

Unclassified

GOV/PGC(2006)10/ANN3

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

17-Oct-2006

English - Or. English

PUBLIC GOVERNANCE AND TERRITORIAL DEVELOPMENT DIRECTORATE
PUBLIC GOVERNANCE COMMITTEE

GOV/PGC(2006)10/ANN3
Unclassified

**Issues in Outcome Measurement for "Government at a Glance"
OECD GOV Technical Paper 3**

OECD project on Management in Government: Comparative Country Data

by Janos Bertok, Jon Hall, Dirk-Jan Kraan, Jana Malinska, Nick Manning, Erica Matthews

**34th Session of the Public Governance Committee
30-31 October 2006
Château de la Muette, Paris**

This is presented to the Committee for information and as background to the discussion under item 8 of the agenda.

Please send comments to:

Nick MANNING: nick.manning@oecd.org; Dirk-Jan KRAAN: dirk-jan.kraan@oecd.org; Jana MALINSKA: jana.malinska@oecd.org

JT03216113

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

English - Or. English

**OECD project on Management in Government: Comparative Country Data
Issues in Outcome Measurement for "Government at a Glance"
OECD GOV Technical Paper 3**

Janos Bertok, Jon Hall, Dirk-Jan Kraan, Jana Malinska, Nick Manning, Erica Matthews¹

1. This paper has been prepared under the guidance of the Informal Public Sector Outcome Editorial Group, coordinated by Jon Hall, OECD Statistics Directorate. It draws on a study commissioned by the OECD Key Indicators Project reviewing the measures of well-being and progress (Matthews, 2006) for the joint OECD-JRC workshop "Measuring Well-being and Societal Progress" held in Milan on 19-21 June 2006. It has benefited from comments from Zsuzsanna Lonti (Victoria University of Wellington), Alex Michalos (Institute for Social Research and Evaluation, University of Northern British Columbia) and Lee McCormack and colleagues (Treasury Board of Canada).

TABLE OF CONTENTS

GLOSSARY	6
1. Introduction	7
2. What are outcome measures?	8
2.1. Outcome and performance measures	8
2.2. What outcome measures do governments use?	9
2.3. Origins of well-being measures	12
2.4. Rapid growth in “well-being” measures	13
3. How are they used?	15
3.1. Outcome measures can “loosely” connect to planning decisions	18
3.2. Outcome measures can “loosely” connect to accountability and control decisions	18
3.3. Outcome measures can provide a frame or a vision for subsequent policy decisions	19
4. Risks in their use	21
4.1. Demanding scarce political and managerial attention	21
4.2. Gaming	22
5. Filling an apparent gap	23
5.1. Developing comparative measures concerning “trust in government”	27
5.2. Developing comparative measures concerning equity	29
5.3. Developing comparative measures concerning economic and fiscal stability	29
6. The value of internationally comparable data on “executive governance outcomes”	30
6.1. Benchmarking and structured practitioner dialogue	31
6.2. Developing measures of institutional effectiveness	31
6.3. Monitoring change through comparisons over time	31
6.4. How could “Government at a Glance” help?	31
7. Summary of the key propositions	32
7.1. How well-being indicators are selected and used	32
7.2. The risks of using outcome measures indicators	33
7.3. The potential contribution of “Government at a Glance”	33
ANNEX	35
BIBLIOGRAPHY	37
Boxes	
Box 1. The need for national outcome indicators	7
Box 2. An increasing focus on "capital" within well-being indicators	11
Box 3. Apparent criteria for "well-being" indicators	11
Box 4. GDP and well-being measures	13
Box 5. Outcome measures can only be loosely connected to budgetary decisions	17
Box 6. Outcomes as a broad focus for policy thinking	19
Box 7. Enthusiasts note that outcome measurement offers performance dividends	21
Box 8. Governance indicators	24
Box 9. Development Goals for State Services in New Zealand	24
Figures	
Figure 1. Disaggregated public sector production process	7

ACRONYMES

COFOG	The classification of the functions of government (a classification used to identify the socio-economic objectives of current transactions, capital outlays and acquisition of financial assets by general government and its sub-sectors). (OECD Glossary of Statistical Terms, 2004)
Eurostat	Statistical Office of the European Commission
GDP	Gross Domestic Product: Gross national income earned in the process of producing output
GNI	Gross National Income: Measure of the contribution that economic activities provide to the well-being of the residents of a country, calculated by adding in income transfers received by residents from abroad to GDP (and transfers made by residents to people in other countries subtracted)
GOV	OECD Directorate for Public Governance and Territorial Development
GPI Atlantic	Genuine progress index of sustainable development for Nova Scotia
GPRA	Government Performance and Results Act (US)
HDI	The Human Development Index
INAC	Indian and Northern Affairs Canada
ISEW	The Index of Sustainable Welfare
MAP	Measures of Australia's Progress
NDP	Net Domestic Product: the maximum amount of output that can be spent on consumption during a year while maintaining the country's future productive capacity unaltered
NNI	Net National Income: GNI adjusted to reflect the consumption of capital, providing a better measure of the economic resources that are available to individuals in a country for current and future consumption
PISA	OECD Programme for International Student Assessment
SNA	United Nations System of National Accounts

GLOSSARY

Terms	Use in this note	Formal meaning
Final (end) outcome	Outcomes significantly reflect the results of government actions (whether intended or unintended), but other factors are also implicated.	The final result desired from delivering outputs. An output may have more than one end outcome; or several outputs may contribute to a single end outcome. (http://www.ssc.govt.nz/Glossary/) See also OECD (2002)
Input (non-financial)	Units of labour, capital, goods and services sacrificed for the production of services	“Taking the health service as an example, input is defined as the time of medical and non medical staff, the drugs, the electricity and other inputs purchased, and the capital services from the equipment and buildings used.” (Lequiller, 2005, p. 4)
Intermediate outcome	A consequence of the outputs or activities of government which contributes towards the final outcome. Can be more directly attributed to public sector activities than final outcomes. <i>Classified as outputs in "Government at a Glance"</i>	An intermediate outcome is expected to lead to an end outcome, but, in itself, is not the desired result. (http://www.ssc.govt.nz/Glossary/)
Output (non-financial)	Output derived from the direct measurement of output volume and associated quality characteristics.	Measures which arise from “the calculation of a volume indicator of output using appropriately weighted measures of output of the various categories of non-market goods and services produced.” (Lequiller, 2005, p. 4)
Performance	Used non-analytically to convey that achievements matter as well as probity and parsimony in resource use	The term "performance" is used to indicate that there is a standard to which managers and agencies will be held to account – beyond complying with constraints on the consumption of inputs. ² The difficulty in the term is that the standard that is to be achieved can refer to anything at all beyond inputs – whether it is in fact classifiable as processes, outputs, or outcomes.
Public sector process	Structures, procedures and management arrangements with a broad application within the public sector	Cross-cutting managerial and institutional arrangements within the public sector. (Andersen, 2004)

2. For example, “Performance-based management is a systematic approach to performance improvement through an ongoing process of establishing strategic performance objectives; measuring performance; collecting, analyzing, reviewing, and reporting performance data; and using that data to drive performance improvement” (Artley, Ellison *et al.*, 2001, p. 3).

1. Introduction

1. This paper discusses the complex question of how outcome measures are constructed and used within government. It defines outcomes as those events, occurrences, or conditions that are the intended or unintended results of government actions. They are generally of more direct importance to customers or the public than other measures of output or process.

2. There is a complication however as, in distinction to outputs, outcomes cannot usually be simply attributed to government actions or processes – other factors often outside of government’s control can be implicated.

3. The paper reviews the outcome measures that governments seem to use, and analyses the various “suites” of outcome measures such as those that comprise key education, health, environment, economic and other indicators. It suggests some criteria that appear to explain the selection of measures employed. It tracks the rapid growth in “well-being” measures, and suggests some of the principal ways in which outcome measures can be deployed by government, while flagging some key risks. The paper goes on to demonstrate that measures of “executive governance” are absent from many sets of well being measures, although logic suggests they should have a home within them.

4. The key conclusion of the paper is that “Government at a Glance” could make a major contribution by developing some outcome measures concerning “executive governance” – meeting the same apparent criteria as other existing “well-being” measures but reflecting activities of the executive – specifically not the legislature or judiciary.

5. The paper concludes that, for “Government at a Glance”, “executive governance outcomes” might be broadly of three types: public confidence, equity and fiscal/economic stability. Public confidence might encompass issues around trust in government, and associated concerns relating to the predictability and acceptability of government policy. Equity might encompass the measured distribution of services and benefits across diverse populations. Fiscal and economic stability might relate to the track record of government in these spheres.

6. The paper suggests how such measures could assist governments and other analysts, and it explores in some detail how trust measures might be pursued.

Box 1. The need for national outcome indicators

"Key national indicators can help us to better understand which programs, policies, functions, and activities are working and which are not. When seen in the aggregate and as part of a broader portfolio, key national indicators can provide a fuller and fairer view of how well a nation is doing as well as whether and, if so, how its political leaders are planning for the future. Such information can educate policymakers and the public about the appropriateness, affordability, and sustainability of a nation’s current path. Key national indicators can also help elected officials make tough but necessary policy choices including facilitating better targeting of government actions while ensuring long-term fiscal, social and environmental sustainability as well as the intergenerational equity of existing and proposed government policies and programs.

There’s simply no substitute for understanding the big picture – that is, the position and progress of a nation as a whole. The challenge and the opportunity before us is to build sophisticated information resources and key indicator systems that yield vital insights that transcend specific economic sectors, public and private institutions, and national borders.

There are many areas in which the stakes are high and better knowledge is needed. In the case of the United States, these areas include ensuring fiscal sustainability, enhancing homeland security, stimulating economic growth, creating productive and fulfilling jobs, improving education and innovation, delivering quality and affordable health care, strengthening competitiveness, protecting the environment, and promoting quality of life."

Source: Walker (2005, p. 32)

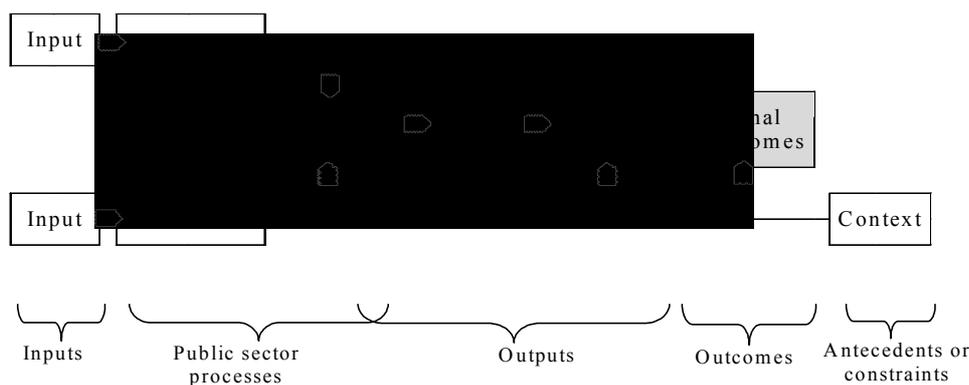
2. What are outcome measures?

2.1. Outcome and performance measures

7. The notion of performance is seen as fundamental to the modern state (Matheson, Weber, *et. al*, 2006; Schick, 2005). However the term itself is somewhat all-embracing – pointing towards various aspects of input, process, output and outcome, and any number of derived ratios between these, and is often used more for rhetorical than practical purposes.³

8. Against this background, “Government at a Glance” will avoid the term performance, and instead will classify measures of public sector activity within six categories of variables.⁴ As Figure 1 indicates, four types of measures characterise the public sector production process (inputs, public sector processes, outputs and outcomes). Antecedents or constraints will provide some key insights into the elements of context that have a significant impact on government efficiency and effectiveness. In addition, revenues (taxation, fees, property income, concessions and transfers, and the incurrence of public debt) will provide insights into the incentives and constraints that governments face in determining how to provide goods and services.

Figure 1. Disaggregated public sector production process



Source: Based on Van Dooren, Sterck, et. al (2006), Hatry (1999), Boyle and Law (2004), Pollitt and Bouckaert (2004) and Algemene Rekenkamer (2006).

3. See Technical Paper 2, “Issues in Output Measurement for ‘Government at a Glance’”, for a more detailed discussion of this point.

4. There is a rather daunting literature on the finer points of classification. Schick highlights this somewhat obsessive concern with arcane questions: “[O]ne of the curious features of this (performance) literature is the endless arguing over what is an output and an outcome; whether a particular measure is an end outcome or an intermediate outcome; whether goals, objectives and targets mean the same things or are different.” (Schick, 2005, p. 9). The issue is clearly not resolvable in any absolute sense – and it is not evident that there is much return on a major discussion of these fine points. Boyle (2005) provides one of the more succinct and pragmatic approaches to these questions.

See Technical Paper 1, “How and Why should Government Activity be Measured in ‘Government at a Glance’?”.

9. Outputs, all products and services delivered by the government, can be readily attributed to public sector activities and processes and their delivery in turn influences the achievement of government outcomes.

10. Outcomes are those events, occurrences, or conditions that are the intended or unintended results of government actions. They are generally of direct importance to customers or the public. For example, in a social policy programme to improve financial management of families, outputs (what the service produces) are the number of counselling sessions or the number of families able to participate in financial management training. However, the desired outcomes include improvements (absolute or relative) in families' financial status, e.g. having more families living within a budget.

11. In distinction to outputs, outcomes cannot usually be simply attributed to government actions or processes - other factors often outside of government's control can be implicated.⁵

12. This classification has the purpose of providing similar units of analysis. Structuring the variables included in "Government at a Glance" within a production process classification does not imply that this idealised flow from inputs to outcomes can always be recognised in practice. There are many situations where the attribution problems between the stages in Figure 1 are so significant that no simple relationship can be identified.

13. To avoid the need to draw fine lines, in "Government at a Glance" outcomes are distinguished from outputs on the rough and ready basis that while outcomes are significantly affected by government actions, unlike outputs, no simple attribution is possible. Contextual factors such as broader social conditions, cultural traditions, and natural disasters that are largely outside of the control of government are often involved and can have a significant bearing on the likelihood that the outcomes will be achieved (though a nation's ability to cope with such disasters is in itself a possible outcome measure of Government).

2.2. *What outcome measures do governments use?*

14. There are many and diverse public sector outcome measures used at the programme, agency or sector level. At the societal or whole of government level, the outcome measures tracked by governments are increasingly described as "well-being" measures. As Summers and Heston (1999) note "[t]he concept [of] well-being... has a number of dimensions [including m]aterial well-being flowing from the availability

5. This definition of outcomes is somewhat at odds with the narrower definition used, for example, in the US Government Performance and Results Act of 1993 (Sec. 2801 – Definitions) "For purposes of this chapter the term... 'outcome measure' refers to an assessment of the results of a program activity compared to its intended purpose".

A further distinction can be drawn between *intermediate* and *final* outcomes. Intermediate outcomes are a result of the public sector activities that are expected to lead to a desired end, but are not ends in themselves. Children attending classes represent an intermediate (and intrinsically valuable) outcome of the activity holding classes. However, the intended final outcome is more likely to be considered the degree to which the children gain competencies in core subjects. Intermediate outcomes can be more directly attributed to public sector activities than final outcomes and, in "Government at a Glance", are included in the discussion of outputs.

Some governments draw a further distinction. The Canadian Government looks at 'immediate', 'intermediate' and 'final outcomes'. Immediate outcomes are those changes that are directly linked to some output (e.g. a reduction in traffic speeding following a road safety campaign). The intermediate and then final outcomes, say a reduction in accidents and a safer road network, respectively, logically follow (Canadian Institute of Chartered Accountants: 2006).

of goods and services, expressed in either current or long-run terms, and [n]on-material well-being, for example, longevity, defined for kinds of welfare conditions that do not necessarily flow simply from the availability of goods and services". Well-being measures are perhaps best exemplified by the datasets in the various national "Suites of Indicators".⁶ Some of the better known are Measures of Australia's Progress (MAP) which has just released its 4th publication; the UK's Quality of Life Counts (first published in 1999); Switzerland's development of Quality of Life indicators; Measuring Ireland's Progress (published in 2003 and 2005); Canada's Performance and the United States' Key National Indicators Initiative which is moving toward becoming a web-based database allowing users to combine variables to measure progress according to their priorities and on a local level (Matthews, 2006).

15. These "suites" include key education, health, environment, economic and other indicators covering, for example, educational attainment, life expectancy at birth, air quality etc. The Annex provides examples.

6 National "Suites of Indicators" are publications containing key measures of national progress together with some discussion of the links between them, but leaving readers to make their own evaluations of whether the indicators together imply that a country is progressing and, if so, at what rate.

The "suite-of-indicators approach" can be contrasted with the "one-number approach", which combines information about progress across a number of fronts (such as health, wealth and the environment) into a single composite indicator, and the "accounting framework approach", which presents social, economic and environmental data in one unified system of accounts, measured in various units (Trewin and Hall, 2005).

Measuring Australia's Progress sets out the approach clearly: "A reader's assessment of whether Australia is, on balance, progressing will depend on the relative importance he or she places on each dimension. For some readers, an improvement in the health and education of Australians might be more important than a decline in our biodiversity. Others might disagree... The suite of indicators presented in this publication suggests progress in some areas of Australian life and regress in others... Overall progress, as explained above, should not be assessed by simply counting the numbers of areas getting better and subtracting those getting worse. Some aspects of progress (especially aspects such as national income and national wealth) are more easily encapsulated in a small number of indicators, than are some social and environmental aspects. And some readers will give greater importance to some progress indicators than others." (Australian Bureau of Statistics, 2006, p. 2).

Box 2. An increasing focus on "capital" within well-being indicators⁷

The Australian Bureau of Statistics is moving towards a "capital" approach as the underpinning framework in its data collection and presentation. A possible approach is under development and will allow sustainability to be assessed as well as progress.

In its budget of 2000 the Government of Canada asked the National Roundtable on the Environment and the Economy to develop recommendations for a small set of environmental sustainable development indicators. A wide range of discussions and consultations led to the adoption of a "capital approach" with the recommendation for a small set of easily understood indicators that would be produced and published annually by Statistics Canada. These indicators cover; air quality, water quality, greenhouse gas emissions, forest cover, extent of wetlands and educational attainment.

In March 2006, Statistics Finland published a statistical review on social capital in Finland. Following largely the OECD definition of social capital, the report contains materials from existing statistical sources. It does not aim to provide any indicators of social capital, simply because there is no common international understanding on the dimensions such indicators should cover. Neither does the report aim at an analysis of the link between social capital and well-being, although such a link has been suggested both in international and national research and in policy debate. The discussion on the role of social capital and well-being in Finland has mainly concerned the assumed positive health effects of social capital.

Statistics Netherlands considers that knowledge is closely linked to the measurement of wellbeing as it affects the future of a society in its core. Sustainability of an economy is highly dependent on the knowledge within a society and so productivity statistics and productivity accounts will be a priority. These productivity accounts will not be limited to labour productivity only but will also include capital.

In the UK, work is likely to commence in the autumn to develop a set of wellbeing measures, with the view to publication in 2007. This is likely to draw on existing statistical work from across government in areas such as social capital and mental health. Another survey will be also commissioned in the autumn which will include questions on life satisfaction.

Source: Matthews (2006)

16. The Annex sets out the well-being measures that are most often included in the "suites". Box 3 suggests some criteria that appear to explain the selection of measures employed.

Box 3. Apparent criteria for "well-being" indicators

A review of the well-being measures used by countries (see Annex) suggests that they are, *de facto*, selected on the criteria that they are:

1. generally desired by citizens;
2. broadly corresponding to the functional areas defined in the Classification of Functions of Government (COFOG);
3. probably significantly affected by government actions;
4. not easily attributable to the outputs from a single sector (although they can be, and often are, readily attributable to the activities of government as a whole);
5. entailing measures that are available at reasonable cost.

17. As will be discussed below, "Government at a Glance" is not proposing to encompass such broad measures of outcome. It will review the possibility of developing some narrower measures which reflect, to some significant degree, activities of the executive arm of government.

7. See OECD Centre for Educational Research and Innovation (2001) for a discussion of social capital in particular.

2.3. *Origins of well-being measures*

18. To some extent, well-being measures have been developed as a reaction to the limitations of Gross Domestic Product (GDP) as a measure of societal progress. The key concern with GDP as a measure is that it fails to identify improvements in, or harm to, social structures or the environment, sustainability of growth, non-market household activities such as unpaid child-care, quality of life issues such as the availability of leisure time, and the distribution of national income.⁸

19. Boarini, Johansson, *et al* (2006) identify three standard ways to translate GDP-based measures into measures at the individual or household level:

1. households' disposable income per capita (a measure of both current consumption spending and of future consumption that could be financed by current saving);
2. final consumption expenditure of households per capita;
3. "actual" household consumption per capita which includes an estimate of government and non-profit institutions services (such as education and health) provided to households for free at point of delivery (or at subsidised prices).

20. All three measures of resources per person are of course lower than GDP per capita, but there is a strong correlation between them and GDP per capita (rank correlations above 0.83 and simple correlations above 0.90). The ranking of countries is also similar for all the indicators (Boarini, *et al*, 2006, p. 15).

21. Since household measures of resource availability are so strongly linked to national measures of GDP, and since these have the intrinsic limitations noted above, then it is probably not surprising that survey-based data on life-satisfaction across OECD are only weakly related to GDP per capita (Boarini, *et al*, 2006).

8. Matthews (2006) points out that higher crime rates, increased pollution, and destruction of natural resources can show up in the GDP as gains. Hurricane Katrina was, therefore, advantageous for the GDP. The Redefining Progress organisation (<http://www.rprogress.org/>) critiques reliance on GDP as a measure of progress on the basis that, in GDP terms, it contributed \$1 billion to the U.S. economy.

The somewhat more refined measures of national income such as GNI (Gross National Income), NNI (Net National Income), or NDP (Net Domestic Product) may be conceptually preferable as measures of well-being, but in practice offer few advantages as they are so closely correlated with GDP that they lead to a very similar ranking of countries and developments over time.

Box 4. GDP and well-being measures

Boarini *et al.* (2006) undertake an analysis of the correlations between levels of 16 indicators of social outcomes and GDP per capita, for each of the four main categories of the OECD social indicators published in different issues of *Society at a Glance*, both in rankings and in annual changes. They find that social indicators covering a broad range of components of well-being have significant cross-country correlations with GDP per capita in several cases, but there are insignificant correlations between changes in GDP per capita and changes in various social outcomes.

"With respect to self-sufficiency, there is in general a significant correlation across OECD countries between levels of GDP per capita and of employment rates (as well as for other self-sufficiency indicators... relating to the labour market). The correlation is not significant, however, for measures of how employment opportunities are shared within society (*i.e.* the extent of joblessness at the household level)..."

With respect to equity, the Gini coefficient for the distribution of household disposable income tends to be higher in countries with lower per capita GDP, but the correlation is insignificant when excluding Mexico and Turkey. Similarly, since the mid-1980s, income inequality declined the most in countries that have recorded stronger growth in GDP per capita, but this mainly reflects the experience of Ireland. OECD countries with lower GDP per capita also record higher relative poverty, both among the entire population and among children, but patterns are less consistent when excluding Luxembourg and when looking at changes. A measure of gender equity in the paid labour market (the differences in median full-time earnings between women and men) is not correlated to GDP per capita...

A number of indicators of health outcomes are significantly correlated with per capita GDP: These include life-expectancy at birth, health-adjusted life-expectancy (a measure that combines information on both mortality and morbidity) and mortality risks, both among the entire population (*i.e.* estimates of the potential number of years lost) and among children. Differences in country performance in health outcomes are generally much smaller than those in GDP per capita but remain large: for example, infant mortality rates differ by a factor of around two between countries with similar per capita GDP. Correlations between changes in GDP per capita and in health conditions are generally insignificant.

Indicators of social cohesion can refer to both its positive manifestations or to some of the pathologies that are informative about lack of cohesion, and which do have resonance as objectives of social policy. On the positive side, social relationships and participation in community life are important determinants of the extent to which individuals feel part of the broader community where they live. Research on "social capital" has stressed the importance of social relationships for subjective and community well-being, and of well-functioning democratic institutions for economic growth. While many of these findings remain controversial, the measure of participation in community life... – the share of people reporting that they volunteer in community groups – is higher in OECD countries with higher GDP per capita. On the negative side, indicators of crime victimisation, prisoners and suicides – as well as of divorces, drug use and road accidents – are not significantly correlated with GDP per capita. In terms of changes over time, OECD countries with higher growth in GDP per capita recorded statistically significant declines in the share of persons who had been victim of crime offences, although data refers to only 20 OECD countries."

Source: Boarini *et al.* (2006, pp. 28-29).

2.4. Rapid growth in "well-being" measures

22. Matthews (2006) notes that the early 1970s saw the "social indicators movement" create considerable academic and social policy attention concerning the construction and use of key measures of societal progress (Lucas, 1985). Subsequent developments included:

- The creation, in 1970, of the "OECD Programme of Work and Social Indicators" as a result of concerns about the limitations of economic measures of progress. OECD (1982) provided 33 indicators together with accompanying statistical specifications and guidelines for data collection, for member countries to measure their social progress. However, amongst some methodological uncertainty and possibly changing political interest, the programme was phased out between 1979 and 1981.
- The launch, in 2002, of the OECD product *Society at a Glance*. This publication is a bi-annual compendium of social indicators intended to shed light on two issues: "how far

OECD countries have progressed in terms of social development; and to what extent this progress has been the result of deliberate policy actions, either by government or institutions” (OECD, 2005b). Indicators are grouped into five main categories: background indicators to provide the overall context in which social policy operates, and four categories that reflect the main objectives of social policy – self-sufficiency, equity, health, and social cohesion.

23. In 1983, the UN General Assembly created the UN World Commission on Environment and Development (the Brundtland Commission). This stimulated a greater emphasis on environmental statistics: evidence was needed more than ever to track the progress of the protocols and assess changes in key aspects of the environment. Key milestones have included:

- Progress in environmental reporting, accomplished through the measurement of greenhouse emissions. Internationally comparable measurements are now provided by almost all the major supranational organisations dealing with the environment and many national statistical agencies produce their own measures.
- The publication of the handbook “Satellite System for Integrated Environmental and Economic Accounts” in 2003. This was an extension of the 1993 version of the System of National Accounts, and was jointly published by the UN, OECD, IMF, EC and the World Bank. It gives an overview of various ways to put into practice the definition of sustainable development proposed by the Brundtland Commission.
- Developments in Material Flows Accounting which track the amounts of materials – as classes or individual substances – that enter the economy, accumulate in capital stock such as housing or automobiles, or exit to the environment as waste. Material flows indicators have been adopted in a growing number of countries and are now regularly published by the EU and some member countries.
- The Global Footprint Network’s measure of Ecological Footprint which indicates how much land and water area a human population requires to produce the resources it consumes and to absorb its waste.

24. Other developments that could contribute to more comprehensive measures of well-being and progress have included:

- United Nations (1975) details a technical statistical system which it claimed, would have developed key datasets that measure progress at the individual and societal levels, and offered a framework for organising them to provide an information system for analysis and policy making. However, this work has not been implemented.
- Development of satellite accounts that can be used to augment information in the national accounts to provide a more complete picture of economic activity in key non-market areas.
- Development of social accounting matrices also based upon the national accounts which link the mainly macro-statistics of National Accounts with the mainly micro-statistics of labour markets and households. One of the most advanced uses of this tool is the Dutch System of Economic and Social Accounting Matrices and Extensions which links economic, social and environmental data in one unified system of accounts measured in various units.
- National “Suites of Indicators” in which key aspects of progress are published with some discussion of the links between them, generally leaving readers to make their own

evaluations of whether the indicators together imply that a country is progressing and, if so, at what rate. Measures of Australia's Progress (MAP) which has just released its 4th publication; the UK's Quality of Life Counts (first published in 1999); Switzerland's development of Quality of Life indicators; Measuring Ireland's Progress (published in 2003 and 2005); Canada's Performance and the United States' Key National Indicators Initiative which is moving toward becoming a web-based database allowing users to combine variables to measure progress according to their priorities and on a local level (Matthews, 2006).⁹

- On a similar "suite" approach, the *OECD Factbook* provides a global overview of world economic, social and environmental trends, bringing together 100 indicators for evaluating the relative position of any OECD member country concerning: population and migration, macroeconomic trends, economic globalisation, prices, labour market, science and technology, environment, education, public policies, quality of life and globalisation.
- The compilation of "sustainable development indicators for the European Union" from Eurostat sets out 110 indicators addressing: climate change and clean energy; threats to public health; management of natural resources; transport and land use; poverty and social exclusion; and the economic and social implications of an ageing society (Eurostat, 2005).

25. Various single number composite measures have been developed:

- The Index of Sustainable Welfare (ISEW) is a measure of economic welfare and sustainable development (Cobb and Daly, 1989) and has been calculated for nine countries (Australia, Austria, Chile, Germany, Italy, the Netherlands, Sweden, the United States and the United Kingdom).
- The Genuine Progress Indicator is a modification of the ISEW, eliminating expenditure on government and investment and adjusting for inequality, adding or subtracting categories of spending based on whether they are considered to enhance or detract from the nation's well-being.
- The Human Development Index (HDI) was proposed in 1990 by the United Nations Development Programme (UNDP) as a composite indicator to assess the three main dimensions of development: longevity, knowledge and the standard of living. Longevity is measured by life expectancy at birth; knowledge is measured by adult literacy rate combined with gross enrolment ratio for primary, secondary, and tertiary schools; and standard of living is measured by GDP per capita in purchasing power parity US dollars. HDI is the unweighted average of the three (UNDP, 2005).
- The Index of Economic Well-being, developed by the Centre for the Study of Living Standards in 1998, comprises four components of economic well-being: consumption flows, wealth stocks, income distribution and economic security.

3. How are they used?

26. Technical Paper 2 in this series highlights that output measures can contribute to a planning discussion¹⁰ or can be employed in decisions concerned with accountability and control. Drawing on the

9. See: <http://www.keyindicators.org/>

10. Planning can be around the nature of the outputs to be provided, the processes to be followed, or around how capacity is to be built.

frameworks proposed by Pollitt (2001) and Talbot (2003), outcome measures can also contribute to these two types of decisions, although to a somewhat lesser extent, but can also contribute towards a broader set of “framing” decisions.

27. Thus in summary, as discussed below, there are three principal ways in which outcome measures can be deployed by government:

1. As a contribution to planning decisions
2. As a contribution to accountability and control decisions:
 - providing a basis for various discussions concerning results achieved;
 - providing a basis for various discussions about the quality of government outputs.
3. As a frame or a vision for subsequent policy decisions:
 - to focus policy thinking;
 - to stimulate a political or public debate;
 - to develop a common agenda within government when diverse sectors and agencies within government are involved;
 - to inspire additional effort by appealing to a broader good.

28. Technical Paper 2 noted that the distinction between the use of output measures for planning and for control and for accountability arises because of the different nature of the incentives that are at stake (Van Dooren, 2006). Accountability emphasises sticks while planning and service improvement suggest carrots.

29. Technical Paper 2 also makes a distinction between two different ways in which measures can be connected to decisions:

- TIGHT: measurement leads to the decision in a direct way. Decisions are driven mainly by the measurement. Other sources of information play a negligible role.
- LOOSE: measurement is one source of information to be incorporated with others. Other sources of information are used to interpret the measurement data and decisions are informed by the measurement, but also by other sources of information such as experience, qualitative information etc.

30. It observed that in both areas output measures are usually only loosely connected to decisions.

31. Box 3 suggested some criteria that appear to explain the selection of outcome measures employed by countries. One criterion seems to be that such measures are not easily attributable to the outputs from a single sector (although they can be, and often are, readily attributable to the activities of government as a whole). This is not particularly surprising as outcome measures reflect broad goals concerning societal progress, and movement towards such goals is likely to be dependent on activities in many sectors. One consequence of this is that it is difficult to connect such measures tightly to any key decisions – either for

planning or for accountability – as there are always several sectoral outputs to be considered, and each of these is likely to be connected to other well-being goals.¹¹

32. In *planning* public sector activities, even if the attribution problem were solved then, as Technical Paper 2 highlights in the case of outputs, there are then significant conceptual problems (does underperformance suggest a higher or lower allocation?) and equally significant gaming problems. Box 5 provides a further discussion of this point in relation to budgetary allocations, although others argue that a robust system of programme evaluation, which uses a sound logical model to link inputs with outputs and outcomes, can provide the basis for results-based management.

Box 5. Outcome measures can only be loosely connected to budgetary decisions

Pollitt (2005) notes that "[i]n a number of countries, governments and experts have recognised the need to move beyond measures of outputs (usually efficiency measures) to measures of outcomes (effectiveness), e.g. East (1997) ... The difficulty arises, however, if there is an attempt directly to link budgetary allocations to effectiveness measures. While this may sound just common sense, in fact it is fraught with problems. The primary obstacle to the integration of performance measurement and budgeting is that the required outcome measures are difficult to construct for public sector programmes... The reasons for this are several. To begin with, for many programmes, outcomes change over a much longer time cycle than the budgetary year. So this year's change of outcomes probably does not reflect the efforts of the current managers at all. Second, it is also frequently the case that outcomes are only partially determined by government programmes – that there are other determining variables which are beyond the control of the managers – and that linking resources to outcomes is, therefore, to greater or lesser degree unfair (Pollitt, 1997). Thus there is an "attribution of outcomes" issue.

Finally, it is necessary to recognise that, politically, for many programmes, failure to achieve outcomes does not mean that resources should be withdrawn and the programmes abandoned. The original political objectives (reducing poverty, reducing crime, creating jobs) will remain as important as ever. There may even be a case for allocating more resources to the task, whilst modifying the programmes in the hope of achieving increased effectiveness. In short, automatic or formulaic links between measures of effectiveness and budgetary allocations will rarely be either technically or politically acceptable.

Source: Pollitt, 1997, pp. 15-16.

33. In *accountability and control* decisions, if the outcomes achieved are not as targeted, then this can be attributed to other factors outside of the immediate control of government, such as lifestyle changes that increase life expectancy or technological changes that increase the effectiveness of particular types of skilled labour.

34. Given that the achievement of outcome targets can not be attributed to single sector-specific government actions, it is not surprising that recent OECD surveys indicate that outcome measures are used considerably less than output measures in the budget process. When OECD senior budget official representatives were asked in 2005 about the types of performance measures that they have developed in relation to the budget process, a half of the respondents said that they tracked data that combined output and outcomes. Over a third indicated that they (also) tracked the unit cost of outputs as a performance measure. About 10% said that they collected output data only (Currstine, 2005). Only around 3% indicated that they used outcome measures alone. The responses were not mutually exclusive.

35. This might be changing in some settings however. Comptroller and Auditor General (2001) demonstrates the changing balance between these categories of performance measures in the UK. Public Service Agreement targets during the period 1999-2002 were categorised as: inputs – 7%; process – 51%; outputs – 27%; outcomes – 15%. During the period 2001-04, they were categorised as: inputs – 5%;

11. This is despite the strong theoretical arguments in favour of using outcome information rather than relying solely on outputs (O'Mahony and Stevens, 2004). Outcomes are, in principal more useful signals for reviewing the effectiveness of public actions because of the lack of market valuations (prices) for outputs, exacerbated by incomplete information.

process – 14%; outputs – 13%; outcomes – 68%.¹² Arguably, there is a tradeoff between the attention given to output and outcome measures – and so the increased focus on outcomes has a price.

3.1. Outcome measures can “loosely” connect to planning decisions

36. Outcome measures can be used as the starting point for the development of an intervention strategy through the use of a “Logic Model” or similar framework (Hatry, 1999; Logic Model Development Guide, 2004).¹³ In essence, such models start from the problem to be solved or the outcome to be achieved, and then work logically towards the required public policy intervention. One version of this planning use of outcome measures highlighted in the literature emphasises a “bottom-up” perspective on evaluation and learning, seeking a more experimental approach in which diverse approaches are tried and tested as various attempts to impact on the outcome (Davies, Nutley *et al.*, 2000).

37. Internationally, the use of the Millennium Development Goals appears to be an example of the use of outcome measures as broad contribution towards a more detailed set of national level planning decisions (Cheung, 2005, p. 376).

3.2. Outcome measures can “loosely” connect to accountability and control decisions

38. If the purpose of public policy action was to contribute towards the achievement of specified outcome targets, then it is reasonable to ask whether those outcomes were subsequently achieved. However, assessing performance against outcome targets can usually be done only generally. In considering, for example, the UK Foreign and Commonwealth Office outcome targets which included “a reduction in the number of people whose lives are affected by violent conflict and a reduction in potential sources of future conflict, where the UK can make a significant contribution” (“Public Service Agreements 2001-2004: Chapter 9: Foreign and Commonwealth Office”, 2001) then the problem of attribution becomes clear.

39. One aspect of this use of outcome measures for accountability and control concerns their use in determining the quality of outputs. Eurostat (2001, paragraph 3.1.2.2) notes that the most appropriate way of adjusting output measures for quality is to investigate changes in outcome indicators. For example, if the level of crime decreases, this could be due (probably in part) to improved effectiveness of the police. If the number of graduates of universities increases, while the total number of students does not change, the

12. Note that time periods overlap.

This movement in the UK is somewhat counter to Graham Scott’s observation that: “The experience of GPRA (Government Performance and Results Act) in the United States shows, in my view, that the attempt to commit an entire government to outcome-based performance accountability is problematical. The practicality of implementation forces the use of proxies for outcomes and measures of government activity that result in a situation that is very close to an output-oriented system that is augmented by measures and indicators of performance. A decade of experience across countries at the leading edge of public management leads me to conclude that outputs are superior to outcomes as the tool for linking the management plans of public organisations to the processes of budgeting and financial accountability, both between the executive and the parliament, and within the executive.” (Scott, 2001, pp. 200-201).

13. In New Zealand such analyses are called the “Intervention Logic”. “Intervention logic starts with a clear definition of an outcome and uses logic and evidence to link outcome goals to departmental outputs. A well-developed intervention logic helps justify the choice of outputs and to improve outcomes on the basis of evidence of effectiveness. When evaluative information can be produced on the outputs, coverage, near-term results and impacts, intervention logic can be used to prioritise outputs so as to maximise departmental outcomes, and confirm outputs delivered by the department work as planned.” (Guidance on Outcomes Focused Management – Building Block 3: Intervention Logic (Version 2.1, July 2003), 2003, p. 1. Gregory (2004) offers a critique of the rationalist claims of Intervention Logic approaches.

universities seem to be doing a better job in educating people. In such cases it will always be necessary to investigate carefully whether there are any other reasons for such changes, for example if universities have lowered their standards. If no changes of outcomes can be observed, it is difficult to argue that a particular product or service has improved in quality. A problem with this approach is the possible time lag between a change in the quality of the output and a change in a particular outcome indicator.

3.3. *Outcome measures can provide a frame or a vision for subsequent policy decisions*

40. Outcome measures can be used as inspiration or, more cynically, rationalisation for a broad set of policy initiatives. This focus on outcomes arguably creates “a ‘mental model’ for awareness, understanding, and thinking prior to choosing” (Riche, 2003, p. 7).

41. By providing a target on the horizon, outcome measures are intended to achieve two key results. First, they are intended to focus policy thinking – providing a framework or an orientation within which other planning and accountability decisions will be made. Second, it is intended to inspire extended organisational and individual effort.

Box 6. Outcomes as a broad focus for policy thinking

Canadian Government departments use a broad framework of strategic outcomes to measure their performance. The following is one example.

Analysis of Performance by Strategic Outcome: Good governance and effective institutions for First Nations, Inuit and Northerners, built on co-operative relationships

Indian and Northern Affairs Canada (INAC) and Canadian Polar Commission

This Strategic Outcome encompasses helping First Nations and Inuit communities to develop effective governance and institutions that support achieving an improved quality of life. First Nations governments and Inuit communities are increasingly responsible for their own affairs as evidenced by devolution, self-government agreements, and new intergovernmental and treaty relationships. These developments support First Nation and Inuit communities' efforts to develop clear accountabilities to citizens, and to help improve community social and economic conditions. INAC's continued support of good governance and effective institutions assists First Nation and Inuit communities to benefit from economic development. At the same time, federal, provincial and territorial governments are building foundations for co-operative relationships with First Nation and Inuit communities.

Progress on the Aboriginal agenda and on northern institution-building contributes to the beginning of a social and cultural revitalization in northern communities and to strengthened partnerships essential to current and future economic development. Further attention to these priorities and to supporting the capacity of territorial governments to deliver programs and services within their jurisdiction is essential. Strong northern governance also has an important role in addressing circumpolar issues.

Source: http://www.tbs-sct.gc.ca/rma/dpr1/04-05/INAC-AINC/INAC-AINCd4503_e.asp

42. The first use has two aspects. Outcome measures such as the European Structural Indicators create a mental “frame” – “pre-digested reasoning that creates an interpretation for facts” (De Biase, 2005, p. 499). Such mental frames can lead to voluntary policy alignment. “The use of Structural Indicators is an example of this new way to design common policies at the European level and to promote convergence among member states through, among others, a common definition of objectives and a common set of indicators to monitor the progress and the convergence” (Munoz, 2005, p. 386).

43. By creating such “frames”, outcomes are particularly useful for public and political policy debate. Ultimately they reflect the issues that engage popular concern, and sidestep the more technical questions about processes and outputs. Politicians think and work in terms of outcomes and the purpose of government policies is often phrased in terms of outcomes.

44. This use of outcome measures to focus a debate also has particular significance in the case of complex outcome objectives which involve contributions of diverse sectors and agencies within

government (Kristensen, Groszyk *et al.*, 2001). In those circumstances, the outcome measure is intended to act as a high-level disciplining device to encourage a collaborative discussion on the relevant contributions of each actor involved. Lonti and Gregory (2006) describe the effect of the New Zealand Public Finance Amendment Act 2004 which came into force in January 2005. By allowing more flexibility in the use of resources, within a tightened requirement that appropriations specify the outcomes that government actions are to achieve or contribute towards, the act is intended to counter the “silo effect”. Parliamentary appropriations do not now have to be confined to one output class and the legislation allows a department effectively to transfer its appropriation for outputs to another department.

45. The Government of Canada “whole of government framework” has a similar purpose. It is structured around three key policy areas (sustainable economy, social foundations and Canada’s place in the world). Each policy area is subdivided into “Government of Canada Outcomes”, which are the “long-term and enduring benefits to Canadians that federal departments, agencies and Crown corporations are working to achieve”.¹⁴ Each outcome is then associated with departmental and agency level outcomes and Crown corporation mandates (see Box 6 above). Bellefeuille-Prégent and Wilson (2005) note that the Canadian Treasury Board Secretariat is considering the development of a whole-of-government planning report that could make the link between societal indicator reporting and government-wide planning more explicit. Similarly, the state level suite of indicators prepared by the State of Victoria, Australia – Growing Victoria Together – had as one of its primary aims to encourage government departments to work better together when tackling crosscutting issues of concern (Hall, Carswell *et al.*, 2005).

46. Outcome indicators have a role not just in focusing political debate, but in broadening the involvement of the wider public in policy-making, though this use goes beyond the ambit of Government at a Glance. Those who develop outcome indicators for this purpose argue that advancements in Information and Communication Technologies (ICT) have changed the way in which markets and societies work, with the Internet in particular making information more accessible to citizens than before. They suggest that the ideal of the “fully informed decision maker” could be a reality. But if that reality is to be achieved, the information must become knowledge. Sets of outcome indicators are seen as one way to provide a wide audience with a digestible set of information with which they can understand the performance of their country, or government. This information can encourage them to better influence policy-makers and help them make more informed decisions when holding policy-makers accountable. (Giovannini and Uyshal, 2006).

47. The second framing purpose of outcome indicators is to direct managerial and individual attention towards a set of higher-level goals or purposes that few would disagree with and that might inspire additional effort See Box 7 and Hatry (2004); Hatry, Morley *et al.* (2003); Perrin (2006, p. 24-25).

14. See: http://www.tbs-sct.gc.ca/rma/krc/so-rs_e.asp.

Box 7. Enthusiasts note that outcome measurement offers performance dividends

"This (outcome measurement) dividend doesn't take years to occur. It often starts appearing early in the process of setting up an outcome measurement system. Just the process of focusing on outcomes-on why the program is doing what its doing and how it thinks participants will be better off-gives program managers and staff a clearer picture of the purpose of their efforts. That clarification alone frequently leads to more focused and productive service delivery... It can, for example, help programs:

1. Recruit and retain talented staff.
2. Enlist and motivate able volunteers.
3. Attract new participants.
4. Engage collaborators.
5. Garner support for innovative efforts.
6. Win designation as a model or demonstration site.
7. Retain or increase funding.
8. Gain favourable public recognition."

Source: Outcome Measurement Resource Network, <http://national.unitedway.org/outcomes/resources/What/why.cfm>

48. Tallis (2005) considers that this framing purpose is where outcome measures can contribute most readily. He notes that, in principle, outcome measures can contribute towards:

- Policy focus – drawing attention to aspects of society that merit attention or intervention.
- Policy design – developing a detailed strategy for a social intervention or for encouraging an environment in which social improvement can occur.
- Policy evaluation – assessing the effectiveness of interventions in achieving desired social outcomes.

He concludes that their strength is in contributing to a sharpened policy focus.

4. Risks in their use

49. There are two primary risks associated with the use of outcome measures within government. First, the literature suggests that scarce political, managerial and practitioner time can be diverted towards consideration of measures which, although important, can only represent one contribution to decisions concerning planning or accountability and control. Second, there is the ever-present risk of gaming or self-serving manipulation of measures of performance.¹⁵

4.1. Demanding scarce political and managerial attention

50. The challenge in outcome measures is one of balancing managerial and political time. To the extent that outcome measures provide an attractive topic for policy makers, not least because they are high level and visionary, then they are particularly seductive – even at the risk of requiring a disproportionate amount of limited management time. Thus while outcome measures can usefully contribute to decisions

15. Gaming has been usefully defined as “reactive subversion such as ‘hitting the target and missing the point’ or reducing performance where targets do not apply” (Beyan and Hood, 2005, p. 8). It is discussed in more detail below.

concerning planning and concerning accountability and control (both in providing a basis for discussion about results achieved and about the quality of government outputs), and can equally usefully provide a frame or a vision for subsequent policy decisions, they do not remove the need for government to control spending and to select modes of production which enhance efficiency.

51. The Minister of Finance for the Netherlands has noted this risk in arguing for cautious sequencing in budget reforms: “I believe that things should be done in the proper sequence. I think that, for example, introducing accrual accounting for central government when accounting capacity still falls short and spending controls are weak is not the way to go forward... The same is true when budgets are primarily based on outputs and outcomes while fiscal rules are non-existent and the position of the Minister of Finance is relatively weak” (Zalm, 2004, p. 101-2).

52. The use of outcome measurements and the other concerns for managing inputs and processes are not intrinsically in tension – except for the problem of finite management and political attention. In Australia, performance budgeting reforms were initiated in order to facilitate discussions on output and even outcome in parliament. Yet, the output and outcome information that was provided by the departments was very broad whereas input information was reduced. As a result, parliament felt that it lost some control over the executive branch (Van Dooren and Sterck, 2006).

4.2. *Gaming*

53. Gaming refers to the strategic reaction of individuals, organisations or countries to the use of measures. Two kinds of reactions can be distinguished. One entails the manipulation of the measures that are selected. In this case, the operations remain the same but the representation of these operations by means of the indicators is deliberately skewed. This results in a loss of the quality of the data. The alternative is to alter the output itself. This usually results in a loss of the quality of the output. A combination of both is also possible.

54. In principle, gaming might be considered a less significant concern in relation to outcome measures than output measures for two reasons. First, there are relatively few attempts to make a tight connection between outcome measures and planning and control/accountability decisions. To the degree that they are connected to decisions somewhat loosely, used as the basis for a discussion to be considered along with other factors rather than as a single driver, the incentives to game are correspondingly less. Second, outcome data are generally produced by entities other than those directly affected by the use of the measures, as suggested by Burgess, Propper *et al.* (2002).

55. However, the literature suggests that this apparently loose relationship to decision-making does not remove the risk of gaming completely in relation to the upstream manipulation of measurement. James (2004) identifies situations in the United Kingdom when ministers seemingly responded to a somewhat “punitive” environment and set their departmental targets low. He notes that the initial outcome targets in relation to criminal justice included reducing the fear of crime, speeding up case processing and ensuring that no local authority area has a crime rate of three times the national average. Subsequently, these were revised to a target of reducing robbery by 14%.¹⁶

56. This problem of cherry-picking of outcome measures, in order to set a low hurdle or to create a particular impression about progress (or lack of it) is exacerbated by the latitude that is inevitable in selecting appropriate outcome measures (Plantz, Greenway *et al.*). Such judicious selection of measures can hide major problems – including, for example, inequitable impact by gender or on different social or ethnic groups.

16. See also De Biase (2005, Appendix A).

57. The key challenge with outcome data is to maintain a loose relationship to decision-making in practice. There is an intrinsic problem with outcome targets in this regard. While for any formal planning or accountability purposes it can be carefully noted that they represent just one contribution to a debate, the media is unlikely to adopt such a nuanced position. As James (2004) notes, there is significant asymmetry in that the media disproportionately report missed targets.¹⁷

56. Although the risks around gaming are clear, it seems reasonable to believe that those risks can, be mitigated through ensuring that the data are collected, analysed and disseminated using transparent processes. In many countries, such data are likely to receive considerable public scrutiny by central agencies, parliament, the media or civil society.

5. Filling an apparent gap

57. In looking at how *de facto* outcome measures seem to be selected, it was noted earlier (see Box 3) that they correspond broadly to the functional areas defined in the Classification of Functions of Government (COFOG). There is one particular gap however. In addition to the obvious functional sectors (defence; public order and safety; economic affairs; environmental protection; housing and community amenities; health; recreation, culture and religion; education; social protection) COFOG includes a set of activities identified as general public services. This is something of a catch-all, intended to reflect the activities of the core machinery of the public sector which can not readily be reflected in the sectoral activities, including the work of the legislature etc. It is striking that there is no equivalent in the well-being measures reviewed in the Annex. In effect, there is a gap in the “well-being” measures around the area of governance.

58. In some ways this is hardly surprising, for two reasons: the entire field of governance is ill-defined, and the policy relevance for governments is hard to establish. For these reasons the recent experiences in the use of governance indicators have not been promising (see Box 8).

17 Questions to witnesses from the UK Select Committee on Public Administration (Minutes of Evidence, Examination of Witnesses (Questions 980-999), Monday 24 March 2003 elicited the comment that “‘Government meets targets’ is not a newspaper story, it is not a newspaper story at all. The fact that we reached our five As to Cs targets was barely covered in the newspapers. Had we not reached it, it would have had much greater coverage... When that happens in business people do not have to resign as a result of not reaching a target but in politics it causes a lot of problems, not just because the media are evil but because politicians themselves go round bragging about all of these wonderful targets we have hit, using them in that way when they are successful but then when there is a failure finding excuses.” <http://www.publications.parliament.uk/pa/cm200203/cmselect/cmpublicadm/62-x/3032404.htm>

It is possible, however, that this reflects a distinctly Anglo-Saxon hostility of the press towards government.

Box 8. Governance indicators¹⁸

Over the last 30 years, there has been a significant increase in the number and range of indicators, purporting to measure a rather broader notion of governance.¹⁹ The theoretical basis behind the indicators is uncertain. Public management as a field of study and practitioner debate is fraught with definitional problems, but there is some general consensus as to what is included. Few would take issue with the proposition that: "*public management reforms consist of deliberate changes to the structures and processes of public sector organizations with the objective of getting them (in some sense) to run better*" (Pollitt and Bouckaert, 2004, p. 8). In essence, *public management* focuses on the plumbing of the public sector – how money is used, how people are motivated and how organisations are structured.

By contrast, the parameters of governance are less clear. "...[W]hile empirical research links governance-related institutions and development, there is not yet a consensus as to how to approach governance and its measurement" ("Global Monitoring Report 2006: Strengthening Mutual Accountability – Aid, Trade and Governance", 2006, p. 11). The idea of governance is undoubtedly identifying a real and pressing issue – and all practitioners with experience of developing or post conflict countries would accept that the roots of problems concerning corruption or major service delivery failures are not be found within the public sector management plumbing. Broader political economy issues, such as capture of the state apparatus by entrenched interests, deliberate opacity in political party financing and corrupt judiciary, are both the consequences and the causes of the performance failures. However, exactly what is within this concept and, equally significantly, what and how any elements of it can be changed is far from clear.

The absence of a well-accepted theoretical framework for governance ensures that any composite indicators are largely devices for communication – for crystallising concerns about corruption etc. into a single short and pithy summary. This is undoubtedly a useful contribution to structuring the dialogue between donor agencies and developing countries – particularly when accompanied by more explicitly qualitative assessments (Doig, Mclvor *et al.*, 2006). However, where the concerns about broader notions of governance are less pressing, where egregious corruption and massive service delivery failure are relatively rare, then the value of the indicators is somewhat limited for two reasons.

First, the accuracy of the indicators is often insufficient for a more nuanced policy debate. Individual governance indicators tend to be perception-based, which can be problematic unless triangulated with some more fact-based indicators. Composite indicators generally combine the problems inherent in an over-reliance on perception-based data, with additional difficulties resulting from the aggregation process. Without a clear theoretical framework, it is not clear what the indicator is attempting to measure other than in a general and intuitive sense (Besançon, 2003).

Second, and as a consequence of the over-reliance on perception-based indicator, and the resort to intricate aggregation methodologies, the resulting indicators provide very little policy guidance. Arguably, they indicate that a particular institutional area within a particular country is indeed in poor (or good) shape – however, they offer little or no insights into the steps that should be taken to improve the situation.²⁰

59. Reducing the vast and somewhat formless territory of governance down to “executive governance” might offer more hope.²¹ Part of the problem is that broad measures of governance such as transparency or rule of law implicate all parts of the state – including electoral institutions and the legislature, and the functioning of the judiciary. It might be possible to identify a set of outcomes for “Government at a Glance” that meet the same apparent criteria as other existing “well-being” measures” (Box 3), but which reflect activities of the executive – but specifically not the legislature or judiciary.

60. Measuring Australia’s Progress offers some examples of possible outcome measures along these lines (Australian Bureau of Statistics, 2006). The focus within New Zealand on “Development Goals for State Services” is somewhat narrower and thus closer to the concept of “executive governance outcomes” (see Box 9).

18. This box is drawn from Annex 3 in Technical Paper 1: How and Why should Government Activity be Measured in "Government at a Glance"?, OECD/GOV.

19. See Knack and Manning (2000).

20. See Knack, Kugler *et al.* (2003) for a more detailed discussion of this point.

21. There is a small literature on this term. See Goetz (2001).

Box 9. Development Goals for State Services in New Zealand

In March 2005 the Government of New Zealand agreed an ambitious set of Development Goals for the State Services. The 2006 progress report notes that the ideas behind these goals are not new, but, by making them clear and visible, and by setting timelines for progress, starting with milestones for 2007 and 2010, a clear agenda has been set for the State Services (State Services Commission, 2006) although of course the problem of obtaining supporting hard data remains.

Development goal	Indicator
<p><i>Employer of choice</i> Ensure the State Services is an employer of choice attractive to high achievers with a commitment to service</p>	<p>1 Reputation: Perceptions of the State Services as a place to work 2 Staff engagement: Engagement levels of staff in State Services agencies</p>
<p><i>Excellent state servants</i> Develop a strong culture of constant learning in pursuit of excellence</p>	<p><i>Staff:</i> 1 State servants' perceptions about the effectiveness of development plans, processes and opportunities <i>Chief executive/agency:</i> 2 The extent to which agencies are specifying the competencies they require of their people to meet current and future organisational needs <i>Sector:</i> 3 The extent to which a learning and development framework is used across the sector to improve agencies' performance</p>
<p><i>Networked State Services</i> Use technology to transform the provision of services for New Zealanders</p>	<p>1 Grouping of services/transactions that apply technology to allow an individual – from one place at the same time – to access multiple programmes 2 Channel synchronisation of government transactions – within an agency or across government 3 The extent to which technology supports a user having to give the same information to government only once</p>
<p><i>Coordinated State Agencies</i> Ensure the total contribution of government agencies is greater than the sum of its parts</p>	<p>1 The extent to which behaviours exhibited by state servants support coordination in pursuit of results 2 The extent to which systems support strategy, design and service delivery staff to work together</p>
<p><i>Accessible State Services</i> Enhance access, responsiveness and effectiveness, and improve New Zealanders' experience of State Services</p>	<p><i>Accessible State Services:</i> 1 Target group uptake of services <i>Responsive State Services:</i> 2 Appropriateness of referral <i>Effective State Services:</i> 3 Users' experience and expectations inform service design and improvement</p>
<p><i>Trusted State Services</i> Strengthen trust in the State Services, and reinforce the spirit of service</p>	<p>1 Trustworthiness exhibited by state servants 2 The extent of New Zealanders' confidence in the integrity of state servants when delivering services</p>

Source: State Services Commission (2006, pp. 14-15).

61. For “Government at a Glance”, “executive governance outcomes” might be broadly of three types: public confidence, equity and fiscal/economic stability (see Box 10). Public confidence might encompass issues around trust in government, and associated concerns relating to the predictability and acceptability of government policy. Equity might encompass the measured distribution of services and benefits across diverse populations. Fiscal and economic stability might relate to the track record of government in these spheres. This paper offers some initial thoughts on the development of comparative measures of trust in government. Measures of equity and economic and fiscal stability will be developed later.

Box 10. Selecting Executive Governance Outcomes

The literature concerning the outcomes that might be more distinctly attributable to the executive (executive governance outcomes) than to other contributors to governance (electoral institutions and the legislature, the judiciary, free media etc.) suggests that they fall into three broad and overlapping categories: equity, fiscal/economic stability and public confidence.

Michalos (2006) sets out the key dimensions of equity in service provision, and social outcomes. This is clearly a key concern for OECD governments and relates to a traditional value of the public sector: impartiality. Many commentators have associated this with representativeness within the public sector on the basis that impartiality is all but impossible in practice without this.²²

Fiscal and economic stability in the sense of avoiding, or at least managing, shocks and in ensuring inter-generational equity are well-recognised as goals. Providing greater fiscal room for manoeuvre through reducing budgetary deficits is often identified as a concern (Schick (2003) for example). Inter-generational inequity is most often discussed in terms of state pension arrangements which provide extensive benefits but which may defer costs to future generations of tax-payers (see OECD (2005d). It should of course be kept in mind that fiscal discipline is certainly not simply attributable to isolated policy or institutional reforms within the executive. Hallerberg (2004) points out that some member states with the ideal institutional arrangements for fiscal discipline are still incurring deficits in violation of the Maastricht Treaty's excessive deficit procedure and the Stability and Growth Pact because of political pressures and electoral and party systems.

Overarching these two key categories of "executive governance outcomes", public confidence in the public sector is a key concern, and closely relates to the perceived legitimacy of government and the willingness to pay taxes (Bird, 2004 and OECD, 2005c).

The empirical aspects of many measures of trust are hard to unravel largely because of the conceptual problems about what is being trusted (OECD, 2000). However, it seems increasingly the case that confidence can be improved through demonstrated responsiveness on the part of the public service to political priorities. Responsiveness to political priorities is now widely seen as a legitimate way of being responsible to the citizens (Dunn, 1997; Hood and Peters, 2004; Self, 1972). At a time of increasingly frequent public-opinion polls, e-mail, call-in radio and television surveys, greater responsiveness is expected of the administration (Rosenthal, 1997). As Schick (2005b) has pointed out, governments must increasingly earn their legitimacy through delivering on their service delivery promises.

²² An early exploration of the relationship between impartiality and representativeness was provided by Kaufman (1956). Others have picked up the baton and debated the risks and benefits of active representation and generally agree with Mosher (1968) in coming down on the side of passive representation – i.e. the belief that if the rules are fair and balanced, then selecting merit will (more or less) automatically lead to representativeness.

5.1. *Developing comparative measures concerning “trust in government”*

62. Gaining (regaining, some would argue) the trust of citizens has emerged as a core concern for governments.²³ Babb (2005) identifies trust in government as an aspect of social capital along with civic participation, membership of social networks, and social participation through involvement in groups and voluntary activities.

5.1.1. *Could trust measures be equivalent to other well-being measures?*

63. Can some measure of trust meet the criteria set out in Box 2? Noting that it is broadly in line with the COFOG functional classifications, and taking the remaining criteria in turn:

64. *Is it generally desired by citizens?* It is hard to imagine a preference for living in a country with an untrustworthy government or civil service any more than one can imagine a preference for particularly poor air quality or short life expectancy. Improving trust in government is certainly one of the common features of discussions concerning the need for various types of public sector reform (Pollitt and Bouckaert, 2004), and so by popular conviction at the very least, it must be regarded as an outcome. There is certainly a technocratic consensus that it is desirable. Van de Walle, Van Roosebroek *et al.* (2005) point out from their review of the literature that high levels of public trust stimulate public sector productivity, since trusting citizens are more willing to comply with regulations and procedures, lowering transaction costs. They also note from that review that having trusting citizens may influence the willingness to make sacrifices during a crisis, to obey the law, to pay taxes or to serve in the military.

65. *Is it significantly affected by government actions?* This is both intuitively plausible and empirically supported. Chanley, Rudolph *et al.* (2000) show that, in the United States, negative perceptions of the economy, scandals associated with Congress, and increasing public concern about crime each lead to declining public trust in government. However, the connection is, to say the least, somewhat obscure. For example, Barnes and Gill (2000) and Bok (1997) found that confidence dropped while performance improved. It is probably significantly affected by government sector policy actions and outputs. For example, Breeman (2003) found evidence that trust was affected by government actions in the case of agriculture.

66. *Is it not easily attributable to the outputs from a single sector?* It is probably significantly affected by cross-cutting structures and systems within government. Yackee and Lowery (2005) for example, in the case of the United States federal bureaucracy, demonstrate that scandals or policy failures do have an impact on overall public approval ratings – and not just for the individual agency concerned (others are tarred with the same brush). Killerby (2005) shows, from World Values Survey Data, that confidence in the bureaucracy is not correlated with social trust or with life satisfaction, diminishing the argument that the public are taking out frustrations on government.²⁴

5.1.2. *Would such measures be policy relevant?*

67. If it could be clarified, then the notion of trust in government as an institutional outcome could have distinct policy relevance. The proponents of particular institutional reforms suggest that there is an undoubted chain of connections:

23. See: http://www.oecdobserver.org/news/categoryfront.php/id/1269/Spotlight:_Trust_in_government.html

24. Although Christensen and Læg Reid (2003) finds that attitudes formed by political cynicism have the strongest overall effect on variation in people's trust in government.

- OECD work on ethics and conflict of interest argues that trust is enhanced by the reality of perceived integrity in public decision making (OECD, 2000, 2003, 2005a). Others maintain that improved service delivery arrangements are key to improvements in trust (Cowell, Downe *et al.*, 2006; Heintzman and Marson, 2005; Humphreys, 2003; Kampen, Van de Walle *et al.*, 2003; Van Ryzin, Muzzio *et al.*, 2004) – although this is challenged on both empirical and theoretical grounds by Bouckaert and Walle (2003) and Zouridis (2003).
- Heintzman and Marson (2005) have also suggested that citizen satisfaction is correlated with public service staff job satisfaction.
- OECD (2001) suggests that greater public participation is key to improved trust.

68. Others argue the reverse – with Suleiman (2005) concluding from an international comparative survey that more than two decades of unrelenting reforms of the state have produced no change (except for the worse) in voter turnout, confidence in politicians, and respect for public institutions. This raises a question of whether declines in some measures of trust are actually linked to “reform”, or whether they happen independently. Others argue (Nevitte: 2006) that a reduction in trust might be a natural manifestation of a more robust democracy, with greater participation in less formal political activity.

69. Yet others have argued that the question is moot until there are better measures of service delivery and performance (Yang and Holzer, 2006).

5.1.3. *Would they introduce a risk of gaming?*

70. As a measure, trust does seem somewhat amenable to methodological cherry-picking as it is conceptually very elusive, with two dimensions of choice concerning measures. First, there is the question of *what the phenomenon engenders in the citizen* – is the citizen reporting that, in retrospect, she/he was *satisfied* with one or several services, or that she/he is prospectively *trusting* that future actions from government will be acceptable? Second, there is a set of very important questions concerning *the unit of analysis* – is the citizen referring to a particular service or entity, or to a major institutional area (e.g. civil service or parliament), or to some abstract notion of government as a whole?

71. On the first dimension, Van de Walle *et al.* (2005) point out from their review of survey data that trust and satisfaction are frequently linked but far from identical. Negative overall views of government often coincide with quite positive evaluations of specific services.

72. On the second dimension concerning the unit of analysis, the World Values Study distinguishes between: the press; parliament; labour unions; government; major companies; civil service; European Union; NATO; justice system; churches; social security system; UN; armed forces; health care system; police; education system; and political parties. However, the survey results indicate that different groups in the countries surveyed have very different views about these institutions. This may very plausibly mean that they have different experiences of these institutions – but it could also suggest that there is some uncertainty about what each of the institutional units represents. Van de Walle (2005) also notes that distrusting attitudes towards the civil service may coexist with very positive evaluations of some specific agencies, such as the fire department, the municipal administration or the postal system. They point out that this suggests that while actual experience with a service will dominate in a customer satisfaction survey, citizens are more likely to refer to their overall image of government or to stereotypical images of the bureaucracy when expressing an overall opinion of the public administration. Arguably, there is also a level of romanticisation concerning specific services that citizens rarely interact with. This might explain why certain types of services consistently receive higher scores than others. Fire departments are almost always evaluated much more positively than others, such as road repair services are.

73. Van de Walle *et al.* (2005) reasonably, take the civil service to be the primary institutional unit when interpreting trust within the World Values Survey results. The National Election Studies use various measures of trust in government in the United States, but it is possible that the term government in the United States is interpreted more inclusively (traditionally the term refers to all three “branches” – executive, legislature and judiciary) and more politically (including elected officials) than in other countries.

74. The ambiguities in the way that trust can be measured seem to have facilitated some flexibility in interpretation which could verge on gaming as apparent declines in trust are “hyped” at key moments in the political cycle. In particular, there is a frequent assertion that confidence in government (possibly interpreted as confidence in the civil service) has been in decline and it is the task of reformers or incoming governments to correct this trend through some radical reform programme (Nye, Zelikow *et al.*, 1997; Perry and Webster, 1999). “[T]he global reform movement is a symptom of – and a reaction to – the decline of public confidence in governmental institutions and performance” (Kettl, 2000, p. 57). This persistent drumbeat of concern has been heard for over 30 years (Crozier, Huntington *et al.*, 1975). Others note that measures of trust are sensitive over the short-term to the crises of the day and to media commentary (Barnes and Gill, 2000).

75. Bouckaert, Lægreid *et al.* (2005, p. 460) find that decline in trust has been used to legitimize public sector reform. Similar points are made more passionately by Suleiman (2005) who argues that this has been a somewhat cynical approach by politicians who are keen to deflect criticism from their own inability to avoid inflation, deficits, and economic instability. Similar concerns have been expressed in relation to the United States by Garrett, Thurber *et al.* (2006).

76. In fact, the World Values Survey data do not reveal clear trends over time (Van de Walle *et al.*, 2005, Technical Appendix).

77. There are also country specific surveys which suggest diverse patterns:

- National Election Studies data in the United States do indeed show a decline in trust since the late 1950s. There has however, been a significant recovery since 1994.
- In most European countries, recent trends can be mapped using Eurobarometer data. Trust in the civil service was included several times since 1997, with the last measurement in spring 2002. Of the EU15 countries included in the 1997 and 2002 Eurobarometer surveys, only three face a decline in trust (Van de Walle *et al.*, 2005, Technical Appendix).

5.2. *Developing comparative measures concerning equity*

78. Consideration will be given to the development of outcome measures in this area over the coming year. Such measures might reasonably cover the distribution of key services by socio-economic or other group. Michalos (2006) provides a useful summary of possibilities.

5.3. *Developing comparative measures concerning economic and fiscal stability*

79. In later stages of the project, consideration will also be given to the development of outcome measures concerning fiscal/economic stability. Such measures might reasonably cover issues such as budgetary deficits (as a contributor to economic and fiscal instability) and other budgetary outcomes. There has already been extensive work examining the relationship between these outcomes and the institutional arrangements within government. Savage (2005) notes that OECD and other governments have attempted to restrain large-scale budgetary deficits and debt by institutional reforms which have created processes

which are more centralized, “top-down” and “front-loaded” in order to reduce the influence of interest groups and budgetary claimants.²⁵ There have been extensive studies of the contribution of institutional arrangements to fiscal discipline.²⁶ Hallerberg (2004) has classified the key institutional drivers of these particular budgetary outcomes under three broad headings:

- “Delegation governance” entails centralized budgetary systems with strong finance ministries, with the ability to formulate budgets, monitor ministry behaviour, and enforce spending rules against spending ministries and free-spending ministries.
- “Commitment governance” refers to situations where “fiscal contracts” set a variety of spending, deficit, and other targets that require budgetary players to adhere to commitments.
- “Fiefdom governance” refers to situations where finance ministries are unable to control the demands of spending ministries and legislators for distributive project funding.

6. The value of internationally comparable data on “executive governance outcomes”

80. There are several ways in which internationally comparable public management data can assist governments and other analysts:

- *For individual countries*, such data can enable robust benchmarking between countries, using common units of analysis, facilitating a structured practitioner dialogue and moving away from simplistic best practice.
- Comparable data can contribute to *OECD-wide lesson-learning* concerning:
- Sector efficiency and broader measures of institutional effectiveness, providing insights into the results of providing services via different institutional and managerial arrangements.
- Causal relationships (which changes in public sector processes are associated with which changes in outputs?)
- Management of resource changes (identification of absorptive capacity constraints following significant increases in sector expenditures, and the converse).

81. Maintaining a database of outcome measures in key sectors will assist in many areas of this agenda. This includes most particularly benchmarking for individual countries. At the OECD-wide level it could assist in the development of measures of institutional effectiveness and, through monitoring change over time, it could assist in unpacking causal relationships.

25. “Macrobudgetary guidelines set at the beginning of the budgetary process constrain micro-decisions during the remainder of the process. Governments employ hard targets, ceilings, and caps to limit spending; calculate long-term inflationary and program costs through baseline budgeting; enhance program evaluation and financial management systems; create new legislative committees with budgetary oversight responsibilities; establish support agencies to impose greater oversight on the budgetary process; enact budget resolutions and reconciliation devices to control entitlements; and strengthen the powers of finance ministers in constraining spending demands.” (Savage, 2005, p. 1)

26. Campos and Pradhan (1996) was one of earliest studies on this issue. More recently there has been an extensive array of empirical work. See particularly de Hann, Moessen *et al.* (1999); Hallerberg, Strauch *et al.* (2001a); Hallerberg, Strauch *et al.* (2001b); Poterba and Hagen (1999); von Hagen (1992); von Hagen and Harden (1995); Wanna, Jensen *et al.* (2003), particularly in relation to EU countries.

6.1. *Benchmarking and structured practitioner dialogue*

82. Benchmarking is a structured debate between practitioners, agencies or governments concerning how and why things are different between them. The purpose of benchmarking is to open up issues for subsequent investigation – to provoke interest in deeper examinations. Benchmarking can be used to compare inputs, processes, outputs or outcomes.

83. National or governmental policy stances tend to be defined in rather broad terms. Setting out comparative data on what governments are measuring as outcomes would allow policy concerns to be explored more concretely.

6.2. *Developing measures of institutional effectiveness*

84. While efficiency or productivity builds on the relationship between inputs and outputs, institutional effectiveness examines the relationships between, on the one hand, public sector structures and processes and, on the other, executive governance outcomes. Thus, for example, the inclusive scope of “government” proposed for the publication series could, over time, provide insights into the results of different oversight arrangements in relation to changed measures of public trust?

6.3. *Monitoring change through comparisons over time*

85. Robust comparative measures of outcome would contribute to OECD-wide lesson learning concerning the complex attribution problems in improving public sector outputs and outcomes. Since agency or programme designs generally can not be adjusted experimentally to assess impact, the challenge is how to determine whether and in what proportion programme activities and public sector processes contribute to outputs, and similarly which outputs contribute significantly to which outcomes. Time series output data would improve understanding of the associations between output changes and changes in outcomes.

6.4. *How could “Government at a Glance” help?*

86. If a tighter set of executive governance outcome measures could be developed, then benchmarking results between countries and over time would undoubtedly be of tremendous interest to OECD member countries and policy analysts.

87. Key data relevant to trust that are currently available offering some international comparability are:

- World Values Study (WVS). This survey has been organised in four waves since 1981: 1981, 1990, 1995-1997 and 1999-2000. All OECD member countries have participated in one or more waves.
- Eurobarometer data provides data from survey questions concerning trust in the civil service, which was included several times since 1997, with the last measurement in spring 2002.²⁷
- Many of the underlying data used in the construction of the World Bank Institute "Worldwide Governance Indicators" can be interpreted as measures of public and business trust in government (see: <http://info.worldbank.org/governance/kkz2005/notes.html#kk>).

27 See Eurostat (2006)

- The World Bank "Doing Business" database (see: www.doingbusiness.org) provides a broad array of data concerning the regulatory costs of business, and can provide insights into business confidence in government .

88. However, there are many conceptual uncertainties in relation to the question that is being addressed in these surveys. On trust/confidence, survey questions would need to be framed which distinguish between the public's views of government policy and their views of the quality of the administration. It is the latter which is of institutional interest – but this would then need to be associated with some comparable units of analysis (which level of government, etc.), and some further distinction made between satisfaction concerning responses in the past, and trust in the quality of services in the future.

89. Ideally, an effort on the scale of the OECD PISA survey would be useful. This might not be completely impracticable as it seems at least feasible that the institutions already undertaking relevant surveys currently might be prepared to discuss collaboration within a larger effort.

90. A consistent data set from the resulting question would also allow some movement towards assessing institutional effectiveness – addressing the question of which reforms of structures and processes really are associated with concrete changes in particular aspects of trust.

7. Summary of the key propositions

7.1. How well-being indicators are selected and used

91. The seeming *de facto* criteria for how governments select “well-being” indicators are:

1. Generally desired by citizens.
2. Broadly corresponding to the functional areas defined in the Classification of Functions of Government (COFOG).
3. Probably significantly affected by government actions.
4. Not easily attributable to the outputs from a single sector (although they can be, and often are, readily attributable to the activities of government as a whole).
5. Entailing measures that are available at reasonable cost.

92. There are three principal ways in which outcome measures can be deployed by government:

1. As a contribution to planning decisions.
2. As a contribution to accountability and control decisions:
 - Providing a basis for discussion concerning results achieved.
 - Providing a basis for discussion about the quality of government outputs.
3. As a frame or a vision for subsequent policy decisions:
 - To focus policy thinking.
 - To stimulate a public and political debate.

- To develop a common agenda within government when diverse sectors and agencies within government are involved.
- To inspire additional effort by appealing to a broader good.

7.2. *The risks of using outcome measures indicators*

93. There are two primary risks associated with the use of outcome measures within government. First, the literature suggests that scarce political, managerial and practitioner time can be diverted towards consideration of measures which, although important, can only represent one contribution to decisions concerning planning or accountability and control. Second, there is the ever-present risk of gaming.

7.3. *The potential contribution of “Government at a Glance”*

94. Outcome measures correspond broadly to the functional areas defined in the Classification of Functions of Government (COFOG). There is less activity in relation to “well-being” measures around the area of governance. This is likely to be for two reasons: the entire field of governance is ill-defined, and the policy relevance for governments is hard to establish. Developing measures of “executive governance” outcomes might offer a practical way forward for “Government at a Glance”. These would meet the same apparent criteria as other existing “well-being measures”, but would reflect activities of the executive – but specifically not the legislature or judiciary.

95. For “Government at a Glance”, “executive governance outcomes” might be broadly of three types: public confidence, equity and fiscal/economic stability. Public confidence might encompass issues around trust in government, and associated concerns relating to the predictability and acceptability of government policy, or measures of public perception about government services or policies. Equity might encompass the measured distribution of services and benefits across diverse populations. Fiscal and economic stability might relate to the track record of government in these spheres.

96. In relation to trust, some data are available, but there are many conceptual uncertainties in relation to the question that is being addressed in the surveys. Survey questions would need to be framed which distinguish between the public’s views of government policy and their views of the quality of the administration, and some further distinction made between satisfaction concerning responses in the past, and trust in the quality of services in the future. Ideally, and if we could be sure that “trust in government” was a useful diagnostic, then an effort on the scale of the OECD PISA survey would be useful. This might not be completely impracticable as it seems at least feasible that the institutions already undertaking relevant surveys currently might be prepared to discuss collaboration within a larger effort.

98. If robust and comparable outcome measures can be developed, then there are several ways in which such internationally comparable public management data can assist governments and other analysts:

- *For individual countries*, such data can enable robust benchmarking between countries, using common units of analysis, facilitating a structured practitioner dialogue and moving away from simplistic best practice.
- Comparable data can contribute to *OECD-wide lesson-learning* concerning:
 - Sector efficiency and broader measures of institutional effectiveness, providing insights into the results of providing services via different institutional and managerial arrangements.

- Causal relationships (which changes in public sector processes are associated with which changes in outputs?)
- Management of resource changes (identification of absorptive capacity constraints following significant increases in sector expenditures, and the converse).

ANNEX

WELL-BEING MEASURES WHICH ARE TAKEN INTO ACCOUNT AT LEAST IN TWO
"NATIONAL SUITES"²⁸

INDICATORS	AUSTRALIA	CANADA	FINLAND	NEW ZEALAND	SWITZERLAND	UK	EUROSTAT
1. Health	✓	✓	✓	✓	✓	✓	✓
1.1. Life expectancy at birth/healthy life expectancy	✓	✓	.	✓	✓	✓	✓
1.1.1. Incidence of all cancer/skin cancer	✓	.	✓	.	.	.	✓
1.1.2. Infant mortality rate	✓	✓
1.2. Obesity	.	.	✓	✓	.	✓	✓
1.3. Expenditure on prevention and health promotion	✓	.	✓
2. Education & training	✓	✓	✓	✓	✓	✓	✓
2.1. Public expenditure on education as a % GDP	✓	.	✓
2.2. Education levels	.	✓	✓	.	.	✓	.
2.3. Educational attainment of the adult population	.	✓	.	✓	.	.	.
2.4. Adult literacy	.	✓	.	✓	.	.	.
3. Work	✓	✓	✓	✓	✓	✓	✓
3.1. Unemployment rate	✓	.	✓	✓	✓	.	.
3.2. Employment rate	.	✓	.	✓	.	✓	✓
4. National income	✓	✓	✓	✓	✓	✓	✓
4.1. Real net national disposable income per capita	✓	.	.	✓	.	.	.
4.1.1. Real GDP per capita	✓	✓	.	.	✓	.	.
4.2. Investment as % of GDP, by institutional sector	✓	✓	✓
4.3. Inflation rate	.	.	✓	.	.	.	✓
5. Financial hardship	✓	.	✓	✓	✓	✓	✓
5.1. Inequality of income distribution	✓	.	✓
6. Housing	✓	✓	.	✓	✓	✓	✓
7. Productivity	✓	✓	✓	.	✓	✓	✓
7.1. Labour productivity	✓	.	.	.	✓	.	✓
7.2. Expenditure on R&D	✓	.	✓	.	✓	.	✓
8. The natural landscape	✓	.	✓	.	✓	✓	✓
8.1. Threatened species trend	✓	✓
8.2. Fish stocks	✓	.	✓	.	.	✓	.
8.3. Forest resources	.	.	✓	.	✓	.	.
8.4. Biodiversity	✓	✓	.

28. Developed from original work by Ayhan Uysal, OECD, Statistics Directorate.

INDICATORS	AUSTRALIA	CANADA	FINLAND	NEW ZEALAND	SWITZERLAND	UK	EUROSTAT
9.The human environment	✓	✓	✓	✓	✓	✓	✓
9.1.Air quality	.	✓	✓	✓	.	.	.
9.2.Emissions of air pollutants	✓	✓
10. International environmental concerns	✓	✓	✓	.	✓	✓	✓
10.1.Net greenhouse gas emissions	✓	.	✓	.	✓	✓	.
10.2.Thickness of ozone layer	.	.	✓	.	✓	.	.
10.3.Mean temperatures	.	.	✓	.	✓	.	.
11. Family community and social cohesion	✓	✓	✓	✓	✓	✓	.
11.1.Suicide and drug-induced death rates	✓	.	✓	✓	✓	.	✓
12.Crime	✓	✓	✓	✓	✓	✓	.
12.1.Violent and drug crime	.	.	✓	.	✓	.	.
12.2.Killed and injured persons in road traffic	.	.	.	✓	✓	✓	✓
13.Governance, democracy and citizenship	✓	✓	✓	✓	.	.	✓
13.1.Voter turnout and informal votes cast	✓	.	.	✓	.	.	.
13.2.Women in Federal Parliament	✓	.	.	✓	✓	.	.
14. Cultural identity	.	.	✓	✓	✓	.	.
14.1.Attendance at cultural events	.	.	.	✓	✓	.	.
15.Transport	.	.	✓	.	✓	✓	✓
15.1.Modal split of passenger transport	✓	.	✓
15.2.Modal split of freight transport	✓	.	✓
15.3.Road length	.	.	✓	.	✓	.	.
16.Global Partnership	.	✓	✓	.	✓	✓	✓
16.1.ODA	.	✓	✓	.	✓	.	.
17. Energy	.	.	✓	.	✓	✓	✓

Sources: Australia: Australian Bureau of Statistics (ABS), Measures of Australia's Progress;

Canada: Treasury Board of Canada Secretariat, Canada's Performance, 2004;

EUROSTAT: European Commission, Sustainable Development Indicators to monitor the implementation of the EU Sustainable Development Strategy SEC (2005) 161 final;

Finland: Finnish Sustainable Development Indicators 2003, Finnish National Commission on Sustainable Development, Finnish Ministry of the Environment;

New Zealand: Statistics New Zealand, The Social Report 2005;

Switzerland: Monitoring Sustainable Development, MONET, Final Report Methods and Results, Swiss Federal Statistical Office, 2004.

United Kingdom: The UK Government Sustainable Development Strategy, 2005;

BIBLIOGRAPHY

- Algemene Rekenkamer (2006), *Performance Audit Manual*. The Hague: European Affairs & Government-wide Performance Audit Division, Netherlands Court of Audit.
- Andersen, Kim Viborg (2004), *E-Government and Public Sector Process Rebuilding*. New York: Kluwer.
- Artley, Will, D.J. Ellison and Bill Kennedy (2001), *The Performance-Based Management Handbook – Volume 1: Establishing and Maintaining a Performance-Based Management Program*. Washington D.C.: Training Resources and Data Exchange (Performance-Based Management Special Interest Group).
- Australian Bureau of Statistics (2006), *Measuring Australia's Progress* ([http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/47132EE72AC3581DCA25717F0004ACE8/\\$File/13700_2006.pdf](http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/47132EE72AC3581DCA25717F0004ACE8/$File/13700_2006.pdf)). Canberra: Australian Bureau of Statistics.
- Babb, Penny (2005), "Measurement of Social Capital in the UK", in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD, pp. 532-5.
- Barnes, Cheryl and Derek Gill (2000), *Declining Government Performance? Why Citizens Don't Trust Government* (http://www.ssc.govt.nz/upload/downloadable_files/working_paper_9.pdf). Wellington: State Services Commission.
- Bellefeuille-Prégent, Louise and Tim Wilson (2005), "Societal Indicators and Government-Wide Reporting in the Government of Canada" in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD, pp. 306-18.
- Besaçon, Marie (2003), "Good Governance Rankings: The Art of Measurement". World Peace Foundation, John F. Kennedy School of Government, Harvard University. Cambridge, Massachusetts.
- Bevan, Gwyn and Christopher Hood (2005), *What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System*. The Public Services Programme. London: Economic and Social Research Council.
- de Biase, Luca (2005), "Agree on Numbers While Discussing Words. Or Is It Just the Opposite?" in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 494-500.
- Bird, Richard M. (2004), *Administrative Dimensions of Tax Reform*. Asia-Pacific Tax Bulletin (March 2004). 134-50.
- Boarini, R., A. Johansson and M. Mira d'Ercole (2006), "Alternative Measures of Well-Being", *OECD Social, Employment and Migration Working Paper No. 33*. Paris: OECD.
- Bok, D. (1997), "Measuring the Performance of Government" in S. Nye, P. D. Zelikow and D. C. King (eds.) *Why People Don't Trust Government*. Cambridge, Massachusetts.: Harvard University Press, pp. 55-76.

- Bouckaert, G., P. Lægheid and S. Van de Walle (2005), "Trust, Quality Measurement Models, and Value Chain Monitoring: Symposium Introduction". *Public Performance and Management Review*. 28 (4), pp. 460-4.
- Bouckaert, G. and S. Van de Walle (2003), "Quality of Public Service Delivery and Trust in Government" in A. Salminen (ed.) *Governing Networks: AGPE Yearbook*. Amsterdam, Netherlands: IOS Press.
- Boyle, Richard (2005), *Civil Service Performance Indicators*. Dublin: Institute of Public Administration.
- Boyne, George and Jennifer Law (2004), "Designing Performance Measurements to Be Drawn on in the Second Generation of Local Public Service Agreements (Local PSAs)" (www.idea-knowledge.gov.uk/idk/aio/384232). Office of the Deputy Prime Minister. London.
- Breeman, Gerard (2003), "Explaining Failures of Regaining Public Trust: The Case of Agricultural Policy-Formation in the Netherlands (Paper to Be Presented on the EGPA Annual Conference, Oeiras 3-6 September 2003)" (http://soc.kuleuven.be/io/egpa/qual/lisbon/paper_lisbon_breeman.pdf). EGPA. Oeiras, Portugal.
- Burgess, Simon, Carol Propper and Deborah Wilson (2002), *Does Performance Monitoring Work? A Review of the Evidence from the UK Public Sector, Excluding Health Care (CMPO Working Paper Series No. 02/49)*. Bristol, United Kingdom: The Centre For Market And Public Organisation.
- Campos, Ed and Sanjay Pradhan (1996), *Budgetary Institutions and Expenditure Outcomes (Policy Research Working Paper No. 1646)*. Washington, D.C.: World Bank.
- Chanley, V.A., T.J. Rudolph and W.M. Rahn (2000), "The Origins and Consequences of Public Trust in Government: A Time Series Analysis". *Public Opinion Quarterly*. 64 (3).
- Cheung, Paul (2005), "Millennium Development Goals: Measuring and Monitoring Global Progress", in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp 374-84.
- Christensen, Tom and Per Lægheid (2003), "Trust in Government – the Significance of Modernism, Political Cynicism and Integration (Paper to Be Presented on the EGPA Annual Conference, Oeiras 3-6 September 2003)" (http://soc.kuleuven.be/io/egpa/qual/lisbon/paper_lisbon_breeman.pdf). EGPA. Oeiras, Portugal.
- Cobb, J. and H. Daly (1989), *For the Common Good: Redirecting the Economy toward Community, the Environment, and a Sustainable Future*. Boston: Beacon Press.
- Comptroller and Auditor General (2001), *Measuring the Performance of Government Departments*. London: National Audit Office.
- Cowell, R., J. Downe, S. Martin and A. Chen (2006), "Making Sense of Service Improvement: An Empirical Analysis of Public Confidence in Local Government (Paper Presented to the 56th Political Studies Association Annual Conference, University of Reading, April 2006)". Centre for Local and Regional Government Research. Cardiff.
- Crozier, Michel, Samuel P. Huntington and Joji Watanuki (1975), *The Crisis of Democracy*. New York: New York University Press.

- Curristine, Teresa (2005), "Performance Information in the Budget Process: Results of OECD 2005 Questionnaire". *OECD Journal on Budgeting*. 5(2). Paris: OECD.
- Davies, H.T.O., S.M. Nutley and P.C. Smith (2000), *What Works? Evidence-Based Policy and Practice in Public Services*. Bristol: The Policy Press.
- Doig, Alan, Stephanie McIvor and Robin Theobald (2006), "Numbers, Nuances and Moving Targets: Converging the Use of Corruption Indicators or Descriptors in Assessing State Development". *International Review of Administrative Sciences*. 72(2). pp. 239-52.
- Dunn, Delmer D. (1997), *Politics and Administration at the Top - Lessons from Down Under*. Pittsburgh: University of Pittsburgh Press.
- van Dooren, W. (2005), "What Makes Organisations Measure? Hypotheses on the Causes and Conditions for Performance Measurement". *Financial Accountability and Management*. 21(3). pp. 363-83.
- van Dooren, W. (2006), "Performance Measurement in the Flemish Public Sector: A Supply and Demand Approach". *Faculty of Social Sciences*. Leuven, Belgium.
- van Dooren, Wouter, Miekatrien Sterck and Geert Bouckaert (2006), "Recent Developments in Output Measurement within the Public Sector: Report Prepared for the OECD". Public Management Institute, Katholieke Universiteit. Leuven, Belgium.
- van Dooren, Wouter and Miekatrien Sterck (2006), "Financial Management Reforms after a Political Shift: A Transformative Perspective". *International Journal of Productivity and Performance Management*. 6(55).
- Dunn, Delmer D. 1997. *Politics and Administration at the Top - Lessons from Down Under*. Pittsburgh: University of Pittsburgh Press.
- East, P. (1997), "From Outputs to Outcomes: Opening Address to the Public Service Senior Management Conference", Wellington, New Zealand, 9 October.
- Eurostat (2001), *Handbook on Price and Volume Measures in National Accounts*. Luxembourg: Office for Official Publications of the European Communities.
- Eurostat (2005), *Measuring Progress Towards a More Sustainable Europe* (http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-68-05-551/EN/KS-68-05-551-EN.PDF). Luxembourg: Office for Official Publications of the European Communities.
- Eurostat (2006), *Public Opinion*. http://ec.europa.eu/public_opinion/index_en.htm. Access Date: 10 September 2006. Last Update: n/k.
- Garrett, R. Sam, James A. Thurber, A. Lee Fritschler and David H. Rosenbloom (2006), "Assessing the Impact of Bureaucracy Bashing by Electoral Campaigns". *Public Administration Review*. 66(2). pp. 228-40.
- Goetz, Klaus H (2001), "Executive Governance in Central and Eastern Europe (Special Issue)". *Journal of European Public Policy*. 8(6).

- Giovannini, E, and Uyshal, A. (2006), *Statistics, Knowledge and Policy: What do We Know about What People Know*, OECD. Presented to the OECD Workshop on Business and Consumer Tendency Surveys, Rome September 2006.
- Gregory, Robert (2004), “Political Life and Intervention Logic: Relearning Old Lessons?” *International Public Management Review*. 5(2). pp. 1-12.
- Guidance on Outcomes Focused Management – Building Block 3: Intervention Logic (Version 2.1, July 2003) (2003), (http://io.ssc.govt.nz/pathfinder/documents/pathfinder-BB3-intervention_logic.pdf).
- von Hagen, Jurgan (1992), *Budgeting Procedures and Fiscal Performance in the European Communities*. Brussels: Directorate General for Economic and Financial Affairs. Brussels: Commission of the European Communities.
- von Hagen, Jurgan and Ian Harden (1995), “Budget Processes and Commitment to Fiscal Discipline”. *European Economic Review*. 39(3-4). pp. 771-9.
- Hall, Jon, Chris Carswell, René Jones and David Yencken (2005), “Collaborating with Civil Society: Reflections from Australia”, in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 434-52.
- Hallerberg, Mark (2004), *Domestic Budgets in a United Europe*. Ithaca: Cornell University Press.
- Hallerberg, Mark, Rolf Strauch and Jurgan von Hagen (2001a), “The Use and Effectiveness of Budgetary Rules and Norms in EU Member States (Report Prepared for the Dutch Ministry of Finance)” (http://www.minfin.nl/binaries/minfin/assets/pdf/old/fiscal_rules_report.pdf). Institute of European Integration Studies. Bonn.
- Hallerberg, Mark, Rolf Strauch and Jürgen von Hagen (2001b), “The Use and Effectiveness of Budgetary Rules and Norms in EU Member States” (<http://www.minfin.nl>). Report Prepared for the Dutch Ministry of Finance.
- de Hann, Jakob, Wim Moessen and Bjorn Volkerink (1999), “Budgetary Procedures – Aspects and Changes: New Evidence for Some European Countries”, in J. M. Poterba and J. v. Hagen (eds.) *Fiscal Institutions and Fiscal Performance*. Chicago: University of Chicago Press. pp. 265-300.
- Hatry, H.P. (1999) *Performance Measurement: Getting Results*. Washington, D.C.: Urban Institute Press.
- Hatry, H.P. (2004), *Public and Private Agencies Need to Manage for Results, Not Just Measure Them* (www.urban.org/url.cfm?ID=900731). Washington, D.C.: Urban Institute.
- Hatry, Harry P., Elaine Morley, Shelli B. Rossman and Joseph S. Wholey (2003), *How Federal Programs Use Outcome Information: Opportunities for Federal Managers*. Washington, D.C.: IBM Endowment for the Business of Government and the National Academy of Public Administration.
- Heintzman, Ralph and Brian Marson (2005), “People, Service and Trust: Is There a Public Sector Service Value Chain?” (<http://ras.sagepub.com/cgi/content/short/71/4/549>). *International Review of Administrative Sciences*. 71(4). pp. 549-75.
- H.M. Treasury (2001), “Public Service Agreements 2001-2004: Chapter 9: Foreign and Commonwealth Office”. (<http://www.hm->

treasury.gov.uk/Documents/Public_Spending_and_Services/Public_Service_Agreements_2001_2004/pss_psa_fco.cfm). London.

- Hood, Christopher and B Guy Peters (2004), *The Middle-Ageing of New Public Management: Into the Age of Paradox?* *Journal of Public Administration Research and Theory*. 14 (3). 267-82.
- Humphreys, Peter C. (2003), “Quality, Satisfaction and Trust from the Perspective of Irish Public Service Modernisation (Paper to Be Presented on the EGPA Annual Conference, Oeiras 3-6 September 2003)” (http://soc.kuleuven.be/io/egpa/qual/lisbon/paper_lisbon_humphreys.pdf). EGPA. Oeiras, Portugal.
- International Monetary Fund and World Bank (2006), “Global Monitoring Report 2006: Strengthening Mutual Accountability – Aid, Trade and Governance”. ([http://siteresources.worldbank.org/DEVCOMMINT/Documentation/20892317/DC2006-0004\(E\)-GMR2006.pdf](http://siteresources.worldbank.org/DEVCOMMINT/Documentation/20892317/DC2006-0004(E)-GMR2006.pdf)). Development Committee of the International Monetary Fund and the World Bank. Washington, D.C.
- James, Oliver (2004), “The UK Core Executive’s Use of Public Service Agreements as a Tool of Governance”. *Public Administration*. 82(2). pp. 397-419.
- Kampen, Jarl K., Steven Van de Walle and Geert Bouckaert (2003), “On the Relative Role of the Public Administration, the Public Services and the Political Institutions in Building Trust in Government in Flanders (Paper Prepared for the ASPA 64th National Conference ‘the Power of Public Service’, Washington D.C., 15-18 March 2003)”. Public Management Institute, Catholic University. Leuven.
- Kaufman, Herbert (1956), *Emerging Conflicts in the Doctrines of Public Administration*. *American Political Science Review*. 50 (4). 1057-73.
- Kettl, D. F. (2000), *The Global Public Management Revolution: A Report on the Transformation of Governance*. Washington, D.C.: Brookings Institution Press.
- Killerby, Paul (2005), “‘Trust Me, I’m from the Government’: The Complex Relationship between Trust in Government and Quality of Governance”. *Social Policy Journal of New Zealand* (25).
- Knack, Stephen, Mark Kugler and Nick Manning (2003), “Second Generation Governance Indicators”. *International Review of Administrative Sciences*. 69(3). pp. 345-64.
- Knack, Stephen and Nick Manning (2000). “Towards Consensus on Governance Indicators: Selecting Public Management and Broader Governance Indicators”. Paper presented at the joint UN/OECD/World Bank/IMF International Development Forum, March 2000. Washington, D.C.
- Kristensen, J.K., W.S. Groszyk and B. Bühler (2001), “Outcome-Focused Management and Budgeting”. *OECD Journal on Budgeting*. 1(4). pp. 7-34. Paris: OECD.
- Logic Model Development Guide (2004). (<http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf>). Michigan: W. K. Kellogg Foundation.
- Lonti, Zsuzsanna and Robert Gregory (2006), “Accountability or Countability Performance Measurement in the New Zealand Public Service, 1992-2002”. *Australian Journal of Public Administration*.
- Matheson, Alex, Boris Weber, Nick Manning and Emmanuelle Arnould (2006), “Managing the Political/Administrative Boundary – Draft Report”. Paris: OECD.

Matthews, Erica (2006), "Measuring Well-Being and Societal Progress: A Brief History and the Latest News (Prepared for the Joint OECD-JRC Workshop "Measuring Well-Being and Societal Progress", 19-21 June 2006, Milan)". Paris: OECD.

Michalos, A.C. (2006). "'Political Culture" and Well-Being: Beyond Government Services (First Draft for Discussion at the JRC/OECD Workshop on Measuring Well-Being and Societal Progress - Milan, 19-21 June 2006)" (http://farmweb.jrc.cec.eu.int/Crell/Well-being/papers/Michalos_Culture%20and%20well-being.pdf). OECD. Paris

Mosher, Frederick C. (1968). *Democracy and the Public Service*. New York: Oxford University Press.

Munoz, Pedro Díaz (2005), "Indicators for EU Policy Making: The Example of Structural Indicators", in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 385-98.

Nevitte, Neil. (2006) *The Decline of Deference: Canadian Value Change in Cross National Perspective*. Broadview Press.

Nye, S., P.D. Zelikow and D.C. King (1997), *Why People Don't Trust Government*. Cambridge Massachussettes: Harvard University Press.

O'Mahony, Mary and Philip Andrew Stevens (2004), "International Comparisons of Performance in the Provision of Public Services: Outcome Based Measures for Education (Paper Presented to the Royal Economic Society Annual Conference, Swansea)".

OECD (1982), "The OECD List of Social Indicators". Paris: OECD.

OECD (2000), "Trust in Government: Ethics Measures in OECD Countries". *Community Development Journal*. Paris: OECD.

OECD (2001), *Citizens as Partners: Information, Consultation and Participation in Policy-Making*. Paris: OECD.

OECD (2002), "Overview of Results-Focused Management and Budgeting in OECD Member Countries". Paris: OECD.

OECD (2003), *Managing Conflict of Interest in the Public Service. OECD Guidelines and Country Experiences*. Paris: OECD.

OECD (2005a), *Managing Conflict of Interest in the Public Sector: A Toolkit*. Paris: OECD.

OECD (2005b), *Society at a Glance: OECD Social Indicators 2005 Edition*. Paris: OECD.

OECD (2005c), *Modernising Government: The Way Forward*. Paris: OECD.

OECD (2005d) "Solving the Pensions Puzzle" (<http://www.oecd.org/dataoecd/53/19/34587956.pdf>). OECD. Paris

OECD Centre for Educational Research and Innovation (2001), *The Wellbeing of Nations: The Role of Human and Social Capital, Education and Skills*. Paris: OECD.

- OECD (2004), *Glossary of Statistical Terms 2004*. (<http://stats.oecd.org/glossary/glossary.pdf>). Paris: OECD.
- Perrin, B. (2006), *Moving from Outputs to Outcomes: Practical Advice from Governments around the World. Managing for Performance and Results Series*. Washington, D.C.: IBM Center for the Business of Government and the World Bank.
- Perry, P. and A. Webster (1999), *New Zealand Politics at the Turn of the Millennium: Attitudes and Values About Politics and Government*. Auckland: Alpha Publications.
- Plantz, Margaret C., Martha Taylor Greenway and Michael Hendricks (1997), *Outcome Measurement: Showing Results in the Nonprofit Sector*. United Way of America. <http://national.unitedway.org/outcomes/library/ndpaper.cfm>. Access Date: 2 August 2006. Last Update: N/k.
- Pollitt, C. (1997), "Looking Outcomes in the Face: The Limits of Government Action (Address to the Public Service Senior Management Conference, Wellington, New Zealand, 9 October)".
- Pollitt, C. (2001), "Integrating Financial Management and Performance Management". *OECD Journal on Budgeting*. 1(2). pp. 7-38.
- Pollitt, Christopher and Geert Bouckaert (2004), *Public Management Reform: A Comparative Analysis*. Oxford, UK: Oxford University Press.
- Poterba, James and Jurgen von Hagen (1999), "Fiscal Institutions and Fiscal Performance (Conference Report, National Bureau of Economic Research)". University of Chicago Press. Chicago.
- Public Sector Accounting Board (2006), *CICA Public Sector Accounting Handbook* (as updated at October 2006). Canadian Institute of Chartered Accountants. Toronto.
- Riche, Martha Farnsworth (2003), *Developing Key National Indicators* (<http://www.gao.gov/npi/usadkni.pdf>). Washington, D.C.: Government Accountability Office.
- Rosenthal, Alan (1997), *The Decline of Representative Democracy: Process, Participation, and Power in State Legislatures*. Washington DC: Congressional Quarterly Books.
- van Ryzin, G. G., D. Muzzio, S. Immerwahr, L. Gulick and E. Martinez (2004), "Drivers and Consequences of Citizen Satisfaction: An Application of the American Customer Satisfaction Index Model to New York City". *Public Administration Review*. 64(3). pp. 331-41.
- Savage, James D. (2005), "Political Development and Fiscal Outcomes (Paper Prepared for the "Long-Term Budget Challenge: Public Finance and Fiscal Sustainability in the G-7" Conference, Washington, D.C., 2-4 June 2005)". University of Virginia. Charlottesville, VA.
- Schick, Allen (2003), "The Performing State: Reflection on an Idea Whose Time Has Come but Whose Implementation Has Not". *OECD Journal on Budgeting*. 3(2). Paris: OECD.
- Schick, Allen (2003b), "The Role of Fiscal Rules in Budgeting". *OECD Journal on Budgeting*. 3 (3). 7-34.
- Schick, Allen (2005), "The Performing State: Reflection on an Idea Whose Time Has Come but Whose Implementation Has Not (Paper Prepared for the 2005 OECD Senior Budget Officials Meeting in Bangkok, Thailand, 15-16 December 2005)" (<http://www.oecd.org/dataoecd/42/43/35651133.pdf>). OECD. Paris

- Self, Peter (1972). *Administrative Theories and Politics - an Inquiry into the Structure and Processes of Modern Government*. London: Allen and Unwin.
- Scott, Graham (2001), *Public Management in New Zealand: Lessons and Challenges* (http://www.nzbr.org.nz/documents/publications/publications-2001/public_management.pdf). Wellington: New Zealand Business Roundtable.
- State Services Commission (2006), “State of the Development Goals Report 2006” (http://www.ssc.govt.nz/upload/downloadable_files/sdg-report06.pdf). State Services Commission. Wellington.
- Suleiman, Ezra (2005), *Dismantling Democratic States*. Princeton: Princeton University Press.
- Summers, Robert and Alan Heston (1999), “The World Distribution of Well-Being Dissected”, in A. Heston and R. E. Lipsey (eds.) *International and Interarea Comparisons of Income, Output, and Prices*. pp. 479-503.
- Talbot, Colin (2003), “Reinventing Government: Performance, Evaluation and Outcomes”. (<http://www.nira.go.jp/publ/seiken/pdf/talbot.pdf>). *NIRA Policy Research*. 16(5).
- Tallis, Ken (2005), “Social Indicator Systems – Foundations for Policy Design and Evaluation?” in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 198-216.
- Trewin, Dennis and Jon Hall (2005), “Measures of Australia’s Progress – A Case Study of a National Report Based on Key Economic, Social and Environment Indicators”, in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 88-101.
- United Nations (1975), *Towards a System of Social and Demographic Statistics*. New York: United Nations.
- van de Walle, Steven, Steven Van Roosebroek and Geert Bouckaert (2005), “Annex: Data on Trust in the Public Sector (Paper Prepared for the OECD Ministerial Meeting ‘Strengthening Trust in Government’ 28 November 2005)” (http://www.modernisinggovernment.nl/html/oeso/document_download.cfm?doc=A5091E33-F206-B8C5-1E0E14BF3BE25BD6.PDF&doc_name=5%20Analytical%20Annex%20E). Paris: OECD.
- Walker, David M. (2005), “Key National Indicator Systems: An Opportunity to Maximize National Progress and Strengthen Accountability”, in *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*. Paris: OECD. pp. 29-32.
- Wanna, John, Lotte Jensen and Jouke de Vries (2003), *Controlling Public Expenditure*. Northampton: Edward Elgar.
- Yackee, Susan Webb and David Lowery (2005), “Understanding Public Support for the US Bureaucracy”. *Public Management Review*. 7(4). pp. 515-36.
- Yang, Kaifeng and Marc Holzer (2006), “The Performance – Trust Link: Implications for Performance Measurement”. *Public Administration Review*. 66(1). pp. 114-26.

Zalm, Gerrit (2004), "Introductory Speech for the Meeting of Senior Budget Officials of Central and Eastern European Countries (10-11 November 2004, the Hague)". *OECD Journal on Budgeting*. 4(4). pp. 99-102.

Zouridis, Stavros (2003) "Satisfied Customers and Distrusting Citizens in the Netherlands (Paper to Be Presented on the EGPA Annual Conference, Oeiras 3-6 September 2003)" (http://soc.kuleuven.be/io/egpa/qual/lisbon/paper_potsdam_zouridis.pdf) EGPA. Oeiras, Portugal.