
Introduction

Felix Knüpling¹

Forum of Federations

This volume seeks to shed light on the varying experiences of federal or federal-type systems experimenting with benchmarking as a technique to improve performance and foster learning across their constituent units.² Benchmarking arrangements are being widely adopted across federal systems. All federations face the issue of balancing the interests of the central government³ in key areas of public policy with the desire of constituent units to have autonomy or at least flexibility in terms of how they manage major programs. As part of the ‘new public management’ agenda and the drive towards evidence based policies, benchmarking is emerging as a way of escaping some of the rigidities of traditional conditional grant programs or injecting a new dynamism into federal practices, as well as shifting the focus to outcomes achievement and ‘best practice’.

The contributions covered in this volume have been presented at a conference the Forum of Federations held in cooperation with the Productivity Commission of the Australian Government in October 2010 in Melbourne, bringing together experts and government representatives from both orders of government in Australia as well as from five countries and the European Union. The objective was to share experiences and discuss the applicability of benchmarking exercises in a federal context.

As part of a multi-year research and knowledge-exchange program run by the Forum of Federations on ‘Benchmarking in Federal Systems’, this volume aims to fill a knowledge gap. A lot of research has been conducted on the techniques of benchmarking already, but, as of yet, there has been no systematic comparison drawing out comparative experiences or lessons learnt with a special focus on federal systems.⁴

¹ Felix Knüpling is the Head of Programs and Partnerships at the Forum of Federations.

² The *States, Provinces, Länder* or *Cantons* of which the federation is constituted.

³ This book uses both ‘central’ government as well as ‘federal’ to refer to the government with national responsibilities.

⁴ To my knowledge, there are only two major comparative studies which are related to this topic:
1) The OECD has published a report on *Promoting Performance: Using Indicators to Enhance*

The Forum of Federations began its program in 2008, when it was invited to provide international expertise to the German Government, when Germany was going through a process of constitutional reform. Accordingly, this program is designed both for the academic world as well as for those who are involved in shaping or executing policies, law-makers and civil servants. It intends to present and identify comparative experiences that could inform and stimulate ongoing debates on benchmarking in other federal countries.

This volume examines current practices and identifiable trends in Australia, Canada, Germany, the United Kingdom, the United States, Switzerland and the European Union (EU). This list of countries contains five classical federal countries, but also one unitary state (the UK) and a quasi-federal international organisation (the EU). In the case of the UK, the focus is primarily on national–local relationships, while in the case of the European Union it is on relations between the European Commission and the EU member states.

Special attention is given to the Australian case. Australia has been the first federation to systematically employ benchmarking techniques in its intergovernmental relations and can look back to almost two decades of experiences in this area. For this reason the book devotes one section to the Australian experience and one to the ‘international’ case studies.

Every country or political system is, of course, different in nature and any comparisons or generalisations are always difficult to justify. It is important to keep in mind the contexts of the different cases covered in this volume. What works in one case is not automatically applicable in another. However, this should not preclude us from learning from one another — be it from ‘good’ or from ‘not so good’ experiences. It is with this intention that this book has been produced: to open perspectives through an international comparative exercise and to stimulate thinking outside the conventional boxes.

We invited both academic experts as well as government representatives to contribute to the volume. Thematically, the focus of the cases covered in this volume is on the implications of using benchmarking as an alternative to existing modes of coordination in federal systems, as well as on the political and

the Effectiveness of Sub Central Spending (Working Paper 5, Fiscal Relations Network (<http://www.oecd.org/dataoecd/43/56/40832141.pdf>).

2) The Council of Europe’s Committee of Local and Regional Democracy commissioned a comparative study on ‘Performance Management at Local Level’ in its member countries in 2005 (<https://wcd.coe.int/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=1163686&SecMode=1&DocId=1345000&Usage=2>).

administrative dimensions of implementation — how to design and operate benchmarking arrangements successfully.

A number of questions arise for federal systems.

- How do benchmarking arrangements affect intergovernmental relations and the functioning of the federal system?
- To what extent might benchmarking practices enhance federalism and what form of benchmarking is most conducive to effective federal practice?
- What are the challenges in moving from performance monitoring to active policy learning?
- Does benchmarking actually lead to improved outcomes?

Authors contributing to this volume were asked to address these questions. They do not give, however, an encompassing overview of federal benchmarking exercises in their country. As experts in specific policy fields they present and discuss individual benchmarking examples and how they relate to and/or are affected by federalism. Given that in many benchmarking cases it is still ‘early days’, it will be no surprise that contributors can often advance only tentative answers to some of these questions.

Contributions in this volume

The book starts off with *Alan Fenna’s* overview on ‘Benchmarking in Federal Systems’. *Fenna* bases his observations not only on the contributions of this volume, but also on the research he was mandated to carry out for the Forum of Federations on this subject.⁵ He provides a conceptual clarification of the main variables discussed in this volume, ‘federalism’ and ‘benchmarking’, and how they relate to each other. In characterising the main differences between benchmarking designs in federal systems along a continuum from top down/coercive benchmarking to bottom up/consensus benchmarking his chapter also lays down the analytical framework for the subsequent chapters.

Part I: International contributions

Following this analytical framework, the cases covered in the next two chapters can be categorised as top-down or coercive benchmarking, with the main objective to increase accountability in specific areas of public policy. In chapter 2, *Clive Grace*

⁵ <http://www.forumfed.org/en/global/thematic/benchmarking.php>.

analyses the changing nature of local government benchmarking in the UK, most visibly noticeable in the abolition of the principal benchmarking and performance management regime for local government in England, the Comprehensive Area Assessment, and the main actor of that regime, the Audit Commission, in 2010–11. Despite some scepticism about the added value of local government benchmarking in the UK, Grace argues that performance assessment will remain relevant in the context of UK public policy. He concludes by drawing out the underlying fundamental relationship that connects benchmarking to service improvement, and suggests a rudimentary framework through which policy makers and practitioners should approach the use of benchmarking methods.

Chapter 3 looks at federal benchmarking in the area of education in the US. *Kenneth Wong* describes the 2001 *No Child Left Behind Act* (NCLB), introduced during the second term of US president George W. Bush to improve the quality of primary education in the United States, as the largest federal benchmarking exercise in the US. The author argues that NCLB, marked the beginning of a serious effort toward performance-based federalism. For some, he writes, NCLB has changed the terms of intergovernmental relations in the US to such an extent as to represent a ‘regime change’. However, as NCLB evolves further under the Obama administration, it remains to be seen whether and if so how the performance-based paradigm will be fully institutionalised in the intergovernmental policy system and to what extent it will continue to produce intergovernmental conflicts, Wong concludes.

Gottfried Konzendorf writes about the emerging process of public policy benchmarking in Germany in chapter 4. Compared to all other cases covered in this volume, Germany is unique in the sense that benchmarking of public service delivery of the *Länder* was made a constitutional provision during a recent constitutional amendment in the context of a major overhaul of German federalism. Thus, this chapter is about the attempt to introduce a holistic federal benchmarking exercise driven through a constitutional provision, and not so much about benchmarking in a specific policy area or sector. Compared to other case covered in this volume benchmarking is still in its infancy in Germany. It is also facing resistance, particularly on the side of the *Länder*. Still, *Konzendorf* expects that further benchmarking exercises in various policy fields will be launched soon and that in the medium term these projects will improve the policy coordination of the two orders of government in Germany.

Taking the decentralised nature of Canadian federalism, and its different legal tradition a constitutionally enshrined provision for inter-provincial benchmarking would be unimaginable in Canada, the focus of chapter 5 by *Patricia Baranek*, *Jeremy Veillard* and *John Wright*. Not surprisingly, the Canadian federal government has no overall strategy to introduce benchmarking in the

intergovernmental context. However, there is a very extensive benchmarking exercise going on in health policy, which the authors attribute to intergovernmental events in the 1990s and the move of policy decision-makers to introduce a performance assessment framework to ensure accountability and efficiency in the health care sector. Although initially Canadian constituent units were rather reluctant to embark on this journey, benchmarking in the health sector is expanding through the comparison of performance with peer groups and the learning from better performers, explain the authors. They conclude by providing an outlook on what requirements are needed to improve the current benchmarking framework further.

Like Canada, Switzerland is also a decentralised federation, as *Daniel Wachter* notes in chapter 6 on benchmarking to promote sustainable development. This is a bottom-up benchmarking regime where a federal agency — the Federal Office for Spatial Development (ARE) — plays a facilitative and cooperative but not a directing role. The collaborative and participative nature of the project is integral to its success, explains *Wachter*. He acknowledges that since participating entities do not have to fear punishment because of inferior performance compared to other the regime has been able to expand in the numbers of participants. The collaborative nature is emblematic of the extent to which this regime is more about learning and sharing best practices in a specific policy area than about exerting some sort of control by the federal government over the use of financial transfers. Thus, it can be considered a ‘soft’ benchmarking regime, since there is no ‘hard’ legislation involved.

The same applies to the European Union’s ‘Open Method of Coordination’ (OMC) on social protection and social inclusion, as *Bart Vanhercke* and *Peter Lelie* write in chapter 7. Although the EU is not a federation, it exhibits important federal features and is therefore relevant for our purposes here (see *Fenna*, this volume). The OMC is designed as communication process of performance assessment of the social policies of EU member states according to common indicators and objectives. In its set up it also falls in the category of what *Fenna* calls collegial benchmarking. The chapter on the OMC provides an overview of a range of benchmarking tools and the way a variety of EU and domestic actors are involved in them. The OMC has been applied by European institutions and stakeholders as a mechanism for coordinating domestic policies in a range of issue areas for which the EU has no formal authority, but also for monitoring and supplementing EU legislative instruments. The authors argue that this benchmarking regime has in effect not been as toothless as many critics have argued. It has evolved as an instrument for learning that proves to be of value for decision makers.

Part II: Australian contributions

The following chapter sets the stage for the contributions on the impact of benchmarking regimes within the Australian federal system. *Gary Banks, Alan Fenna* and *Lawrence McDonald* describe the underlying features of Commonwealth–State relations, with a particular view as to how they relate to benchmarking. They touch on constitutional provisions, Australian fiscal federalism, the cooperative nature of Australian federalism and its high degree of centralisation as well as recent reform steps in the area of fiscal federalism that provide the background for new benchmarking regimes. It is the move away from strict controls over federal transfers to the constituent units that provides the main motive for benchmarking regimes with a focus on outcome performance.

In chapter 9, *Gary Banks* and *Lawrence McDonald* describe and analyse benchmarking in the context of the annual *Report on Government Services* (RoGS) as carried out by the Productivity Commission. Similar to our Swiss case study, RoGS is a collaborative and consensual exercise in which the Commonwealth government plays a facilitative role rather than a directive or coercive one. However, it is a much bigger and more comprehensive exercise in performance reporting covering a wide range of services delivered by Australia's constituent units. They amount to almost \$150 billion covering over two-thirds of total government recurrent expenditure. Banks and McDonald describe how RoGS evolved — when they started in 1995 they already embraced a range of different public services (including education, health and justice). The 2011 *Report* contains performance information for 24 'overarching' services, encompassing 12 specific services. Describing the machinery of this vast exercise, including the intergovernmental context, the authors conclude that RoGS can overall be seen as a success in that it contributes to stimulate decision making and that it also reflects the cooperative nature of Australian federalism. However, they also see room for further improvement.

In Chapter 10, *Ben Rimmer* focuses on the COAG reform agenda from the perspective of the Commonwealth government and how it could potentially transform or remodel Australian federalism. COAG — the Council of Australian Governments — is the prime body dealing with dialogue, disputes and funding arrangements. At the core of the COAG reform agenda is the objective, notes *Rimmer*, to improve service delivery through three related provisions: funding linked to the achievement of outcomes and outputs (rather than inputs) in areas of policy collaboration; devolution of decision making and service design to the frontline wherever possible and effective; and competitive tensions between the constituent units ('competitive federalism') and competitive tensions between service providers. The use of benchmarking to measure performance underpins

these three elements, forming a cornerstone of the COAG reform agenda. The success, so far, of this ambitious agenda has been modest, argues *Rimmer*, since the Australian states that have not yet fully delivered on improving public services while at the same time receiving increased funding from the Commonwealth government.

An insider view on the machinery of the COAG Reform Council (CRC) is provided by *Mary Ann O’Loughlin* in Chapter 11. The CRC is the independent benchmarking assessor in Australian federalism’s new performance regime. She explains how more than 90 different payments from the Commonwealth to the States for specific purposes were combined into five new National Specific Purpose Payments. These are underpinned by National Agreements, concluded between the Commonwealth and the States, on key service delivery sectors of schools, skills and workforce development, health care, affordable housing, and disability services. Noting that the whole exercise is still in its early stages, *O’Loughlin* points out that there are some technical challenges CRC is seeking to overcome in cooperation with the States — notably regarding the conceptual adequacy of indicators as well as the availability of adequate data for reporting progress. She also mentions the issue of causality: ‘A comparative analysis does not explain why there are differences between the jurisdictions or why performance has improved or declined over time.’ However, CRC remains committed to pressure governments to take action in response to performance feedback in order to improve service delivery.

In Chapter 12, *Peter Dawkins* and *Sara Glover* discuss the National Agreements on Education in more detail from the perspective of the State of Victoria. They explain how through this agreement benchmarking has become firmly embedded in national policy through the setting of incentives and rewards for States and Territories. This involves payments from the Commonwealth linked with an outcomes framework. Their analysis concludes that there is a very important role for benchmarking in seeking to improve the educational system. However, they also identify a wide range of challenges in undertaking successful benchmarking and — agreeing with *O’Loughlin* — they emphasise the need to take account of the different contexts: when seeking to improve educational outcomes with the assistance of benchmarking, it is important to develop an understanding of what causes improvements in outcomes. Over time *Dawkins* and *Glover* expect that there will be a learning experience and that significant progress can be made to improve educational outcomes in Australia.

A view from Queensland on the value of benchmarking is provided in Chapter 13 by *Sharon Bailey* and *Ken Smith*. They analyse four benchmarking exercises that Queensland is involved in with the Commonwealth. Their main argument is that in employing benchmarking exercises, the issue of context needs to be taken very

seriously — an argument that is echoed also by other authors in this volume. Otherwise, so they claim, the opportunities for policy learning are lost. Describing how in 2010 Queensland had been penalised in the context of one of the Commonwealth–State benchmarking operations, they argue that sanctions might have an adverse impact on the relationship between States and the Commonwealth and hence on performance in the future. According to them, sanctions are a strong, but potentially blunt, tool that requires supplementation. Benchmarking, they contend, has the potential to be misused or to bring about unintended consequences; its impact is dependent on context and the way it is used.

In Chapter 14, *Helen Silver* emphasises the merits of the collaborative nature of intergovernmental relations in Australia. Good processes, so she argues, lead to good outcomes. Given that a decisive feature of federalism in Australia is its vertical fiscal imbalance, she points out that benchmarking exercises need to be shared exercises between the States and the Commonwealth. In this vein, she regards the ongoing institutional reform of COAG itself as important and argues for the need of an intergovernmental agreement to enshrine COAG’s principles and governance. This reform should entail some basic procedural disciplines, such as planning for a small number of regular meetings each year. As part of such an agreement, COAG should be provided with an independent secretariat to coordinate a more focused agenda and allow for the States and Territories to put issues on the table for discussion and action.

Outlook

This volume is attempting to draw out some preliminary comparative conclusions about the relation between ‘federalism’ and ‘benchmarking’. This is an ambitious exercise, and we realise, that this volume can only be a first step in a larger project. There needs to be more empirical research to determine the exact nature of that relationship — and to what extent benchmarking ‘delivers’.

Australia has been at the forefront of experimentation with benchmarking as a tool to improve policy performance. Commonwealth and State governments have, across a wide range of policy sectors, negotiated intergovernmental agreements that identify outcomes, goals, targets or guidelines, and include obligations on the part of participating governments to report to the public on the achievement of these measures. This form of benchmarking is attempting to avoid (or at least alleviate) the hierarchical and prescriptive relationship between Commonwealth and State governments associated with the traditional conditional grant programs that characterised Commonwealth–State fiscal arrangements prior to the ‘new public management’ era.

Also in other federal systems central governments and constituent units have to balance the centripetal and centrifugal impulses for country-wide policy outcomes, on the one hand, and policy outcomes that respect state autonomy or at least promote flexibility, on the other. The contributions by the international authors show that other federations are embarking on a similar route as Australia, without employing or copying its very comprehensive and systematic approach to public service benchmarking.

All chapters in this volume show that benchmarking has become an important aspect of federal governance and we believe that there is much that we can learn about inter-governmental benchmarking by looking across different federal systems. However, the contributions also show that benchmarking comes in many forms and that there are different drivers for benchmarking regimes. In practice, benchmarking is used to describe a wide variety of arrangements, and the objectives vary. While some are about accountability and transparency in intergovernmental relations and public service delivery, some focus on learning and improvement, and sometimes it is a mixture of both.

We also need more research on the impact of politics in general and the broader institutional context (including fiscal arrangements) on benchmarking regimes. Benchmarking can be viewed as an instrument of governance, but the issue of how to set up the governance of benchmarking regimes is also emerging as a key issue from many of the contributions of this volume requiring further investigation. One preliminary conclusion is that the models of a collegial nature, that are not based on hierarchy, targets and reputation effects (naming and shaming), encourage the greatest willingness of constituent units to participate. However, the jury stands out whether it is those arrangements that lead to performance improvement.

Another preliminary conclusion is that all benchmarking systems seem to face considerable challenges in creating and capturing robust and comparable indicator data. Producing good comparative data is only one step in the benchmarking process, and an equally important step is to ensure adequate analysis and interpretation of those data.

Many authors in this volume argue that benchmarking of public services matter because it is critical for governments and communities who need to know whether services are effective, efficient, who is accountable for service delivery, and whether the outcomes of service delivery are in the interests of the citizenry. They also argue that it is an important framework for policy decision-making. However, we would need to more research to find out what works, for what purposes, and with what opportunity costs.

And finally: What is the role of citizens and service users in benchmarking? Although the political rhetoric surrounding benchmarking and the putative benefits of federalism makes some assumptions about improved service delivery, in practice citizens and service users are often only marginal participants in many benchmarking systems. They are rarely involved in discussions about what the indicators should be, what they mean, or what should be done in response to benchmarking results.