

Nuno Ferreira da Cruz, Rui Cunha Marques

Scorecards for sustainable local governments

Article (Accepted version)
(Refereed)

Original citation:

da Cruz, Nuno Ferreira and Marques, Rui Cunha (2014) *Scorecards for sustainable local governments*. *Cities*, 39. pp. 165-170. ISSN 0264-2751

DOI: [10.1016/j.cities.2014.01.001](https://doi.org/10.1016/j.cities.2014.01.001)

© 2014 [Elsevier Ltd.](#)

This version available at: <http://eprints.lse.ac.uk/59670/>

Available in LSE Research Online: November 2014

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Scorecards for Sustainable Local Governments

Abstract:

The current paper addresses the issue of performance measurement and benchmarking in local government. It briefly reviews some relevant international experiences and categorizes the existing assessment tools according to their general features. The authors argue that more than the benchmarking results themselves it is the evaluation process that matters. Furthermore, it is suggested that focusing only on effectiveness (e.g. ensuring the availability of certain public services) or efficiency (e.g. achieving good economic performance), is no longer consistent with the challenges that local decision-makers are now facing. Based on this, a conceptual model for benchmarking the “sustainability” of local governments is proposed. The illustrative municipal scorecard is presented for the city of Lisbon. The results show that using a comprehensive approach and a set of simple and carefully selected quantitative and qualitative indicators may empower citizens to act as “armchair auditors” and encourage local governments to realign their objectives.

Keywords: benchmarking; governance; indicators; performance.

Assessing local government performance

Attempting to compare the performance of governments might be overly ambitious, since they are highly complex structures. Indeed, benchmarking entities with similar and very specific objectives or limited competences presents enough challenge on its own (Berg, 2007). Moreover, many critics argue that benchmarking public sector entities is too expensive and that it produces several unintended effects, such as distorting priorities (Lowe, 2013). Other pragmatic obstacles to the development of sound performance evaluation models are related to the presence of many explanatory factors, and the unobserved linkages between these factors and the perceived/measured outcomes (Fried et al., 2008). All these issues often render performance evaluation or benchmarking unfeasible, or its results meaningless. The fact that it is challenging, however, does not make benchmarking in the public sector any less relevant.

If within the same country or state there are few disparities regarding the rules, governance structures and competences of municipalities, carrying out systematic comparisons of the performance of local governments may be less complex than implementing benchmarking at a state-level (i.e. comparing the

performance of different states or countries). In particular, municipalities from the same country or state usually operate under the same institutional or regulatory environments and are subject to similar financing conditions (Cruz and Marques, in press). Evidently, benchmarking across local governments from different countries is a much more ambitious and complex proposition.

Assessing local government performance is particularly relevant because, while it is bound to differ substantially from one municipality to another, the visibility (e.g. national media attention) and accountability mechanisms are generally weaker at the local level (Martin et al., 2013). In fact, arguments supporting the usefulness of well-designed benchmarking tools at the local or regional level are widespread in the literature. As Ammons (2012, p. vii) argues: “city governments need performance benchmarks if they are serious about the efficient delivery of quality services. And their citizens need municipal benchmarks if they are not!”

The measurement of local government performance is currently a topic of international importance. The UK has been the forerunner regarding the design and wide-ranging implementation of performance measurement models at the municipal level. Highly embedded in the New Public Management (NPM) movement, initial efforts began in the 1980s, particularly after the creation of the Audit Commission in 1982. The Comprehensive Performance Assessment (CPA) introduced in 2002, revised in 2005 and replaced in 2008 by the Comprehensive Area Assessment (CAA) in England, were annual assessments with the objective of naming and shaming poor performers; in the Scottish Best Value Audits (BVAs), introduced in 2003 and revised in 2009, local governments were evaluated once every three years and there were no overall scores (to avoid ranking in league tables); the Wales Programme for Improvement (WPI) was mainly based on self-assessments and the results were not publicly disclosed (Martin et al., 2013). All these assessment systems were mandatory. However, recent developments have been in the direction of voluntary programmes.

Other European countries followed the UK example, each adopting their own benchmarking schemes. What is interesting to note is that, while the outcomes of the “performance measurement regime” in this country have been mixed (and the more “muscular” approach to local government performance measurement is slowly being dismissed), Western European countries are increasingly leaning towards

compulsory large-scale benchmarking projects (possibly as a result of the current fiscal crisis, Kuhlmann and Jäkel, 2013). A constraint to be considered is often the local autonomy principle laid down in the Constitutions of many European countries which may somehow shield local authorities from central government evaluation/intervention. At the Community level, the Common Assessment Framework (CAF) was launched in the year 2000 as the first European quality management instrument specifically tailored for and developed by the public sector itself (EIPA, 2013). The CAF uses nine criteria (five are “enablers” and the remainders denote “results”) divided in 28 sub-criteria to be scored with an ordinal scale. Although this might be a useful diagnostic tool for self-assessment, the CAF is not suited and has not been used to compare and contrast the performance of local governments.

Outside Europe the Australian example is quite relevant. First, the annual Report on Government Services (ROGS), which assembles detailed performance data from the six states, is the result of a collaborative work between the several governments; however, being a tool for government, not the general public, league table type reporting is avoided and no effort is made to establish benchmarks (Grace and Fenna, 2013). Second, Australian states have also established mandatory performance reporting approaches applicable to local governments (e.g. see the case of Tasmania, Local Government Division, 2011). These reports are quite comprehensive and publicly disclose benchmarks for municipalities. In the US and Canada, benchmarking efforts at the municipal level are often voluntary (yet quite common) and differ from state to state (or province). There are, nevertheless, some mandatory initiatives. For instance, the Municipal Management Indicators regime of the Canadian province of Quebec was implemented in 2004 and currently includes 14 mandatory standardized indicators (Charbonneau and Bellavance, 2012).

The initial motivations behind performance assessments of local authorities were mainly rooted in service failures that eroded confidence in professionals to protect the public interest (Grace and Fenna, 2013). Being the closest link to the citizens and usually responsible for the provision of essential services, the failures of local governments are highly harmful for social welfare. The recent reforms towards more decentralisation, emphasis on localism and additional autonomy gave rise to a need for more self-regulation and scrutiny by citizens. Therefore, this study discusses an approach for comparing the performance of local governments in a given country, in terms of their sustainability, where the citizens are the main users. In the authors' opinion, discussing the current approach is of international importance

because local governments around the world face similar threats to their sustainability, namely: “skills shortages, financial stress, limited revenue-raising capacity, asset deterioration and replacement costs, ageing populations, growing community demands, and governance issues” (Local Government Division, 2011, p. 3).

Concepts, approaches and methodologies

In simple terms, benchmarking can be defined as the process of pursuing excellence through the systematic comparison of certain performance measures with the predefined reference levels (Marques and Witte, 2010). Depending on the sector, purpose, focus, methodology and so forth, benchmarking approaches may have different designations, such as: “metric benchmarking” (quantifying the relative performance of units), “process benchmarking” (involving detailed analysis of operating characteristics), and “customer survey benchmarking” (identifying customer perceptions, Berg, 2007). The current chapter provides a general classification of the tools and techniques that have been used to carry out benchmarking in various fields (e.g. based on the method, data type, ownership of the assessment, purpose and target audience).

Having its origins in the private sector, benchmarking acquired its own jargon. By comparing its performance against that of industry leaders in order to learn and improve, a firm would be carrying out external benchmarking. Otherwise, if the management decides to assess the performance of the various departments/staff members by establishing benchmarks and perhaps impose a system of awards and sanctions, the firm would be carrying out internal benchmarking. Regarding local government benchmarking, it is said to be vertically intergovernmental when the central government mandates and/or facilitates performance reporting and horizontally intergovernmental when there is an implicit or explicit comparison between individual municipalities (Grace and Fenna, 2013). Also related to these issues, the taxonomy set forth by Kuhlmann and Jäkel (2013) classifies benchmarking schemes as compulsory hierarchical management (the design of the assessment model, the auditing and the information exchange is carried out under the supervision of the central or state government), vertically co-ordinated management (central or state government cooperates with local authorities to develop the assessment model, gather data, and report and act upon the results), voluntary local self-management (without the intervention from the state, several local authorities decide to measure and compare their performance),

and independent monitoring (NGOs, civil society associations or academics perform benchmarking using perception-based or publicly available data).

Irrespective of the sector, the process of designing performance measurement models can either be technocratic or participatory. Technocratic models are developed by experts or consultants on behalf of top management (or central governments, in our case) and involve a normative or prescriptive approach regarding what should be measured and what should be the targets. Contrary to this top-down approach, a participatory process can also be adopted to involve practitioners, users and other stakeholders (Cruz and Marques, 2013). After the design phase, adherence to the performance evaluation programme can still be voluntary or coercive. Depending on the legal framework, results can be used for naming and shaming (there are no direct consequences, although the publicity is expected to drive change and improvement) or even to enact awards and sanctions (to promote good examples and censor bad practices). Evidently, the underlying objective is always to determine the sector standards and ensure continuous improvement.

Regarding the methods used to measure performance, the array of options available is also quite wide. First of all, the performance descriptors (often referred to as, “metrics” or “indicators”) do not have to be necessarily quantitative. Complex phenomena can be assessed through constructed descriptors using categories that clearly describe and define (qualitatively) the different performance levels. Evidently, direct (or natural) quantitative descriptors are preferable and should be used whenever it is possible to avoid ambiguities. Occasionally, however, data is only available to feed indirect (or proxy) descriptors. The quality and type of data used in the assessment is a determining factor for its success. Benchmarking models based on objective, factual, hard data tend to be less controversial than the ones relying on perception-based data, such as surveys or expert coding (Bouckaert and Van de Walle, 2003). Each descriptor (performance indicator) is a partial measure of performance (since not all relevant variables are taken into account). These performance measures may refer to input (resources required), output (amount of work or services provided), efficiency (relationship between output and input), outcome (results, consequences or impacts), effectiveness (quality, responsiveness and achievement rate) or productivity (combines efficiency and effectiveness or is related to changes in efficiency through time). However, if the intention is to gauge the overall performance, global measures must be employed. Usually these assessments focus on the relative efficiency or the productivity of the units. Both parametric and non-

parametric methodologies may be used to carry out this type of benchmarking. While parametric methodologies (such as Stochastic Frontier Analysis – SFA, Aigner et al., 1977) require an a priori definition of the cost or production function, non-parametric methodologies (such as Data Envelopment Analysis – DEA, Charnes et al., 1978) use the information within the data to estimate the overall scores. These methods are further divided into frontier (e.g. DEA and SFA) and non-frontier techniques (e.g. regression analysis). In frontier techniques all observations are compared with the “best practice” or “efficient” frontier (Fried et al., 2008).

The local government scorecards

Assessing the performance of municipalities as a whole instead of focusing on each public service individually (e.g. water services, waste management, urban transportation and licensing services) is certainly a more ambitious undertaking. However, though dealing with each competence of local governments might be extremely useful to public servants and even to political leaders (to determine where the room for improvement is exactly), to the general public an excessive number of measures or too much detail may cause “information overload” (Cruz and Marques, 2013). One must ensure that the tool fits the users’ needs. Thus, if the objective is to craft an accountability mechanism which relies on citizens acting as “armchair auditors”, it is crucial that the general public is able to understand and use the benchmarking results.

Global measures of performance are attractive because they are not so prone to “blame avoidance” (Charbonneau and Bellavance, 2012) or to the discretionary presentation of results by local governments as partial measures (where only the performance indicators that “look good” are disclosed). Nonetheless, local governments are subject to numerous explanatory or non-controllable factors that affect their performance (either negatively or positively). Since it is virtually impossible to take into account the operational environment (or context) to its full extent, the credibility of these assessments is seriously affected. Considering this limitation, and adding the fact that local authorities are often defensive and unwilling to embrace performance assessment initiatives, the robustness and usefulness of the results must be the primary concerns. Therefore, in the spirit of the balanced scorecard approach developed by Kaplan and Norton (1992), a wide-ranging scorecard of simple and carefully selected quantitative and

qualitative indicators may empower citizens to effectively monitor local governments and encourage them to adopt better practices.

Input and output measures have some value but the information they provide may be limited for citizens. For some crucial aspects, however, this is the only data available and it should therefore be included together with other efficiency, effectiveness and outcome indicators. The outline of the benchmarking approach suggested by the authors is presented in figure 1. The aim is to develop a four-field scorecard for every local government. The assessment perspectives have been tailored to fit the “municipal scorecards”. The top three perspectives correspond to the triple bottom line (TBL) outset (Elkington, 1997). The authors argue, however, that the social, environmental and economic perspectives are necessary, but not sufficient to benchmark local governments. In the last decade, the argument that the performance of all organizations must be seen through these three (often conflicting) lenses has gained traction both in the academic and the practitioner settings (Tanguay et al., 2010). While excelling in the three dimensions is of utmost importance for local governments, a crucial element is being left out: good governance.

Indeed, whereas achieving good governance might not be an end in itself, it is certainly instrumental and necessary to sustain the TBL pillars (Stewart, 2006). Being the closest link to the citizens and usually responsible for the provision of essential services, misconduct and maladministration in local governments are particularly harmful to social welfare. Nowadays, the social, economic and environmental performance is not sufficient to judge the actions of a local authority; it should also be evaluated by its conduct and the way it actually carries out its responsibilities (Hill & Lynn, 2004). Governance relates to the behaviour of institutions, the governing processes and the relations between the state and the citizens and other stakeholders (6, 2004), and good governance demands high levels of public participation in policymaking, transparency, accountability and respect for the rule of law (Callahan, 2007).

The rationale for using these perspectives instead of the well-known model using the four dimensions of the balanced scorecard (Kaplan and Norton, 1992) is linked to the targeted final-user. The focus here is on the use of the information by citizens. The TBL dimensions (now well-entrenched into society) plus the

governance perspective (increasingly regarded as a crucial matter in both developed and developing countries) may help to communicate the attained performance to the general public. Indeed, while the model skilfully discussed in Kloot and Martin (2000) may also be used by interested citizens (if the results are publicly disclosed) it is mainly useful for the internal management of local governments.

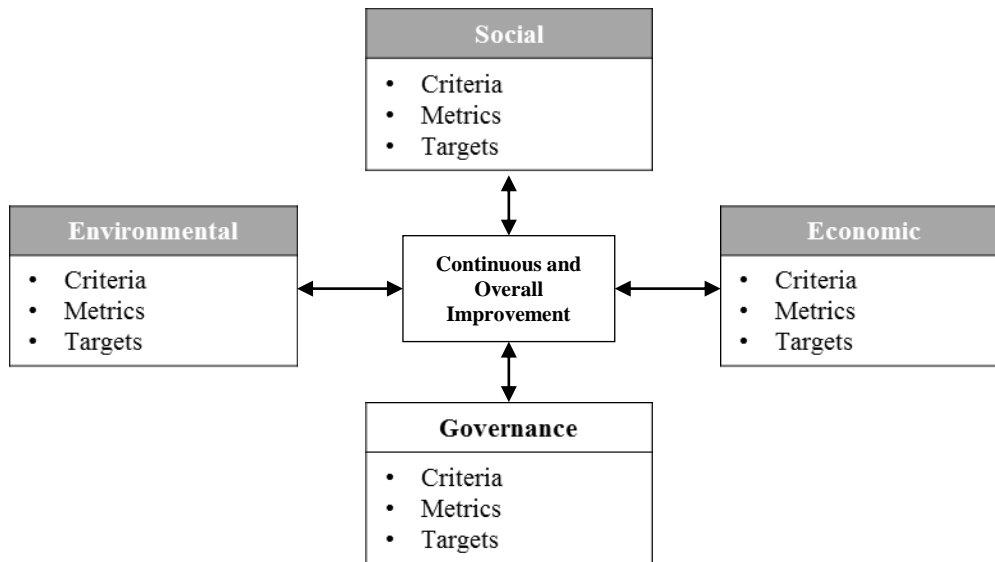


Figure 1 – Conceptual approach for benchmarking local governments

The criteria selected for each of the four dimensions of the municipal scorecard must be in line with the role and competences of local governments. In theory, the same conceptual scorecard could be used to compare municipalities from different countries; however, since the local government structures and functions assigned vary substantially across countries, this would probably render unfair or distorted results (because it is very difficult to select meaningful criteria that are not prone to be influenced by the operational environment). Although the aggregation of the indicators from the four dimensions is possible, it would require the rescaling of the performances attained (to convert impacts in scores) and the definition of performance intervals (two performance references) rather than performance targets (one performance reference). Discretionary aggregation methods (such as equal weights for all criteria or any other arbitrary elicitation) are theoretically incorrect and in general produce meaningless scores (Cruz and Marques 2013). A suitable weighting procedure takes into account the preferences of the legitimate stakeholders and the interactions between criteria (which should be preferentially independent, Cruz and Marques, 2013). Bearing in mind that aggregation is useful to operationalize and measure complex

phenomena it may not be desirable for the specific case of measuring local government performance (as mentioned above, ranking in league tables may be counterproductive).

To truly inform local policies and to be widely used by citizens from all municipalities within a country, the criteria and respective descriptors should be discussed among and selected through the participation of civil society organizations, audit institutions, local government representatives, and other interested stakeholders (Holden, 2006). This group of stakeholders ought to carry out this work (e.g. facilitated by NGO, central government or local government entities) taking into account several practical constraints, namely: the criteria should only capture aspects that are controllable (or influenced) by local governments, be concise and as simple as possible (for a clear perception of what is being measured), and rely on data that is available for all local governments (if the intention is to compare performance between different jurisdictions). The idea is that these criteria reflect the fundamental objectives of citizens (or society in general) so that by scoring in each criterion local governments are actually aligning their own goals with these objectives and consolidating their long-term sustainability.

With regard to the “control” or “influence” requirement mentioned above, the following is worth mentioning. Although currently many local services are organized and delivered by different organizations in different municipalities, from the perspective of the citizen, the actual governance structure of the service (e.g. ownership, regulatory features, etc.) is irrelevant. Indeed, irrespective of the service delivery option (a prerogative of the local executive) the local government is always the “provider” to the eyes of the citizen, even if it is not the “producer”. That is, local governments are always responsible for the quality delivery of essential public services including when the services are produced by a private company. If local governments opt for a model other than the traditional in-house delivery, they ought to design and enforce the necessary mechanisms to ensure the operators commit to the public interest (e.g. during the access to the market phase, through contracts and with continuous monitoring).

Illustration: the Municipality of Lisbon

The current chapter illustrates the application of the conceptual municipal scorecard to the case of Lisbon. The Portuguese capital has 547,733 inhabitants and an area of 84.97 km² (INE, 2012). The selected

criteria are in line with the role and competences of the Portuguese municipalities (in addition to licensing, urban planning and other administrative competences, the main tasks correspond to the delivery of urban infrastructure services such as water, wastewater and urban waste services). However, as mentioned in the previous section, the real operationalization and implementation of the conceptual scorecard would require the criteria, descriptors and targets to be discussed and selected by all interested stakeholders.

The municipal scorecard for Lisbon is presented in Table 1 (data pertains to the period 2010-2012, depending on the indicators). As can easily be seen, the metrics consist of a mix of indicators with special emphasis on results. This example includes input (S2, S3, G1) and output (En2, En5) measures as well as efficiency (Ec1, Ec2), outcome (S1, En1, G3, G4) and effectiveness (Ec3, En3, En4, G2, G5) indicators. Note that the municipality of Lisbon did not (publicly) establish any targets for these indicators and they are merely illustrative. In any case, the established targets must be at least as demanding as current legal standards (when applicable), specified by representatives of the legitimate stakeholders or, for example, correspond to the first or third quartiles considering the performances of all municipalities in a given year (respectively, when the targets correspond to a maximum or a minimum). It is not expected that most municipalities would easily achieve the targets laid down. These levels of performance should correspond to the “best practices” and ought to be systematically revised to foster continuous improvement.

Table 1
Municipal scorecard for Lisbon

Aim	Criteria	ID	Performance descriptors	Target	Performance	Distance to target	
Social	Complaints	S1	Complaints received by the ombudsman in the last 3 years (no. per 10 ⁴ inh.)	6.0 ^b	6.4	-6%	!
	Investment on NESGI ^a	S2	Social, cultural, educational and recreational services (€ per capita)	72	58	24%	!
	Fiscal effort	S3	General local taxes collected (€ per capita)	225 ^b	592	-62%	!!!
Economic	Debt management	Ec1	Debt to total revenue ratio (%)	45 ^b	154	-71%	!!!
	Cost coverage	Ec2	Total revenues to total expenses ratio (-)	1.00	1.05	-4%	✓
	Payment to suppliers	Ec3	Average payment period (days)	60 ^b	95	-37%	!!
Environmental	Water management	En1	Real losses (% of revenue water)	10.0 ^b	8.7	15%	✓
		En2	Water consumption (l/inh./day)	145 ^b	161	-10%	!
	Wastewater management	En3	Wastewater treatment (% of compliant discharges)	97	95	2%	!
	Urban waste management	En4	Waste landfilled (%)	20 ^b	27	-26%	!!
		En5	Waste production (kg/inh./day)	503 ^b	597	-16%	!

Governance	Public participation	G1	Participatory budget (% of total investment)	3.0	3.5	-14%	✓
	Financial accountability	G2	Qualitative/constructed (see Table 2)	Level III	Level V	(-)	!!!
	Credibility	G3	Budget execution (% - weighting each entry by its initial amount)	75	64	17%	!
	Transparency	G4	Documents available online (% of items in Table 3)	100	75	33%	!!
	Procurement procedures	G5	Adoption of competitive tendering (% of procurement expenditure)	80	48	67%	!!!

^a Non-economic services of general interest.

^b Maximum targets (as opposed to the remainder which are minimum targets); in these cases, a positive distance to the target means that the objectives have been surpassed (i.e. negative distances represent underperformance).

A pictorial representation of performance such as the one presented in the last column of Table 1 next to the quantitative information may help presenting and communicating the results. Extreme distances to the targets (as the ones attained in Lisbon for the indicators S3, Ec1 and G5) demand special attention by local decision-makers and/or a clear justification and analysis regarding the contextual factors that might be causing it. For particularly complex criteria the use of constructed descriptors is quite useful. The example provided in Table 2 for the criterion “Financial accountability” is a good illustration of how categorical performance levels (each containing a comprehensive description of what must be respected) may be drafted to operationalize abstract notions.

Table 2
Constructed descriptor for the criterion “Financial accountability”

Performance	Description
Level I	The last annual financial statement includes full and detailed information on revenues, expenditures and financial assets and liabilities. This statement consolidates all the participations of the municipality (corporate and non-corporate). Any contingent liabilities for guarantees or warranties with local PPPs are fully disclosed. The revenues and expenditures of utilities are allocated to each type of service and reflect the costs of the social choices.
Level II	The last annual financial statement includes full and detailed information on revenues, expenditures and financial assets and liabilities. This statement consolidates all the capital participations of the municipality (public and mixed entities). Any contingent liabilities for guarantees or warranties with local PPPs are fully disclosed. The revenues and expenditures of utilities are allocated to each type of service.
Level III	The last annual financial statement includes full and detailed information on revenues, expenditures and financial assets and liabilities. This statement consolidates the majority of the capital participations of the municipality (public and mixed entities). The revenues and expenditures of utilities are allocated to each type of service.
Level IV	The last annual financial statement includes full and detailed information on revenues, expenditures and financial assets and liabilities. All public and mixed municipal companies present their own detailed financial reports.
Level V	None of the above levels is fully respected.

Each descriptor must be carefully devised. The credibility of the benchmarking model depends on it. For instance, the municipality of Lisbon does not display on its website complete information regarding “Procurement” and “Transfer and subsidies” nor does it provide “Sustainability” and “Self-assessment” reports (see Table 3). Would it be the same if, say, a “Sustainability report” was presented but the municipality failed to deliver an “Income statement”? If the answer to this question would be something in the lines of “no, this would be worse for the transparency of the municipality”, then the current

descriptor would not be appropriate (% of items in Table 3 available online) and would have to be revised.

Table 3

Documents that should be available online (for the criterion “Transparency”)

Predictive information	Financial information	Additional information
Activities plan	Balance	Associations
Multi-year investment plan	Consolidated account	Foundations
Budget	Income statement	Local companies and participations
	Unpaid commitments	Procurement
	Budgetary control maps	PPPs
	Budget modifications	Sustainability report
	Management report	Self-assessment performance report
		Transfers and subsidies

Concluding remarks

There are several difficulties connected with performance measurement in the public sector. The obstacles can be of the institutional (mistrust and resistance to reporting “bad news”), pragmatic (lack of usefulness or absence of good data), technical (lack of standards or of a suitable assessment model), and financial types (valuable benchmarking usually demands investing time and resources, Pollanen, 2005). In fact, benchmarking initiatives instigated by the central government and devised to award good performers and sanction “laggard” municipalities have not been consistently successful. Taking this into account, this paper advocates the usefulness of a simple, concise and intuitive performance evaluation model that can be understood and used by citizens. By suggesting a simple, yet wide-ranging approach for local government benchmarking, it intends to provide a good starting point for discussions around performance evaluation at the local level. Although the conceptual model illustrated here prescribes several specific measures, coercive oversight mechanisms are not particularly welcomed by local authorities. However, benchmarking might also offer valuable insights to local governments and inform “cognitive conflicts” in the relationship with central governments (Berg, 2007).

Whereas sustainability is a societal objective and thus it may be the focus of benchmarking at the local level (Tanguay et al., 2010), it is argued that the TBL approach is necessary but not sufficient to assess local government performance. Traditionally, local governments have always placed social concerns as top priorities. And while the environmental agenda is also deeply rooted in the agenda of local decision-makers from Western democracies (mostly due to pressures from NGOs and public opinion) economic and financial concerns also gain visibility in times of crisis. Nevertheless, modern cities need also to be

evaluated in other aspects, including for the conduct of their governments (Yigitcanlar and Lönnqvist, 2013). For this reason, the municipal scorecards also include the “governance” perspective. Although the perspectives of the TBL are regarded as the fundamental ends to achieve excellence, the conceptual approach acknowledges that good governance is a fundamental means and a crucial aspect to foster the confidence of citizens and other stakeholders.

References

- 6, P (2004) *E-governance: Styles of Political Judgment in the Information Age Polity*, Palgrave, New York.
- Aigner, D, Lovell, K, Schmidt, P (1977) Formulation and estimation of stochastic frontier function models. *Journal of Econometrics* 6(1), 21-37.
- Ammons, D (2012) *Municipal Benchmarks: Assessing Local Performance and Establishing Community Standards*, M.E. Sharpe, London.
- Callahan, K (2007) *Elements of Effective Governance: Measurement, Accountability and Participation*, CRC Press, Boca Raton.
- Charnes, A, Cooper, W, Rhodes, E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2(6), 429-444.
- Berg, S (2007) Conflict resolution: benchmarking water utility performance. *Public Administration and Development* 27(1), 1-11.
- Bouckaert, G, Walle, S (2003) Comparing measures of citizen trust and user satisfaction as indicators of ‘good governance’: Difficulties in linking trust and satisfaction indicators. *International Review of Administrative Sciences* 69(3), 329-343.
- Charbonneau, E, Bellavance, F (2012) Blame avoidance in public reporting: evidence from a provincially mandated municipal performance measurement regime. *Public Performance & Management Review* 35(3), 399-421.
- Cruz, NF, Marques, R (2013) New development: the challenges of designing municipal governance indicators. *Public Money & Management* 33(3), 209-212.
- Cruz, NF, Marques, R (in press) Revisiting the determinants of local government performance. *Omega*, in press. Doi: 10.1016/j.omega.2013.09.002
- EIPA (2013) *Improving Public Organisations through Self-Assessment: CAF 2013*, European Institute of Public Administration, Maastricht.
- Elkington, J (1997) *Cannibals with Forks: The Triple Bottom Line of 21st Century Business*, Capstone, Oxford.
- Fried, H, Lovell, K, Schmidt, S (2008) *The Measurement of Productive Efficiency and Productivity Change*, Oxford University Press, New York.
- Grace, C, Fenna, A (2013) Comparing for improvement: recent developments in benchmarking. *Public Money & Management* 33(4), 235-240.

- Hill, CJ, Lynn, LE (2004) Governance and public management: an introduction. *Journal of Policy Analysis and Management*, 23(1), 3-11.
- Holden, M (2006) Urban indicators and the integrative ideals of cities. *Cities* 23(3), 170-183.
- INE (2013) *Censos 2011*, Instituto Nacional de Estatística, Lisboa.
- Kaplan, RS and Norton, DP (1992) The balanced scorecard: measures that drive performance. *Harvard Business Review* 70(1), 71-79.
- Kuhlmann, S, Jäkel, T (2013) Competing, collaborating or controlling? Comparing benchmarking in European local government. *Public Money & Management* 33(4), 269-276.
- Kloot, L, Martin, J (2000) Strategic performance management: a balanced approach to performance management issues in local government. *Management Accounting Research* 11(2), 231-251.
- Local Government Division (2011) *Local Government Sustainability Objectives and Indicators*, Department of Premier and Cabinet, Hobart.
- Lowe, T (2013) New development: the paradox of outcomes: the more we measure the less we understand. *Public Money & Management* 33(3), 213-216.
- Marques, R, Witte, K (2010) Towards a benchmarking paradigm in European water utilities. *Public Money & Management* 30(1), 42-48.
- Martin, S, Downe, J, Grace, C, Nutley, S (2013) New development: all change? Performance assessment regimes in UK local government. *Public Money & Management* 33(4), 277-280.
- Pollanen, R (2005) Performance measurement in municipalities: empirical evidence in Canadian context. *International Journal of Public Sector Management* 18(1), 4-24.
- Stewart, K (2006) Designing good urban governance indicators: The importance of citizen participation and its evaluation in Greater Vancouver. *Cities* 23(3), 196-204.
- Tanguay, G, Rajaonson, J, Lefebvre, JF, Lanoie, P (2010) Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators* 10(2), 407-418.
- Yigitcanlar, T, Lönnqvist, A (2013) Benchmarking knowledge-based urban development performance: results from the international comparison of Helsinki. *Cities* 31, 357-369.