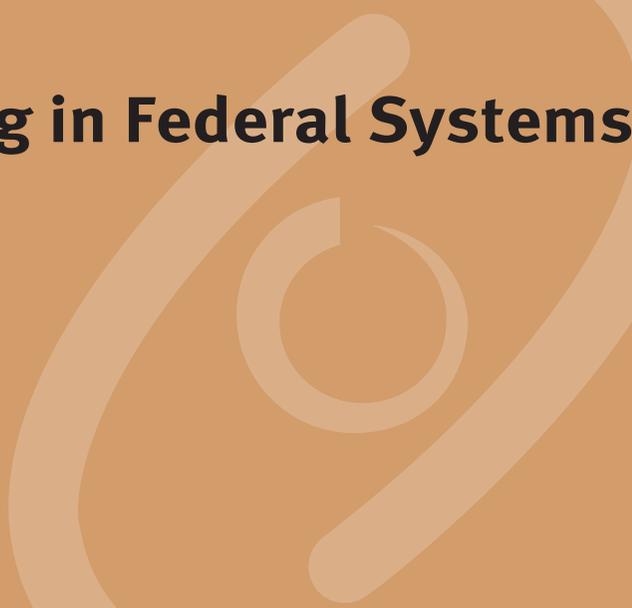


Occasional Paper Series

Number 6

Benchmarking in Federal Systems

Alan Fenna



BENCHMARKING
IN FEDERAL SYSTEMS

ALAN FENNA

© Forum of Federations 2010
ISSN: 1922-558X (online ISSN 1922-5598)

Occasional Paper Series Number 6
Benchmarking in Federal Systems
By Alan Fenna

For more information about the Forum of Federations and its publications,
please visit our website: www.forumfed.org.



Forum of Federations
700- 325 Dalhousie
Ottawa, Ontario (Canada) K1N 7G2
p. 613-244-3360
forum@forumfed.org

Abstract

Recent years have seen a rapidly growing interest in the use of benchmarking arrangements to improve policy performance in federal systems. This is a new development and one that is in its very early stages, but there is no doubting its significance. At issue here is the intersection of two things: a particular form of government and a particular form of management. Each is a complex matter in itself. How compatible is benchmarking with principles of federalism; and to what extent benchmarking can 'add value' to existing federal arrangements either by offering a superior mode of intergovernmental relations and/or by generating better substantive results for citizens? This paper looks at various benchmarking experiences in OECD-type federations and the European Union and draws tentative conclusions as to how complementary federalism and benchmarking might be.

Keywords: Benchmarking; Europe; federal; government; intergovernmental; management; performance; policy; OECD.

Benchmarking in Federal Systems

Alan Fenna

Recent years have seen a rapidly growing interest in the use of benchmarking arrangements to improve policy performance in federal systems. This is a new development and one that is in its very early stages, but there is no doubting its significance. If clear indication were lacking, that changed in August 2009 with the insertion of Article 91d in the German Constitution: “The Federation and the States [*Länder*] may, to establish and improve the performance of their administrations, conduct comparative studies and publish the results.”¹ How much legal or practical impact such an ambiguous clause will have is unclear; what is clear, though, is its symbolic significance — benchmarking is now a recognized device of modern federalism.

At the same time, the issue is far from straightforward. Only a year after the Germans adopted Article 91d, the incoming government of the United Kingdom announced that it would wind up its Audit Commission, the body that for two decades carried primary responsibility for the bulk of performance monitoring and benchmarking of local government in England.² The juxtaposition of these two events suggests something about the complexity of the issue: if benchmarking has been called into question in the UK where it had been instrumental in driving reform, what is its future in a federal context? At issue here is the intersection of two things: a particular form of government and a particular form of management. Each is a complex matter in itself. The question is how complementary they might be. How compatible is benchmarking with principles of federalism; and to what extent benchmarking can ‘add value’ to existing federal arrangements either by offering a superior mode of intergovernmental relations and/or by generating better substantive results for citizens? In answering these questions one is inevitably drawn into consideration of the very different forms that both federalism and benchmarking can take.

This chapter provides an overview of the problem considered as four propositions:

1. Benchmarking is a logical but challenging mode of public sector governance which comes in different forms and carries with it a number of risks or limitations;
2. Federalism is a very specific form of government predicated on

1. Grundgesetz für die Bundesrepublik Deutschland, Art 91d: Bund und Länder können zur Feststellung und Förderung der Leistungsfähigkeit ihrer Verwaltungen Vergleichsstudien durchführen und die Ergebnisse veröffentlichen. Inserted by amendment 1 August 2009.

2. “Eric Pickles to disband Audit Commission in new era of town hall transparency,” London: Ministry of Communities and Local Government, 2010. The decision applies to England only; Scotland and Wales have each had their own equivalents.

well-established norms and promising certain advantages but also one where significant differences in practice from one instance to the next make direct comparison difficult;

3. In principle, there is a strong affinity between federalism and certain types of benchmarking;

4. In practice we find that the experience with benchmarking in federal systems has taken various forms depending on what its purpose is; who carries it out; how hierarchically it is implemented; and what balance is maintained between quantitative and qualitative dimensions.

Challenges of Benchmarking in the Public Sector

‘Benchmarking’ is a term that is used rather loosely and takes on at least two somewhat different meanings, one more demanding than the other. In the broader sense we can understand benchmarking simply to mean the comparative measurement of performance. In the fuller or more specific sense we can understand benchmarking to mean the use of comparative performance measurement as a tool for identifying and adopting more efficient or effective practices.³ In the former sense it is an assessment device; in the latter it is a learning and adjustment tool.

Archetypes: external and internal benchmarking

Benchmarking originated in the private sector, which supplies us with the classic model of what the practice is about. In Type 1 benchmarking, an individual firm finds a way to assess performance of some aspect of its enterprise against industry leaders in other sectors and learns from that comparison how to improve its practices. The assessed and the assessors are effectively one and the same. Such ‘external’ benchmarking is voluntary or self-directed and oriented solely toward learning.

There is another model from the private sector, though, and that is the ‘internal’ one: central management imposes benchmarking requirements on the firm’s constituent units as a way of driving improvement through internal competition. This Type 2 benchmarking thus mimics the market forces that had been displaced internally by the creation of the business enterprise in the first place. In this version, the assessed and the assessors are different, and the former are subject to sanctions imposed by the latter. This is neither self-directed nor focused on learning as far as those individual units are concerned; rather, it is top-down and coercive, focusing on performance monitoring and involving some kind of sanctioning device. Such internal performance monitoring equates to benchmarking in the broad or loose sense of the term.

3. Gregory H. Watson, “A Perspective on Benchmarking,” *Benchmarking: an international journal* 1 (1994):5–10.

From private to public sector

The evident value of performance comparison, identification of best practice, and commitment to learning and improvement makes benchmarking an attractive proposition for the public sector as for the private. In Osborne and Gaebler's oft-cited argument, "what gets measured gets done," and if public sector agencies start measuring what they do they will find ways to do it better.⁴ Moreover, if governments can shine the spotlight of performance measurement onto the things that ultimately count the most – what government *achieves* as distinct from what it merely *does* – then presumably they will find a way to achieve more.

However, while alike in some regards, the public and private sectors are distinctly unlike in others. Indeed, in two very important respects the differences are fundamental. First, governments and their various agencies are not profit-driven enterprises engaged in a competitive struggle for business and survival in the market place. Second, their *raison d'être* is to achieve impact or *outcomes* in society rather than merely output. Thus the public sector has neither the same *imperative* nor the same *capacity* for benchmarking as the private sector.

Overcoming the incentives problem

The first of these challenges, the lack of intrinsic incentive, is in some ways precisely the reason for introducing benchmarking – just as it has been for internal corporate benchmarking. Performance monitoring and the imposition of benchmarking requirements is a public sector surrogate for market forces. This may be initiated by an individual agency to improve its own performance – external benchmarking – but given the lower level of intrinsic incentive and the greater difficulties, such action is likely to be the exception to the rule. In reality, the lower level of incentive means that public sector agencies are more likely to need such requirements to be imposed on them. Thus Type 2, internal, benchmarking is more likely in a public sector environment.

Internal benchmarking operates via sanctions – which, in the private sector, appear in the form of decisions about capital allocation. In that sense, it is a coercive device. In the public sector, sanctions might take a number of forms of which two are particularly prominent. One, following the private sector lead, relies on financial penalties and rewards. There are drawbacks to financial penalties, however — among them the distinct possibility that substandard performance may require more, not less, resource input to address. A multi-site corporation is free to let its underperforming sites wither and die; governments are not. Hence, then, the attraction of a quite different form of sanction: the political device of *naming and shaming*. Here the exercise has an external audience whom, it is assumed, can be reached effectively and will respond in a way that has the desired sanctioning effect.

4. David Osborne and Ted Gaebler. *Reinventing Government: how the entrepreneurial spirit is transforming the public sector*. Reading MA: Addison-Wesley, 1992. p.146

Any form of sanctioning creates incentives for behaviour that is perverse to the intentions of the benchmarking regime. Honest efforts to generate the desired results as reflected in the measurement criteria may take the form of *teaching to the test*. In such cases, the overall purpose is forgotten in efforts to achieve the measured targets. Dishonest efforts take the form of *gaming*, whereby the desired *indication* of results is achieved through manipulation without necessarily any improvement in the *actual* results. Presumably the higher the stakes involved the higher the propensity for perverse behaviour of both honest and dishonest forms.

Difficulties of measurement

The second of the challenges and the one that is most particular to the public sector – having a focus on outcomes rather than outputs – is less amenable to solution. Private enterprise judges its success by outputs; those outputs all have monetary values; and there is no debate about what ultimately the goal is. Private enterprise is not concerned with what its impact might be – if it were, clearly many widely available commodities and services would cease to be produced. Government produces outputs, but these outputs are only a means to an end, the end of addressing some problem in the economy or society. The ultimate goal is *outcomes* and that presents problems of measurement, attribution and direction. Social indicators may exist or be developed for many outcomes but with varying difficulty, particularly for outcomes with longer time horizons. Schools should produce children with identifiable and testable cognitive skills; but to some degree that is an indicative or intermediate outcome. Schools ultimately should produce citizens who over the longer term prove to be capable economic agents and well-adjusted members of society. The trend toward outcomes-focused benchmarking in the public sector spurs greater research on, and investment in, social indicators.⁵

Even if the outcomes are readily measurable, they may not be so readily influenced through policy; to what factors do we attribute performance? Best efforts may well fall short of impressive outcomes given the often intractable and ‘wicked’ nature of policy challenges. And, finally, unlike in the private sector, there are legitimate differences in views about what outcomes the public sector is seeking in many areas.

Of course, there is much utility in measuring public sector outputs and in measuring output efficiency. Even here there are not-insignificant challenges given the complexity of many public sector outputs. The argument of benchmarking advocates is that the creation of such regimes prompts and promotes the progressive improvement in the data and this seems undeniable. One lesson of the UK experience with a performance monitoring reliance

5. On which, see for example: Tony Atkinson, Bea Cantillon, Eric Marlier and Brian Nolan, *Social Indicators: the EU and social inclusion*. Oxford: Oxford University Press, 2002; OECD, ed., *Statistics, Knowledge and Policy: key indicators to inform decision making*. Paris: Organisation for Economic Co-operation and Development, 2005.

on quantitative indicators, though, is that significant qualitative dimensions may slip through the net with potential for quite misleading conclusions to be drawn.⁶

Concerning Federalism

There are four points to be made about federalism that are of particular relevance to this question. These are: federalism's distinctiveness as a mode of governance; its putative benefits; its operational complexity; and its diversity.

1. The distinctiveness of federalism

First, federalism is not just a formalized system of multi-level governance, but a particular form of constitutionalized power-sharing whereby sovereignty is in some sense and to some degree shared and powers divided between two levels of government, *viz.*, the central government and the governments of the constituent units.⁷ It is predicated on three main tenets. The first is that the two levels have a constitutionally-protected autonomy; neither level can unilaterally alter the status or roles of the other. The second is that constituent units have a meaningful degree of responsibility for local matters. And the third is that for matters that affect all, decisions are made nationally not locally. Taken together, these last two principles are similar to the European Union's *subsidiarity* principle: the rule that tasks should be performed by the lowest level of government that can execute them effectively.

There are at least two corollaries of these defining principles. One is that the member governments of a federation are accountable first and foremost to their own political communities and not to each other or to the wider national community. It is not for the national community to punish or over-rule local communities for 'bad' policy or politics. The other is that relations between the two levels of government in a federation be conducted in accordance with principles of mutual respect.

2. The (putative) benefits of federalism

If that is what federalism is, the second question is what's federalism for? Why do we have federalism and why should we go on trying to maintain it?

Federalism has long been seen as offering specific advantages over a unitary state. Traditionally, the first of these has been seen as being the protection of legitimate difference and the ability to have policy tailored to local needs and preferences. Other advantages now capture greater attention in many federations. Most important of these are the enhancement of policy-learning capacity through the multiplication of policy-making sites — so-called *laboratory federalism* — and the ability of citizens to compare the performance of

6. *Nothing But the Truth?* London: Audit Commission, 2009.

7. See Thomas Hueglin and Alan Fenna, *Comparative Federalism: a systematic inquiry*. Peterborough ON: Broadview Press. 2006.

their government with that of governments in other jurisdictions, or *competitive federalism*. These are, however, theoretical or hypothetical advantages. Whether the constructive potential of laboratory or competitive federalism is realized — or is realized to an extent that compensates adequately for the inevitable disadvantages of divided jurisdiction — is another question.

3. The complexity of federalism

The third point concerns the relationship between theory and practice, and in particular the wide gap that tends to exist between federalism in theory and federalism as it actually exists.

Federations have evolved into highly complex and messy arrangements of political and administrative entanglement which conform only very approximately to ideal-typical models. This is particularly the case for the Anglo federations where a constitutional division of powers designed in the eighteenth or nineteenth centuries has had to adapt to modern conditions.⁸ The consequence is a wide range of policy domains where traditional local responsibility has been subject to central government involvement, direction or influence. Taking Australia as a particularly pronounced example, we find the Commonwealth (central government) exercising its influence in a wide range of policy areas that are constitutionally the domain of the States. In an arrangement that is sometimes called ‘cooperative federalism,’ the States typically retain administrative responsibility for service delivery but are subject to some form of Commonwealth steering.⁹ This is accomplished using a range of mechanisms, but predominant among those are conditional grants.¹⁰ The availability of that instrument has been made possible by a strong tendency toward ‘vertical fiscal imbalance’ (VFI) in the Anglo federations. Canada and the United States have a lower degree of VFI but while the Canadian government now makes relatively little use of conditionality in its grants, in the United States, transfers are predominantly conditional, or ‘categorical’, in nature.¹¹

In most federations, social welfare, education and health care have traditionally been a local responsibility but over time they have become

8. Alan Fenna, “The Malaise of Federalism: comparative reflections on Commonwealth–State Relations,” *Australian Journal of Public Administration* 66 (2007):298–306.

9. See John M. Williams and Clement MacIntyre, “Commonwealth of Australia,” in *Distribution of Powers and Responsibilities in Federal Countries*, edited by A. Majeed, R. L. Watts and D. M. Brown. Montreal & Kingston: McGill–Queen’s University Press, 2006.

10. See Alan Fenna, “Commonwealth Fiscal Power and Australia Federalism,” *University of New South Wales Law Journal* 31 (2008):509–29; Alan Morris, “Commonwealth of Australia,” in *The Practice of Fiscal Federalism: comparative perspectives*, edited by A. Shah. Montreal & Kingston: McGill–Queen’s University Press, 2007.

11. See Robin Boadway, “Canada,” and William Fox, “United States of America,” in *The Practice of Fiscal Federalism: comparative perspectives*, edited by A. Shah. Montreal & Kingston: McGill–Queen’s University Press, 2007.

nationalized to one degree or another and in one form or another. This has happened for a variety of mutually-reinforcing reasons, among them the fact that many now have, or are perceived as having, national dimensions that were absent previously. Traditionally regarded as a matter of almost entirely local import, education has in recent years, for instance, come to be seen as integral to the economic vitality of the nation because of the growing importance of human capital to productivity and innovation.

4. The diversity of federal systems and experiences

The fourth and final point is that actually-existing federations each have their own distinct character. Each is unique in a meaningful sense and generalisation is difficult. Even comparing experiences from the five major Western federations Australia, Canada, Germany, Switzerland and the United States raises a number of difficulties. These difficulties escalate when one includes the European Union (EU), an important federation-in-the-making.

The three Anglo federations share some important commonalities. Most importantly, they are all based around the legislative division of powers with the two levels, at least ostensibly, exercising full powers of policy-making, implementation and administration within their assigned spheres. This sets them apart from the German model that assigns policy-making responsibility in many areas to the central government and responsibility for implementation and administration to the constituent units, the *Länder*.¹² As a corollary of that approach, the German model incorporates a heightened degree of representation— via the *Bundesrat*, or Federal Council — for the constituent units in the legislative process of the central government.¹³

On paper, the similarity between the Australian and American federations is particularly strong given the degree to which the Australian founders followed the American example in their design and drafting.¹⁴ However, even among the Anglo federations differences are significant. Of relevance are the fact that the US uses a presidential ‘separation-of-powers’ form of government while Canada and Australia are parliamentary systems; the much larger number of units in the American system; the distinctiveness of Canada as a federation divided between an English-speaking majority and a French-

12. Jutta Kramer, “Federal Republic of Germany,” in *Constitutional Origins, Structure, and Change in Federal Countries*, edited by J. Kincaid and G. A. Tarr. Montreal & Kingston: McGill–Queen’s University Press, 2005; Hans-Peter Schneider, “The Federal Republic of Germany,” in *Distribution of Powers and Responsibilities in Federal Countries*, edited by A. Majeed, R. L. Watts and D. Brown. Montreal & Kingston: McGill–Queen’s University Press, 2006.

13. Stefan Oeter, “Federal Republic of Germany,” in *Legislative, Executive, and Judicial Governance in Federal Countries*, edited by K. Le Roy and C. Saunders. Montreal & Kingston: McGill–Queen’s University Press, 2006.

14. Cheryl Saunders, “Commonwealth of Australia,” In *Constitutional Origins, Structure, and Change in Federal Countries*, edited by J. Kincaid and G. A. Tarr. Montreal & Kingston: McGill–Queen’s University Press, 2005.

speaking minority centred in one of the main provinces; and the absence of any significant underlying federal difference in Australia. They range from the fiscally quite decentralized Canadian case to the highly fiscally centralized Australian one. Once we include Germany, Switzerland and the EU, diversity increases considerably — and it would increase even further were we to include other, less-conventionally structured, federations such as Spain or Belgium. The upshot is that one has to make highly contextualized comparisons.

What might benchmarking do for federalism?

Federalism and benchmarking have, on the face of it, a number of affinities. Both are about utilizing multiple experiences to identify better ways of doing things; and formalized benchmarking may offer ways to harness the policy learning potential of federalism.

Governments as learning enterprises

If we take the original ‘external’ private sector model of benchmarking where independent firms initiate comparative assessment of their performance as a learning exercise that allows them to incorporate elements of ‘best practice,’ then these affinities between benchmarking and federal systems are immediately apparent. We might imagine a federation where the constituent units act like improvement-seeking enterprises, perpetually gauging their performance against fellow governments and incorporating lessons of experience. In this ideal world, everyone is leveraging themselves up, never reinventing the wheel.

The real world is not quite like that. The contributing reasons are manifold, among which are the following:

- Governments have a suboptimal propensity for experimentation;
- Gauging performance and identifying ‘best practice’ is not always easy;
- Policy objectives are often value-impregnated;
- Governments are under electoral pressure not to engage in open self-assessment;
- Mechanisms for cross-jurisdictional learning may be inadequate;
- In many domains there is a homogenizing central government influence.

Seen in that light, the introduction of benchmarking practices and requirements could supply the necessary stimulus and mechanism for competitive improvement and policy learning. Governments that voluntarily enter into benchmarking agreements — ‘benchmarking clubs’ — can create a framework in which more systematic evaluation, greater experimentation, and enhanced learning occurs. Given political realities, this is likely to focus on aspects of service delivery design rather than policy frameworks.

Limitations on the likelihood of sub-national governments engaging in benchmarking of their own volition suggest at least two possible alternatives. One is that benchmarking is done by an independent, non-governmental institution. The advantage is that independence brings in the potential

for neutral assessment. The advantage of deploying a non-governmental institution is that such an arrangement would not impact on the power dynamics of the federal system. However, it also has weaknesses that are the reverse of the advantages: such agencies are at the mercy of governments from whom they seek information; they have no formal leverage. The other alternative, then, is for the task to be executed by an agent with real authority — the central government.

What role for the central government?

In the classic federal model, at least that of the Anglo federations, there is little legitimate role for the central government in overseeing the activities of the constituent units — including via benchmarking. Sub-national governments are accountable for their performance to their own voters, not to the national government or the national community. However, there are a number of reasons to relax that stricture.

Most importantly, there is the animating and facilitating role that the central government can play in generating an optimal degree of benchmarking between the constituent units in areas where they continue to dominate. Given that both laboratory federalism and competitive federalism are goods that are almost inevitably in undersupply, there is a constructive role for the central government in encouraging experimentation, promoting and coordinating comparative performance measurement, and facilitating learning.¹⁵

There are also the various programs directed and funded to one degree or another by the central government that operate in areas of sub-national jurisdiction. Benchmarking in those contexts represents an alternative mode of coordination that potentially exchanges ‘micro-management’ type controls for a set of incentives that focus on what policy is ultimately all about: outcomes. Potentially, a switch from input and outputs to a focus on outcomes would encourage experimentation and learning in the effort to find more effective and efficient means to ends at the service delivery level.

What do we find?

In practice, we find a wide range of benchmarking experiences in federal countries — although in almost all countries the use of benchmarking is to date rather fragmentary. Some experiences exemplify the federal ethos; some of them represent the continuation of un-federal practices by new means; all of them raise questions about efficacy and the translation of performance measurement into policy learning. In a number of cases it is very early days. Three general approaches can be identified: monitoring by independent agencies; coercive, top-down monitoring; and collegial benchmarking.

15. As argued, for instance, in Shelley K. Metzenbaum, *Strategies for Using State Information: measuring and improving program performance*. IBM Center for the Business of Government, University of Maryland, 2003.

Independent monitoring

In a number of countries, performance monitoring of constituent units has been, or is being done, by non-governmental organisations or institutions. In the United States the Pew Center carries out a periodic 'Grading the States' exercise.¹⁶ Summary assessment is presented in 'report card' or 'league table' style using a twelve point scale with information made publicly available through website presentation. In the most recent iteration, grades ranged from A- at the top to D+ at the bottom. The assessment criteria are fiscal condition, human resources management, infrastructure provision and information technology. In its performance assessment of the States, the Pew Center highlights the use of performance measurement by state governments as, in turn, an important contributor to success.¹⁷ In Germany, a similar assessment is carried out by the private-sector Bertelsmann Foundation, focusing particularly on fiscal performance.¹⁸ In Switzerland, a university-based institute, the Databank on Swiss Cantons and Municipalities, creates performance comparisons on a range of fiscal and governance indicators.¹⁹ It also seeks to make its assessments widely available through website presentation.

Such independent monitoring has evident advantages and disadvantages. By their unobtrusive nature and disinterested focus on strengths and weaknesses across jurisdictions, such exercises are entirely consonant with federalism and contribute a degree of comparative performance assessment that would otherwise be lacking. This should contribute to both laboratory and competitive federalism. But how much impact do such exercises have? Operating to a large extent with freely available data, independent monitoring may end up measuring things not because they are important or revealing but simply because the data exist and are available. Having no ownership of the exercise, governments may also disregard the findings.

Coercive monitoring

At the other extreme are cases where the central government uses internal benchmarking much as a large business enterprise would with its operating units: as a means of driving performance improvement through sanctions. Such exercises may represent the continuance in new forms of traditional

16. See *Grading the States 2008* at: <http://www.pewcenteronthestates.org/gpp_report_card.aspx>.

17. Top ranked Virginia, the report says for instance, "uses state and agency strategic plans to achieve social goals by linking program and managerial performance measures. These measures are tracked online so that the public can easily see the results".

18. See, for example: Norbert Berthold, Norbert, Matthias Kullas, and Andreas Müller, *Die Bundesländer im Standortwettbewerb*. Gütersloh: Bertelsmann Stiftung, 2007; Uwe Wagschal, Ole Wintermann, and Thieß Petersen, *Konsolidierungsstrategien der Bundesländer: Verantwortung für die Zukunft* Gütersloh: Bertelsmann Stiftung, 2009.

19. *Base de données des cantons et des villes suisses* (BADAC), at the Institute of Advanced Studies in Public Administration (IDHEAP) in Lausanne.

centralizing trends in federal systems whereby the national government uses particular constitutional powers or fiscal superiority to achieve a *de facto* alteration in the division of powers and responsibilities or the application of a new mode of governance to an existing division of functions between national and sub-national governments. Neither version has been without its difficulties.

In Switzerland, the implementation of performance benchmarking between the Confederation and the cantons in the area of employment services seems to have been largely unsuccessful as a result of resistance to the sanctioning arrangements.²⁰ In Australia, the Commonwealth has launched the My School website providing online reporting of student results in national testing, on a school-by-school basis, of results despite the fact that schooling falls within State-government jurisdiction.²¹

The most prominent example of coercive benchmarking comes, though, from the United States. The U.S. government's *No Child Left Behind* Act of 2001 (NCLB)²² effected a major shift of control over primary and secondary schooling away from the States who traditionally exercised almost all responsibility in the field.²³ This brought the States explicitly into the ambitious performance management fold that Congress had initiated for federal government agencies with passage of the *Government and Performance Results Act* in 1993. NCLB was unilaterally developed and imposed on the States as an extension of Congress's traditional conditional ('categorical') grant approach to extending its reach to matters within State jurisdiction. Resistance was significant and for reasons pertaining to federalism and to the difficulties of governing by performance measurement, many commentators regard achievements as small.²⁴

Collegial monitoring and benchmarking

In between these two extremes of independent and top-down monitoring is a third model: more collegial-style monitoring carried out on the basis of intergovernmental agreement and cooperation between constituent units

20. Christopher Hilbert, "Implementation of Performance Measurement in Public Employment Services in Switzerland," in *The Evaluation of Active Labour Market Policies: measures, public private partnerships and benchmarking* edited by J. De Koning. Cheltenham: Edward Elgar, 2007.

21. See www.myschool.edu.au.

22. *An Act to Close the Achievement Gap with Accountability, Flexibility, and Choice, so that No Child is Left Behind*.

23. See: Paul Manna, *School's in: federalism and the national education agenda*. Washington, D.C.: Georgetown University Press, 2006; Patrick J. McGuinn, *No Child Left Behind and the Transformation of Federal Educational Policy, 1965–2005*. Lawrence KS: University Press of Kansas, 2006; Kenneth Wong, "Toward Federalizing Education Policy," *Publius* 40 (2009): 226–233.

24. For example: Beryl A. Radin, *Challenging the Performance Movement: accountability, complexity, and democratic values*. Washington DC: Georgetown University Press, 2006; Diane Ravitch, "Time to Kill 'No Child Left Behind'," *Education Week*.

in some federations. The central government is also typically involved — playing, however, a more facilitative role. The audience may be primarily the governments themselves, or it may be the community more broadly. The leading example of this is Australia's *Report on Government Services* (ROGS), now in its fifteenth year of publication. A steering committee representing all governments establishes the performance monitoring framework and oversees publication of the *Report*. An arm's-length research agency of the Commonwealth government, the Productivity Commission, acts as the node of the exercise: serving as secretariat, compiling the data and producing the reports. While ROGS does not cover everything State governments are involved in doing, it does cover an ambitiously wide range of public services making up a substantial part of State government activity.

There is no real equivalent to ROGS in other federations, though one can find similar arrangements operating on a sector-specific basis. In Switzerland, for instance, the Confederation's Office of Regional Development facilitates cantonal performance monitoring in the area of sustainability policy. The central government's role in this instance is limited, it would seem, to promoting data quality and thus utility — in particular to promote the comparability of data generated on a local basis. In Canada, the federal government acts as a node for a similar exercise in the area of health and hospital services via the Canadian Institute for Health Information (CIHI).

The question to be asked of these various instances of collegial benchmarking is how effectively their increasingly sophisticated generating and aggregating of performance data feeds back into policy learning and service delivery improvement in the individual jurisdictions. In other words, to what extent does performance monitoring actually translate into true benchmarking? The focus in all three of the cases noted here is on quantitative indicators and comparative performance measurement; mechanisms for qualitative learning are absent or very much secondary. This brings us to the European Union's *Open Method of Coordination* (OMC).

The EU has developed the OMC as a mode of policy coordination for application in areas where it lacks jurisdiction.²⁵ The OMC is described as a form of 'soft law' in contradistinction to the 'hard law' that the EU exercises via its 'directives.' Prominent among those policy domains where the EU lacks authority are the institutions of the welfare state, education systems and labour market programs — in other words social policy broadly conceived. These were originally seen as incidental to the EU's main objective of promoting economic dynamism through economic integration, but they are now regarded as essential factors in economic performance.

Under the Lisbon Strategy proclaimed in 2000, the EU will pursue improvement in those policy areas by establishing performance measurement and benchmarking frameworks, engaging in evaluation and peer review and

25. See, for example, Luc Tholoniati, "The Career of the Open Method of Coordination: lessons from a 'soft' EU instrument," *West European Politics* 33 (2010):93–117.

encouraging mutual learning. Thus the OMC was designed to have strong and complementary quantitative and qualitative dimensions, to be voluntary, and to promote contextualized learning – that is, learning based on recognition of the different circumstances and different cultural and institutional orders prevailing in different jurisdictions. Considerable emphasis is placed on detailed *peer review* exercises that focus on the qualitative dimension necessary for inter-jurisdictional learning to take place.

Dissatisfaction with the OMC's limited impact led after a few years to the recommendation that it switch to a “naming, shaming and faming” approach in the form of a league tables style of ranking that would more aggressively cajole Member States into adopting best practices.²⁶ That recommendation was rejected and there is little reason to think that it would have been successful. While the OMC epitomizes the federal principle of cooperation, mutual respect, autonomous accountability and improvement through mutual learning, its substantive impact remains much debated.²⁷

Conclusion

Federalism and benchmarking are enjoying a tentative, exploratory, relationship that is partly based in good faith attempts to fulfil some of federalism's potential as a learning-oriented governance arrangement and partly reflective of long-running centralisation dynamics. Both logic and evidence suggest that best-practice oriented ‘external’ benchmarking will be very much an undersupplied good in both the public sector in general and federal systems in particular. Central governments would thus seem to have an important role to play in facilitating cooperation of that nature. Benchmarking may also offer an administratively and substantively superior alternative to more directive modes of centralized policy making in federal systems. However, there are real challenges involved in generating reliable and genuinely indicative data; in relating outputs to outcomes; to identifying and incorporating practice improvements; and in employing sanctions. The nature of those challenges places a premium on cooperative and collaborative processes.

26. Wim Kok, *Facing the Challenge: the Lisbon strategy for growth and employment* (2004).

27. Martin Heidenreich and Jonathan Zeitlin, eds. *Changing European Employment and Welfare Regimes: the influence of the open method of coordination on national reforms*. London: Routledge, 2009; Sandra Kröger, “The Effectiveness of Soft Governance in the Field of European Anti-Poverty Policy: operationalization and empirical evidence,” *Journal of Comparative Policy Analysis* 11 (2009):197–211.

Contributions

The Forum of Federations welcomes the submission of articles highlighting new and interesting trends related to federal governance and providing an insight into the practice of federalism. See the notes for contributors below.

Submissions will undergo peer review to ensure they meet high standards for research quality and objectivity.

Papers in this series will also be posted on the Forum's website.

Notes for contributors

Please keep the following in mind when preparing your paper:

1. Submissions should be no more than 6,000 words in length (excluding citations). The Forum of Federations follows the Chicago Manual of Style (with end-notes for citations). See: http://www.chicagomanualofstyle.org/tools_citationguide.html
2. Authors are encouraged to use section headings to divide the paper into segments that are manageable for readers. When doing so, please:
 - a. Do not number (or letter) the section headings
 - b. Do not use too many section levels
 - c. Section headings should be short, descriptive terms or phrases
3. Authors should also provide an abstract of no more than 150 words. This should provide a brief summary of the paper, and include "keywords" that other scholars might use to search for your paper.
4. Authors should enclose a biographical note of no more than 100 words.

Please contact the General Editor as indicated below.

Rupak Chattopadhyay
Vice-President, Governance Programs

Forum of Federations
700- 325 Dalhousie
Ottawa, Ontario (Canada) K1N 7G2
t. 613-244-3360
f. 613-244-3372
forum@forumfed.org

The Forum does not guarantee the return of unsolicited manuscripts.

Alan Fenna

Alan Fenna has taught political science at Curtin University since 1993 and is currently Professor of Politics in The John Curtin Institute of Public Policy (JCIPP) where he is directing research on Australian federalism. He served as 2009–10 President of the Australian Political Studies Association (APSA), been an elected member of local government, and has worked in the Western Australian government. He is the author of four books, including *Australian Public Policy* (Longman 2004) and, with Thomas Hueglin, *Comparative Federalism: a systematic inquiry* (Broadview, 2006).

ISSN 1922-558X



Forum of Federations
325 Dalhousie Street, Suite 700
Ottawa, Ontario Canada K1N 7G2

forumfed.org