programs efficiently. And once money is allocated to a program, there is no requirement to revisit the question of whether the results obtained are solving

> problems the American people care about. … Improvements in the management of human capital, competitive sourcing, improved financial performance, and expanding electronic government will matter little if they are not linked to better results. (OMB, 2001, 27)

Performance Measure Quality

The goal of a performance measurement system is to establish a de facto contract in which the agency receives funding based on actual measured results. Performance-based budgeting "would mean that money would be allocated not just on the basis of perceived needs, but also on the basis of what is actually being accomplished" (OMB, 2001, p. 7).

The criteria for good performance measures reflect the degree to which a program's progress can be assessed. A good metric would allow an outside entity to measure relative progress toward a program target, as well as in comparison with similar actors pursuing similar goals. However, government programs often include an array of complex, multidimensional performance elements. The central budget office would prefer to have a benchmark to compare against annual performance to establish appropriate progress. However, because of the complex nature of the government program, it may be difficult to find comparisons with other programs, whether in the public or private sector.[2]

To evaluate performance, one can rely on either outcome measures or output measures. Each has its advantages and disadvantages. Outcome measures aim to directly measure progress toward the actual program goals. Thus, a program may work to reduce water pollution, raise the incomes of poorer rural communities, or educate the public on public health. What sounds straightforward, however, is anything but. Measuring progress over time (i.e., evaluating progress based on past performance) can be very difficult. For example, a program's contribution may be overwhelmed by environmental factors. Or, it may be invariant in the short run or relates to deterrent behavior (OMB, 2003). Outcomes may be temporally distant or difficult to measure, and program impacts may be commingled with the effects of multiple programs. Although using the measures may move the agency in the right direction, the level of environment noise inherent in the measures causes it to be a poor motivator for improved management (Baker, 1992 & 2000; Feltham & Xie, 1994).

As a result, most outcome measures may not be readily contractible. It is well understood that evaluating agency management is difficult when progress to program goals reflects factors separate from the program's performance. Overall, for a performance measure to function as a useful metric, it must clearly show the effect of agency management decisions, otherwise it is useless.

The alternative to using outcome measures is to use output measures. The purpose of output measures is different. These behavior-based measures seek to reduce measurement uncertainty. Unlike outcome measures, which may be 'noisy,' output measures show specifically what the agency did. They indicate the number of plans written, the numbers of farmers visited, and the value of the fines processed. Measures may be clear, simple, and transparent, even less costly to monitor.

Unfortunately, whether these output measures are very reflective of the agency's mission is a separate question. Thus, the risk associated with relying on output measures to motivate performance is that the agency will be steered toward less-productive tasks. Output measures do not distinguish between unproductive busy-work and valuable contributions. Agencies are likely to be locked into the same production processes, just aiming to produce more of each of the output products (the plans, agreements, site visits, etc.) without regard to how these help accomplish the overall mission of the agency.

So, in using performance measures does the central budget agency have to choose between measures that are unreliable and those that distort the agency mission? Furthermore, if the program has multiple outcome measures, as well as multiple output measures, how should the central budget office treat these measures? How should the budget office respond to the plethora of different measures?

One idea–to exclude some measures–is not a good one. Consider the story of the blind men describing the elephant: when each was able to touch only a part of the elephant, their description was flawed. Only by collectively assessing the creature could they produce a good estimation of the animal's appearance. So too with performance measures; leaving out some measures leads to a distorted understanding of the program (Rose-Ackerman, 1986).

What happens when incentives are tied to an incomplete set of performance measures? The ensuing 'low-powered' incentive scheme reflects the case when an agent is rewarded or punished only a small fraction of his total performance (Holmstrom & Milgrom, 1991). The result is inefficient, with the agent emphasizing the more visible and highly rewarded job elements and ignoring others (Rose-Ackerman, 1986; Dixit, 1997).

Therefore, it is necessary to include all of the program elements in the performance measurement system. This is particularly true in China. In China, the commingled finances and performance measures of the bureaucratic businesses makes performance and financial evaluations difficult. To influence behavior, a central or regional budget office must consider not just the quality of individual performance measures, but how the bundle of measures is tied to the incentives.

One preferred solution is to combine different measures in an index of measures. The budget office can test the index's coherence by using Cronbach's Alpha as a measure of the index's internal consistency and reliability. Still, determining the weights to use in constructing the index is not obvious. Regression analysis (e.g., regressing an outcome measure on the various outputs, as adjusted by environment conditions) may help create weights. But in many circumstances, the index will need to be tested using simulations or other techniques.

One problem in using an index of different measures is that when different portions or aspects of a program show different degrees of efficiency and effectiveness, one cannot respond to the elements separately. For example, one program within the U.S. Department of Agriculture is charged with preventing soil loss, improving farm incomes, reducing water quality degradation, etc. Given these numerous goals, with a single index, it is not possible to react to whether the program components are performing effectively, or should be performing these functions better.

As long as program performance measures are generally isomorphic (i.e., moving in the same direction at a similar pace) and consistent with the same program missions, the use of the index encourages the agency to experiment with program design to maximize the index results. To the extent that the content of an index becomes less isomorphic and the weighting of performance measures in the index, subjective, the program results may be distorted. For example, if an agency is tasked with both responding rapidly to customer concerns and in fully satisfying the customers, the weight placed on each of the two performance elements in an index will determine the emphasis it receives. In the worst case, an agency may be diverted into less productive paths because of a poorly analyzed index.

An index cannot be effective if the program's measures are mutually exclusive, as indicated by a negative correlated in a correlation/covariance matrix. For example, the U.S. Forest Service's goals traditionally have included timber production and environmental protection; not surprisingly, this combination of missions has frequently led to conflicts. When the correlations between performance measures are negative, then combining the measures in an index will not produce meaningful incentives, just agency compromises based on the relative incentivization the measures have received.[3]

If performance measurement is doomed as a tool because the program is tool complex, with program goals operating at cross purposes, one solution is to redesign the program. Relying on program design as a way of creating programs that can be const efficiently monitored is a long standing strategy for the principals in the principal-agent relationship (Macey, 1992; Tirole, 1994). By reconstituting a program into more harmonious units, the resulting program structures may lead to improved performance in each new entity (McCubbins, 1990). The resulting organizational units – designed for isomorphic goal configurations—may develop better institutional cultures, as well, conducive to more enthusiastically pursuing the new program missions.

Management Tools as Incentives

What has been discussed thus far has been the development of a metric: a device to assess performance. It has been argued that an inexpensive monitoring capability and unbiased assessment device is a necessary starting point. This subsection describes how we can use these performance measures by tying it to a management or budget tool and so thereby motivate efficient agency behavior.

A few points to observe, however. First, incentives do not work if there are no opportunities to vary agency management decisions, and alter the program design or delivery. With sufficient incentives, agency heads are motivated to reconstruct the program. If the incentive is linked to a particular outcome, such as improving a river's water quality (i.e., less pollution), then an agency head is motivated to try many different approaches, whether punishing farmers who let cattle go into the river, taxing factories for polluting, or rewarding citizens who clean up waste in the river to best accomplish that outcome. Sometimes, however, an agency is required to carry out a program in a certain way or to deal with a particular population. In that case, the possible benefits of adding incentives are reduced. With reduced freedom of action come reduced potential benefits from incentives.

In the same fashion, the incentives provided have to be sufficient to overcome incentives from other sources. Public bureaus are often subject to multiple principals or interested parties with influence. Thus, the process of determining their funding is 'political' in that the multiple principals may each have different objectives. Consequently, two forms of inefficiency may result. First, the original program mission may be a compromise among the principals with each gaining some elements, but producing confused, contradictory, or complicated goals for the agency. Second, as part of the ongoing interaction among the agency and the principals, the formal program goals are inevitably augmented by hidden goals of the principals. Moreover, in government, with multiple principals, none of the principals acts a residual claimant. Revenues generated (created in the process of the bureau selling or distributing its output) are not to the clear benefit of one party. Thus, the true set of incentives facing the agency managers are based on their maximizing agency utility given the sum of the principals' formal and informal or hidden goals and corresponding potential rewards.

There are different tools that a regional or central budget or planning office can use to influence program management and agency incentives. Ultimately a performance measure would be linked to a specific motivating management tool. The tool can be linked to a single-outcome performance measure or to an index made up of a weighted average of multiple measures. The incentive must be sufficiently large to motivate behavior. If the agency is receiving significant financial resources from nongovernmental or other sources, the potential effect of a government incentive will be reduced. Incentive devices to affect agency management include:

Public disclosure and public pressure. Simply providing increased transparency on program management can augment performance. The goal then is to develop ways of increasing the quality and availability of standardized information about an agency's activities (e.g., the U.S. Government Accountability Organization's list of "High-Risk Agencies;" those agencies with especially poor financial management). The management incentive is brought about by the "sunshine" of public disclosure. For financial management, there should be an audit trail composed of visible actions taken by identifiable officials. Actions taken counter to regulations should be flagged for review. The more transparent the process, the more standardized the information, the cheaper it is to monitor.

Regulation/Mandates. Requiring that agencies adopt regulations that limit their discretion. The controls rule out less-effective program management and operations (e.g., requiring cost–benefit analysis for water projects). The performance measure reflects the status of the agencies' adoption of these regulations, which is relatively easy to measure. Similarly, for financial management, the more objective the process (e.g., an official must select the low bid in awarding a procurement contract), the fewer the rents available. Regulations are only as good as the state's ability and willingness to supervise and enforce the measures.

Material incentives through appropriation or taxation. Directly motivating agencies through the use of incentives. This approach is based on providing incentive payments to encourage procedural adoption, such as the use of a cost–benefit analysis, or to encourage improved program management (e.g., contingent funding available upon achieving program targets). Good management gets a financial or other kind of reward that could go to the program or to the group of individuals (Dixon, 2001).

It is important to distinguish between program needs and management. For example, a hospital that is under-performing may still warrant additional funding, even if the management is poor. One does not stop offering service to patients because of the hospital's poor leadership. One must have a separate incentive for the management, independent of the funding, and based on program or sectoral needs.

Market devices. Establishing competitive internal markets for determining program funding. Relies on the redundancy that ministries create during "empire building." When different bureaus or ministries offer analogous programs, they would then have similar program capacity. Thus, teams across agencies could compete to provide a service, whether managing a geographic site or providing a service, like an environmental review for which a contract would be awarded. A performance measure is needed as the rubric for assessing the basis for awarding the contract.

As initially described in this paper, many Chinese government agencies have created their own funding sources, legally and illegally, through new businesses, fees, fines, and donations. For a performance measurement system to function, all funds must be covered through the system, not just those funds that are placed on-budget. [4] It is difficult to argue that government efficiency is appropriate only for the visible transparent on-budget funds, but is an unnecessary luxury for off-budget accounts. Therefore, funds gained through fees and other sources would have restrictions placed on their use, just as are funds gained from appropriation. Likewise, revenue transfers between funds gained from appropriations, businesses, or fees should be controlled and maintained for their specific purposes. The source of the funding is not important in ensuring the efficiency of the expenditure.

In addition, because agencies are rewarded for their program outputs or outcomes, based on the performance measures and programmatic targets selected, if the agencies themselves are the primary authors of the performance measures, it is unlikely that useful measures will result. Agencies will design measures to best aid the agencies slated to receive additional resources. Thus, an outside entity is necessary to oversee first the design and selection of performance measures, then to allocate rewards so as to encourage agencies to take the measures seriously. The next subsection discusses the role of different outsider auditors or overseers.

Effective Personnel Management and External Oversight

Performance-oriented management (or really any good management) requires certain building blocks. Prerequisites include adequate core administrative support systems, such as public financial and personnel management systems (Reid, 1998). In addition, without adequate compensation for civil service employees, reforms can be undone in a hurry. For a system to be successful, one needs to start with the fundamental laws and rules encouraging good performance and controlling corruption, including: personnel laws (controlling against favoritism and nepotism in hiring); financial management (requiring standardized, auditable records of transactions, inspectors general offices), and procurement laws (no sole source contracting, competitive bidding, etc.).

These laws and rules do not operate on their own. That means that there must be a mechanism for the oversight and enforcement of laws and rules. In addition, there must be a reward for good program management. If the performance quality is ignored, largely in favor of political loyalty for example, then there is little reason for managers to emphasize their job performance as opposed to currying favor through gift-giving. Effective performance measures are only valuable if they are used.

China already has many of these formal laws and civil service procedures in place so that management problems have focused more on improving the implementation of these laws and regulations. For this, a Ministry of Supervision and the bureaus of supervision in the administrative agencies are empowered for oversight and enforcement. However, with respect to management concerns, this institutional arrangement has focused more on the investigations and criminal conduct, and less on administrative efficiency-enhancing procedures. Similarly, while the party discipline inspection committees have the authority to investigate and penalize behavior ranging from political, moral, and personnel misconduct, and fiscal mismanagement (Manion, 2004), they do not focus on general program management.[5]

To gain greater traction in increasing efficiency, performance measures are a beginning, but, in addition, external institutions should motivate their relevance, that is, to assess progress and establish incentives to help control against malfeasance, sloth, as well as corruption. In general, whenever the oversight entity is either internal to the organization or the hierarchy or engaged in repeat transactions and capable of building an extended relationship, then enforcement may be less effective. In China, the bureaucratic autonomy associated with *first-in-command* in a decentralized system means that the same entity is responsible for its own evaluation. Another problem results from the cadre system as program performance evaluation is carried out internal to the system, therefore independent program assessment is lost. In such a circumstance, the evaluation and enforcement of program management, procurement, and accounting procedures should be circulated to different oversight individuals, and rely on an unfamiliar, externalized relationship so that an agency or a cadre has a more difficult time assuming comfortable and dependable relations.

Chinese cities and other political jurisdictions can rely on a variety of different auditing and/or enforcement and reward systems to build efficient, impartial program delivery systems. Which institutional arrangement is chosen will also influence the type of management tools chosen. A few alternatives include:

Outside review by “residual claimant.” Basic procurement regulations should be drafted outside the ministry itself by a disinterested expert agency. For example, in Denmark, the National Audit Office, which is independent of the government and reports only to the Parliament, plays a prominent role in auditing performance information. Similarly, audits should be conducted by an outside entity. Moreover, to ensure appropriate incentives so that the reviewing or supervisory officials are not corrupted, one option is that the auditors receive a portion of the value for any corruption detected. This places the auditors in a comparable position of the residual claimant to the state/organization’s rents, and therefore, eager to find cases of corruption. The nature of the enforcement–reward or punishment–of the agency, after a reviewer’s evaluation and notification can come from various sources, such as the central budget or planning agency.

Competitive oversight. If multiple entities provide oversight over different agencies, the opportunity increases for a group to observe noncompliance with rules or to assess a penalty for poor performance. When the oversight entities are themselves rewarded for their unbiased assessment of poor performance or fiscal inadequacies (through promotions or bonuses), then managers in each agency are attentive to what their counterparts in other oversight agencies might observe and report (because it would reflect badly on them to be 'scooped'). In the U.S., for example, oversight agencies include the White House OMB, the Congressional Budget Office, the Government Accounting Office (GAO), and the Inspectors General in each Federal department. The agencies provide somewhat redundant functions, except are accountable to different principals. These agencies both evaluate and report on management and performance. Some are in a position to directly reward or punish the agency, whereas some rely entirely on disclosure.

Community "fire alarms" as oversight mechanisms. Too often the focus on poor performance or corruption considers only the dyadic environment of the agency and program recipients, or bribe taker and the briber. In fact the environment is larger. Acknowledging the high cost of direct monitoring, the state can instead turn to those who may be more directly knowledgeable, the community. Thus, community members and local community groups can act as watchdogs or "fire alarms," reducing information asymmetries held by the government bureau by involving the public and using their knowledge of government performance (Epstein & O'Halloran, 1995). Nongovernmental community organizations could be selected in various communities to conduct customer surveys and rate their bureaucracies. Directed by an external central government management agency, the process would include personnel awards or agency budget increases going to the highest-rated government bureaus.

Commitment Devices

In providing incentives for program managers, the certainty of the reward or punishment for the agency matters. Unlike a salary, the basis for which is clearly understood by both parties, the performance contracts often do not establish certainty of a reward for good management (or punishment for bad management).

Discretion is the enemy of powerful incentives. The more a performance reward system appears inflexible and objective, the more seriously the agency officials will treat it. The reward for management success, as evidenced by the performance measures, should not vary based on the political interests or current preferences of political officials. If productivity is not rewarded by the system, agency performance will not improve (Klay, 1987; Swiss, 2005).

If a performance management system is meant to generate the visible, predictable consequences for agency behavior the multiple principals must create binding commitments that ensure that they honor the *de facto* performance contract. The more their interests can be harmonized and the reward system formalized, the more effective the management system.

The commitment devices can rely on personnel, administrative, or financial rewards. The principals can embed commitment devices in performance contracts, relying on either informal inter-organizational agreements or formal regulations (e.g., published in an official venue). Finally, agreements can rely on self-enforcing mechanisms, like internal markets, where appropriate.

Conclusion

There is a special environment of for all public sector bureaucracies, which generates incentives, particular to government. The public sector is usually characterized by a combination of funding (at least in part) from taxes and regulatory actions (not just product sales); pervasive asymmetric information between principal(s) and the bureaucratic agent; multiple program goals (not just profitability); and the presence of multiple principals for a given bureaucratic agent. As a result of these influences, government agencies are generally subject to fiscal and managerial incentives that prompt programmatic inefficiencies. Are there ways to overcome these dysfunctional influences so that government agencies can enhance their performance?

This paper examined whether performance measures can serve as a means for evaluating bureaucratic performance and program design. That answer depends on several factors. The first challenge comes from what is being measured. Many of the programs offered by public agencies are distinguished by complex and multiple goals. In addition, these complex goals may be invariant in the short run, complicated by a lack of external benchmarks for comparisons, subject to measurement error, or reflect the contributions of many programs (thus confounding an assessment of individual program contributions). As a result, it is difficult for a principal (or outside watchdog interest groups) to evaluate bureau performance. When one can only poorly assess program performance, the quality of management decisions, or even know which agencies are primarily responsible for outcomes, it is difficult to gauge performance and accordingly, to encourage accountability and efficiency.

As a result of the imperfect, unreliable information generated through the performance measures, measurement error is inevitable. Thus, Type I and Type II errors (funding unworthy programs or not funding worthy programs based on incorrect performance measurements) will lead to funding inefficiencies.

The central budget office faces an additional type of challenge. Governments have an inability to guarantee funding according to performance. When program needs change, based on societal needs, funding for a program generally declines, even if program performance was excellent. Allocations to the highest priority purposes may be inconsistent with maintaining performance incentives. In addition, it is often unclear how a central budget office should react to a program's failure to achieve goals listed in a strategic plan. Should the offending agency be penalized with less money, even though supporters assert that they need more funding to achieve the performance goals?

This paper described an approach that the Chinese government can take to potentially yield improved management and better program outcomes. It is not a foolproof system. A central budget office must assess each program to determine the goals to be monitored. The performance measures generated need to be complete and cover all elements of the program's mission. To leave out any program element is to ensure that it does not get done.

The resulting set of performance indicators need to be contractible. One way to simplify the array of different performance measures is to consider them as part of a package, not viewed separately and independently from each other. The central budget office can use an index combining outcome and output measures.

The central or regional budget office must match the weighted performance measurement index to budget or management instruments. The accomplishment of performance targets/strategic goal is tied to rewards. Part of the unfortunate consequences of a performance budgeting system is the risk of unintended consequences, which lurks whenever a budget line is tied to a performance measure, while neglecting or undervaluing other program performance elements.

On top of the design of the performance measures and incentives, several additional elements stand out for each of the actors in the process:

The agency: Can an agency react to performance information by altering program design elements or vary the program delivery mechanism? It is important to know if the agency supposed to improve performance based on the new incentivized measures is empowered to do so.

The budget office: Can the budget process distinguish between sectorally derived program needs and management-driven rewards, between the consequences of environmental changes and the consequences of administration on program performance? In addition, are the incentive devices conditional or inflexibly committed? Performance must be rewarded. Program managers will discount any incentives that are conditional on exogenous conditions. The commitment devices need to incorporate the reactions of all the relevant principals who are important in motivating agency behavior.

Oversight: Independent, motivated watchdogs are important to reduce informal dealmaking and encourage transparency to help reduce incentives that undermine the performance system.

A point of caution however is useful in viewing these instruments. For the United States, the FY 2002 President's Management Agenda concluded that:

> In 1993, Congress enacted the Government Performance and Results Act (GPRA) to get the federal government to focus federal programs on performance. After eight years of experience, progress toward the use of performance information for program management has been discouraging. According to a GAO survey of federal managers, agencies may, in fact, be losing ground in their efforts to build organizational cultures that support a focus on results … Agency performance measures tend to be ill defined and not properly integrated into agency budget submissions and the management and operation of agencies. Performance measures are insufficiently used to monitor and reward staff, or to hold program managers accountable … [and] the General Accounting Office reported that the majority of federal managers are largely ignoring performance information when allocating resources. (OMB, 2001, p. 27)

The point of including the above description is that the process is difficult and success is uncertain. Contrary to many of the claims of New Public Administration enthusiasts who advocate the widespread use of performance measures, aligned with budget system appropriations as a means of obtaining increased program efficiency and effectiveness, performance budgeting successes are often more anticipatory or anecdotal than actual. In general, there is often little in the way of empirical evaluation of management tools before states have attempted their widespread implementation.

As a result, it is invariably better to phase-in the implementation process. Moreover, it may be useful to rely on an experimental design for testing administrative reforms. Similar programs in different regions can be matched. Then, one of the programs would be randomly assigned to act as control or treatment, in which the treatments would reflect the attempt to tie management/budget tools to program goals. Then, a program under one performance system can be compared to a similar program under another. Using this type of approach, the efficiency of management innovations can be tested before they are implemented, rather than the current approach of implementation and subsequent argument and justification.

Author

Stuart Kasdin is an assistant professor of public policy and public administration at the Trachtenberg School of Public Policy and Public Administration, George Washington University, Washington, DC. He previously worked at the Federal Office of Management and Budget. He can be contacted at skasdin@gwu.edu.

References

Baker, G. (1992). Incentive Contracts and Performance Measurement. *The Journal of Political Economy*, 100, 598–614.

Baker, G. (2000). The Use of Performance Measures in Incentive Contracting. *The American Economic Review*, 90, 415–420.

Baker, G. (2002). Distortion and Risk in Optimal Incentive Contracts. *The Journal of Human Resources*, 37, 728–751.

Congressional Budget Office. (1993). *Using Performance Measures in the Federal Budget Process*. Washington, D.C.: Congressional Budget Office.

Dixit, A. (1997). Power of Incentives in Private versus Public Organization. *The American Economic Review*, 87, 378–382.

Dixon, G.L.R. (1996). "Budgeting Institutions and Expenditure Outcomes in Australia. " Paper prepared for The World Bank study of The Impact of Budgetary Institutions on Expenditure Outcomes. Geoff Dixon and Associates

Epstein, D. & O'Halloran, S. (1995). A Theory of Strategic Oversight: Congress, Lobbyists, and the Bureaucracy. *Journal of Law, Economics, and Organization*, 11, 245–255.

Feltham, G. & Xie, J. (1994). Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations. *The Accounting Review,* 69, 429–453.

Government Accounting Office. (1994). "Management Reforms: Examples of Public and Private Innovations to Improve Service Delivery." GAO/AIMD/GGD-94-90BR. Washington, DC: GAO.

Government Accounting Office. (1994). "Managing for Results. State Experiences Provide Insights for Federal Management Reforms." GAO/GGD-95-22. Washington, DC: GAO.

Government Accounting Office. (1995). "Managing for Results: Experiences Abroad Suggest Insights for Federal Management Reforms." GAO/GGD-95-120. Washington, DC: GAO.

Government Accounting Office. (1997). "Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation." GAO/AIMD-97-46. Washington, DC: GAO.

Government Accounting Office. (1999). "Performance Budgeting: Initial Experiences Under the Results Act in Linking Plans with Budgets." GAO/AIMD/GGD-99-67. Washington, DC: GAO.

Government Accountability Office. (2003). "Statement of Paul L. Posner. Performance Budgeting: Current Developments and Future Prospects." GAO-03-595T. Washington DC: Testimony.

Government Accounting Office. (2004). "Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget." GAO-04-174. Washington DC: GAO.

Hamrin, C.L. & Zhao, S. (Eds.). (1995). *Decision-*

making in Deng's China. Armonk, NY: An East Gate Book.

Holmstrom, B. & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. J*ournal of Law, Economics, and Organization*, 7, 24–38.

Klay, W.E. (1987). Management Through Budgetary Incentives. *Public Productivity Review.* 10(3)59–71.

Lee, P.N.S. & Lo, C.W.H. (Eds.). (2001). *Remaking China's Public Management.* Westport, CT: Quorum Books.

Lü, X. (2000). *Cadres and Corruption*. Stanford, CA: Stanford University Press.

Ma, S. (1996). *Administrative Reform in Post-Mao China*. Maryland: University of America Press.

Macey, J.R. (1992). Organizational Design and Political Control in Administrative Agencies. *Journal of Law, Economics, and Organization*, 8, 93–110.

Manion, M. (2004). *Corruption by Design*. Cambridge, MA: Harvard University Press.

McCubbins, M. (1990). Note: Budget Policy-making and the Appearance of Power. *Journal of Law, Economics, and Organization*, 6,133–153.

Meyers, R. (Ed.). (1999). *Handbook of Government Budgeting*. San Francisco: Jossey-Bass, Inc.

National Association of State Budget Officers. (1999). *Budget Processes in the States.* Washington, DC: National Association of State Budget Officers.

Office of Management and Budget. (2001). *The President's Management Agenda.* Executive Office of the President, Office of Management and Budget. Retreived September 2004 from http://www.whitehouse.gov/omb/budget/fy2002/mgmt.pdf.

Rasappan, A. (2000). *Budget Reform – Malaysia*. Report to World Bank. ARTD (Malaysia). Management, Evaluation, & Research Associates

Reid, G. (1998). Performance-Oriented Public Sector Modernization in Developing Countries Meeting the Implementation Challenge. *PREM* Washington, DC: The World Bank.

Rose-Ackerman, S. (1986). Reforming Public Bureaucracy Through Economic Incentives? *Journal of Law, Economics, and Organization*, 2, 131–161.

Shirley, M.M. & Xu, L.C. (1997). Information, Incentives and Commitment: An Empirical Analysis of Contracts between Government and State Enterprises. *World Bank Policy Research Working Paper No. 1769*. Washington, DC: The World Bank.

Shirley, M.M. & Xu, L.C. (1998). The Empirical Effects of Performance Contracts: Evidence from China. *Policy, Research working paper; no. WPS 1919*. Washington, DC: The World Bank.

Shirley, M.M. (1998). Why Performance Contracts for State-Owned Enterprises Haven't Worked. *Public Policy for the Private Sector*. Note No. 150. Washington, DC: The World Bank.

Swiss, J.E. (2005). A Framework for Assessing Incentives in Results-Based Management. *Public Administration Review*. 65(5)592–602.

Tirole, J. (1994). The Internal Organization of Government. *Oxford Economic Papers*, 46, 1–29.

White House. (2003). "The President's Management Agenda: The Five Initiatives. Budget and Performance Integration." Retrieved September, 2004 from www.results.gov/agenda.

Zheng, S. (1997). *Party vs. State in Post-1949 China: The Institutional Dilemma*. Cambridge: Cambridge University Press.

Notes

[1] An earlier version of this paper was presented at the Conference on Public Governance in Urban Communities at Shenzhen University in 2005.

[2] Agencies have an incentive to choose performance measures strategically to avoid competition. As a result, benchmarks are often not readily available in the private sector or even other programs.

[3] In such a case, where accomplishing one goal requires undermining another program goal, it is often best to try to split the tasks among different parties. By separating responsibilities, each entity can pursue its goals directly.

[4] One approach that some agencies may use with available off-budget accounts is to separate their high-performing projects into the on-budget program account and then set aside their low-performing, less justifiable projects for the off-budget account. Allowing this form of budgetary flexibility for an agency encourages inefficiency, favoritism, and corruption. It also makes the performance measurement system less meaningful.

[5] Moreover, oversight impact may be diluted in anti-corruption enforcement due to the substitution of party disciplinary action instead of harsher criminal penalties (Manion, 2004, p. 139).