Assessing governance: the importance of evaluating policy outcomes in national mission areas

Craig Boardman
Assessing Governance: The Importance of Evaluating Policy Outcomes in National Mission Areas

ARTICLE in GOVERNANCE · JULY 2014
Impact Factor: 1.54 · DOI: 10.1111/gove.12083

1 AUTHOR:

Craig Boardman
35 PUBLICATIONS 499 CITATIONS

Available from: Craig Boardman
Retrieved on: 31 July 2015
Research Note

Assessing Governance: The Importance of Evaluating Policy Outcomes in National Mission Areas

CRAIG BOARDMAN*

This commentary responds to Fukuyama’s (2013) rejection of “final outputs” as measures of government quality. It argues that public administration research should address policy outcomes in “national mission areas.” But public administration scholars should not simply become policy analysts. Rather, they should become policy analysts for policy areas that can benefit from expertise in organizations and management.

Introduction

In his recent commentary on the state of empirical measures of government “quality” or performance, Fukuyama (2013) rejects “final outputs” or policy outcomes as measures of government quality and recommends that public administration scholars turn their collective attention to evaluating government quality in terms of “capacity” (tax rates, bureaucrats’ education levels) and “autonomy” (executive). This rejection is premature because historically there has been very little public administration research focused on policy outcomes.

But this does not mean that this sort of research, while challenging, cannot be done, and in a rigorous manner. Public administration research can and should turn at least some of its focus to national mission areas—which represent collections of policy goals and objectives for which there is national consensus and resources, programs, and activities targeted at achieving those goals and objectives—for two reasons.

First, national mission areas of public policy (which can vary from one country to the next) represent the values of a society and what it expects from its government. In turn a particular government’s quality can and should be assessed not just in terms of its capacity and autonomy (as Fukuyama suggests), but additionally in terms of the outcomes its society values and expects.

*The Ohio State University

Second, national mission areas in many cases are amenable to the valid and reliable measurement of policy outcomes. In many of these areas data for policy outcomes are readily available. And these data do not suffer from the typical ailments of government performance data—they are objective rather than self-reported and tracked with regularity over the course of many years rather than for just a single point in time.

Of course neither of these reasons are reason to jettison Fukuyama’s emphasis on measuring government quality additionally in terms of what amounts to national policy inputs and processes. As he rightly observes, these can be important antecedents to policy outcomes. But bureaucratic professionalism, tax rates, and executive autonomy are at best indirect proxies for the public and quasi-public goods that a government intends to provide.

A Direct Response to Fukuyama

It is understandable why Fukuyama arrives at his conclusion to reject output measures in the study of public administration. To his overarching point with regard to output measures, there are several major drawbacks to their use. However, these drawbacks are not as decisive as he suggests.

First, Fukuyama correctly observes that policy outcomes are not solely the products of government, but rather that they can be relatively messy insofar that a number of socioeconomic and political factors, not just government action or policies, are involved in their production. Yet this particular drawback can be extended to both the capacity and autonomy measures he condones, such as tax extraction rates, the education levels of government officials, and expert perceptions of executive autonomy—all of which, as objects to measure, too are affected by exogenous factors.

What’s more, Fukuyama implicitly suggests that perhaps at least some “final outputs” may be affected by public policies. Then is it not incumbent to at least attempt to understand the extent to which governmental actions are responsible for particular ends, both positive and negative? The field of public policy studies is devoted to this purpose, as evidenced by high-impact journals like *Journal of Policy Analysis and Management* and *Policy Studies Journal* as well as by even higher-impact journals emphasizing policies in national mission areas, such as *Energy Policy* and *Research Policy*.

Next, Fukuyama correctly observes that policy outcomes are at times difficult to measure validly and reliably and not always the most important things to measure. This may in fact be the case for the examples he provides—educational outcomes—and policies allowing the torture of criminals—but it is not necessarily the case for all areas of public policy and in particular for policies in national mission areas like energy and science and technology. In these latter policy areas (which are discussed further in the text), there are readily available outcomes data that are not only valid and reliable, but additionally can be directly affected by particular policies.
Last, Fukuyama correctly observes that econometric techniques can be problematic (though he does not go into specifics). But, again, this charge is so general that it can be applied to the areas of measurement he condones as legitimate operationalizations of government quality. Moreover, the charge implicitly invalidates the field of public policy studies as well as any other field that cannot readily employ experimental research as a standard. Though econometric challenges may be greater in the study of government than in the study of, say, private businesses (for which valid and reliable outcomes data are easier to come by), these challenges are partly what invigorate scholars to study policy outcomes in the first place.

All of that said, it becomes incumbent to provide counterexamples to those provided by Fukuyama. The rest of this article addresses areas of national policy that can benefit from public administration research emphasizing “final outputs.” The article also discusses some specific data sets and research designs for these policy areas. It should be noted that this is a somewhat broader argument regarding outcomes-focused research than that made by Fukuyama in that it emphasizes not just policy outcomes but additionally the importance of understanding organizational and managerial variation at policy implementation.

First, however, it is important to discuss briefly what constitutes a national mission area.

What Is a National Mission Area?

A potential observation to make is that emphasis on the outcomes of national policies talks past Fukuyama’s interest in government quality, specifically executive branch quality, and predominantly for the developing part of the world. This is not the case. Fukuyama acknowledges that policy outcomes are a function (at least in part) of executive branch autonomy and capacity. What’s more, the developing part of the world, though it may not have the same national priorities as, say, the USA or Russia, no less has priorities of its own. National mission areas are direct reflections of the priorities of government, and in particular of its executives, and should be central to any attempt to address government quality.

This begs the question what constitutes a national mission area. There is no exact answer; there are a number of strategies for identifying areas of national priority. Whichever strategy one takes, at best the label “national mission area” is to be used as an heuristic for the policy priorities of an executive branch, which can shift and even change dramatically with the times. The purpose of using the label here is for public administration scholars to think more broadly when identifying areas of government to study, that is, areas where government quality may be measured in terms of policy outcomes as well as inputs and processes (including but not limited to Fukuyama’s measures for autonomy and capacity).
Toward Better Decision Making

There are a number of national mission areas that stand to benefit greatly from public administration research emphasizing policy outcomes. And this benefit can be fully realized only by including as measures of quality for executive branches and their agencies the extent to which they achieve specific policy outcomes. Take efforts by the U.S. National Institutes of Health (NIH), Department of Energy (DOE), and National Science Foundation (NSF) to address socioeconomic problems like AIDS, energy independence, and more broadly U.S. scientific and technological competitiveness.

Longitudinal data for assessing the creation and quality of new knowledge and technological innovations in these policy areas are available to public administration researchers via a number of sources, including but not limited to Web of Science, Scopus, Engineering Village, STN International, MEDLINE, INSPEC, USPTO, and Derwent World Patents. This list is not comprehensive and is only focused on one type of policy (for new knowledge creation and technological innovation), but are examples of readily available data for tracking policy outcomes longitudinally, objectively, and validly and reliably.

The longitudinal nature of these data implies research designs that can address the net effects of policies in these national mission areas (not just for the USA, but for any country). NSF, DOE, and NIH endeavors in these areas may be addressed in a variety of ways that can alleviate most of the concerns about exogenous factors that Fukuyama highlights.

One option is the multiple baseline interrupted time series research design that is often used in clinical trials for new medicines by pharmaceutical companies when randomized group trials are infeasible or too costly. The design is considered relatively strong for accounting for alternate exogenous explanations of outcomes (Campbell and Stanley 1963) and is common in public policy analysis, for example, when multiple states implement the same policy but at different points in time.

For the NIH, DOE, and NSF examples, the design can be used to address similar or the same policy implementation mechanisms for coordinated problem solving in the areas of health, energy, and science—specifically, to address the extent to which these mechanisms lead to new knowledge and new technologies over time. But depending on the focal policy, other more sophisticated designs can be used.

For example, regression discontinuity design—which is as “strong” as randomized experiments (Trochim 2005) and frequent in assessing the educational outcomes Fukuyama identifies as problematic—can be employed, perhaps in cases (in addition to education) of ecosystems management and smart energy grids when failure to meet a quantifiable standard results in a policy intervention. Field experiments are also becoming more common in particular areas of policy analysis.
Why Public Administration Scholars?

Public administration scholars should not simply become policy analysts. Rather, they should become policy analysts for policy areas that can benefit from public administration scholars’ expertise in organizations and management—to connect Fukuyama’s measures for government quality in terms of inputs and processes to measures of policy outcomes. In this sense public administration scholarship can help to bridge the gap between public policy studies and public administration.

Related, public administration scholars should become policy analysts in national mission areas because more often than not these policies are quite difficult to implement, due to the increasing organizational and managerial complexity of policy design and implementation in these areas. Blame for such difficulty is oftentimes attributed to executive branches and their agencies, fairly or not. Whether or not one agrees that executive branches should be judged in terms of policy outcomes, typically they are judged in this manner.

Continuing with the NSF, DOE, and NIH examples, each agency regularly establishes new organizations to coordinate and manage unique sets of experts and other resources to facilitate policy outcomes. These experts and resources come from different institutions and economic sectors and from different scientific and engineering disciplines, and therefore can be difficult to coordinate and manage due to participants’ diverse goals and resource bases.

For policymakers, these (and numerous other, similar) efforts by U.S. agencies beg the question whether all such organizations should be designed and implemented the same way. The answer for public administration scholars is of course “no,” but there is little guidance on how such organizations should be implemented differently.

General theories of policy design and implementation account for only the most basic of contingencies and therefore are of little use. Contingency theories of organizing, moreover, for example, resource and institutional approaches, can in many cases fail to provide insight. Even policy implementation theories emphasizing the processes of new organizations in national mission areas themselves fare no better, at least when it comes to ex ante decision making.

To inform ex ante decision making in national mission areas, public administration scholars must assess both processes and outcomes. Outcomes data are only one side of the model specification and public administration scholars must assess antecedents to policy outcomes, specifically organizational and other structural features affecting policy implementation.

Caveats and Closing Remark

Public administration scholars can and should conduct policy evaluations in national mission areas and in doing so address government quality.
However, this is not to recommend that public administration scholars abandon more traditional emphases on the workaday processes of government agencies. Indeed, it is the latter that makes public administration scholars uniquely qualified to conduct the sort of policy evaluation promoted here.

Moreover, this article should not be interpreted as a rejection of Fukuyama’s emphases on assessing government quality: autonomy and capacity. The research agenda discussed above is a complement to, not a substitute for, Fukuyama’s recommendations, including for developing countries.

For developing countries, national mission areas may be different than for the USA but will include common emphases such as public health and education. Though in some cases “final outputs” may be autonomy and capacity themselves, the establishment of an autonomous and capable executive branch is not the only way to measure government quality for developing countries.

This brings up the issue of comparability of government quality across nations. Though developing and developed countries can have different national missions, as just mentioned there will be some overlap. Moreover, each subset of nations for the most part has common national pursuits, as evidenced by Organisation for Economic Co-operation and Development studies comparing developed nations in terms of scientific capacity and economic performance and the numerous indices for comparing developing nations in terms of democracy and corruption.

In closing, Fukuyama has written a thoughtful piece regarding the measurement of government quality. But public administration scholars should reject his rejection of policy outcomes as measures of national performance. The rigorous study of policy outcomes is not only feasible but is already happening. And policy outcomes in national mission areas have been and will continue to be the benchmarks by which nations self-assess and compete internationally.

Acknowledgments

Thank you to Trevor Brown and Stephane Lavertu for commenting on earlier versions of this article.

Notes

1. To cite a specific example of rigorous research in education policy (which Fukuyama singles out as being too problematic in terms of exogenous factors from a research design perspective), see Heckman et al.’s (2010) assessment of the HighScope Perry Preschool Program.
2. Actually this is not the case for education. See the discussion below.
3. Some potential decision rules come to mind for setting boundaries separating “national mission areas” from other areas of policy. One potential decision rule is to count only policies promoted by cabinet-level departments,
for example by the U.S. Departments of Energy, Defense, State, etc. However, this approach excludes historic mission agencies established legislatively by Congress, like the National Science Foundation, National Institutes of Health (NIH), and NASA. Another potential approach is to “follow the money” and address only policies implemented by the highest-funded U.S. departments and agencies—but this approach excludes both cabinet-level agencies like DOE and independent federal agencies like NIH, suggesting that neither energy nor health are currently of national import. A third potential approach is to set a decision rule based on recent polls of public opinion, but this produces a set of focal departments and agencies so diffuse as to provide little if any guidance; for example, “the economy” as a perceived public problem is inseparable conceptually and practically from a number of socioeconomic issues, government agencies, and public policies, including but not limited to health, energy, environment, entitlements, and foreign affairs.

4. Think NSF Nanoscale Science and Engineering Centers and I-Corps Hubs, DOE Energy Frontier Research Centers and Energy Innovation Hubs, and NIH Nanomedicine Development Centers and Translational Research Centers (to name just a very few).

5. For example, Matland’s (1995) contingencies “ambiguity” and “conflict” for a policy’s “means” and “goals” are useful for distinguishing the management challenges of a new NSF, DOE, or NIH organization from the management challenges faced by other policies not emphasizing boundary-spanning and knowledge-intensive policy implementation, but are not so useful for addressing “within-type” differences, for example, whether a particular organization should be structured and managed at policy implementation as were previous organizations.

6. In many instances NSF, DOE, and NIH research centers can lack conventional governance mechanisms for coordinated problem solving involving diverse actors (Boardman 2012).

7. For example, as each new organization is established in response to a unique problem for which a “uniquely qualified” set of human capital and other resources has yet to be assembled (Zerhouni 2003), in theory (if not in practice) this means that at design and implementation each should be unique structurally and managerially, due to different types and numbers of boundaries spanned, different goals, different if not divergent contextual/institutional factors, and thusly different coordination/transaction costs.

8. This entails in addition to the use of extant outcomes data the collection of new ancillary data specifying what in terms of organization and management theories should covary with outcomes, including but not limited to “treatment” variables tracking variation across the organization and management of policy implementation.

References


