

The International Civil Service Effectiveness (InCiSE) Index

Technical Report

Users are free to copy, download and print InCiSE content and findings for their own use. Excerpts from InCiSE reports and publications may also be used, but must be appropriately referenced. Please cite any uses of this content as: "International Civil Service Effectiveness (InCiSE) Index, 2017".

The Index is comprised of a variety of other data sources. Reuse is not applicable to data subject to intellectual property rights of third parties. Please refer to other organisations' corresponding websites and data licensing restrictions to ensure compliance with their data limitations. Details of data sources can be found in Chapter 4 of this technical report. Every effort has been made to adhere to third party data regulations.

Please contact InCiSE@instituteorgovernment.org.uk for any queries.

Acknowledgements: Our thanks go to those who have given their time to shape this inaugural pilot publication. This has included substantial contributions made by academics, think-tanks, international organisations (particularly the OECD), and civil servants past and present. Our thanks also to those organisations who have allowed use of their data in the Index.

Contents

Chapter 1: Introduction	3
1.1 Why we need civil service effectiveness indicators	3
1.2 What InCiSE is aiming to do	3
1.3 Structure of this report	4
Chapter 2: Defining the civil service	5
2.1 Alternative definitions of the civil service	5
2.2 Scope of the InCiSE Index	11
Chapter 3: Defining the measurement framework	12
3.1 Purpose of the framework	12
3.2 Principles for the framework	12
3.3 Approach to assessing performance	13
3.4 The InCiSE framework	15
3.4.1 Definitions and justification of the attributes and functions	18
Chapter 4: Measuring against the framework	27
4.1 Data availability	27
4.2 Data underpinning the indicators	31
4.2.1 Functions	31
4.2.2 Attributes	44
Chapter 5: Index country coverage	53
5.1 Countries included in the Index	53
5.2: Imputation method	57
5.3: Imputation approach for indicators	59
5.3.1 Functions	59
5.3.2 Attributes	60
Chapter 6: Index results	62
6.1 Normalisation	62
6.2 Reporting of results	64

Chapter 7: Composite	68
7.1 Strengths and weaknesses of composites	68
7.2 The InCiSE composite	69
Chapter 8: Sensitivity Analysis	72
8.1 Introduction to sensitivity analysis	72
8.2 Sensitivity analysis 1: Adjusting for GDP per capita	72
8.3 Sensitivity analysis 2: Aggregation method	75
8.4 Sensitivity analysis 3: Exclusion of input data	78
8.4.1 Subjective assessment	78
8.4.2 Out of date data	81
8.4.3 Public sector proxy	83
8.5 Sensitivity analysis 4: Imputation method	85
Chapter 9: Next steps	87
9.1 Identified limitations of the Index	87
9.2 Priorities for the next 12 months	88
Appendices	90
Annex A: Data limitations	91
Annex B: Reference list	93
Annex C: Reader's Guide	96
C.1 Abbreviations	97
C.2 Country codes	98

Chapter 1: Introduction

.....

This chapter provides a brief background to the InCiSE Index, and sets out the structure of the report.

.....

1.1 Why we need civil service effectiveness indicators

An effective civil service can play an important role in determining a country's progress and prosperity. But what constitutes an "effective civil service" in the 21st Century? And once a consensus has been reached on defining this, how do civil service leaders know whether their organisations are effective, and in which areas are they performing more strongly than others? The InCiSE Index seeks to help answer these questions.

A comprehensive set of international indicators of civil service effectiveness does not currently exist. This subject area is also well recognised in academic, international and practitioner communities as a highly complex area for analysis. This is partly because of data limitations, different views on the definitions of "civil service" and "effectiveness", as well as the need to take account of country context factors when looking at performance issues. Nevertheless, there are many existing surveys and data collections available globally that can be pulled together to provide a view on civil service effectiveness on an annual basis.

The creation of a new and concise set of civil service effectiveness indicators would serve as:

- An accountability tool: allowing citizens, government officials and politicians to establish in a clear and concise way how well their civil service is performing.
- A performance improvement tool: enabling senior decision makers to see which countries perform best in which areas, and learn from them.

1.2 What InCiSE is aiming to do

InCiSE aims to define "effectiveness" more extensively than previously. It draws on a wide range of existing international data sources and brings together a set of indicators each measuring a different dimension of civil service effectiveness – and then produces a composite (an overall) score.

InCiSE has been developed following a literature review and in consultation with many experts, including academics from schools of government, think-tanks that monitor government effectiveness, international organisations, senior civil servants (past and present) and subject experts.

InCiSE has also been the subject of an independent, international peer review process. Three peer reviewers were selected: a senior academic from a major European governance research institute; a recently retired top civil servant with practical experience of civil service performance issues; and a senior governance expert with a distinguished career in two major international institutions. Country results were provided in an anonymous form for the review. Between them, the peer reviewers were asked to examine the measurement framework of ‘effectiveness’, the methodology and approach used to produce the indicators, and the data being used.

The International Civil Service Effectiveness Index (InCiSE) project is a collaboration between the Blavatnik School of Government and the Institute for Government. The project has been supported by the UK Civil Service, and is funded by the Open Society Foundations.

Despite some current limitations, the Index has already brought together a rich volume of data. The launch of this pilot Index will provide an opportunity to engage with a wide variety of stakeholders on a range of issues, stimulating discussions and feedback.

1.3 Structure of this report

This Technical Report on the InCiSE Index is intended to describe the methodology, data, and limitations of the approach used. The results of the Index can be found in the accompanying Main Report.

Given the 2017 publication is a pilot release, this technical paper will be updated as the Index is developed and as feedback is received. Should you have comment on this paper please email InCiSE@instituteforgovernment.org.uk.

Including this introductory chapter, there are 9 chapters of this report:

- **Chapter 2 – *Defining the civil service*** considers the object of assessment, the “civil service” and establishes a definition of the unit of measurement that allows for international comparison.
- **Chapter 3 – *Defining the measurement framework*** sets out a common approach for assessing the effectiveness of a civil service.
- **Chapter 4 – *Measuring against the framework*** describes the data included in the Index and the weighting of data.
- **Chapter 5 – *Index country coverage*** lists the countries included in the Index, and explains our approach to handling missing data.
- **Chapter 6 – *Index results*** describes how country scores are produced and how results are presented.
- **Chapter 7 – *Composite*** discusses the pros and cons of producing a composite and explains how the InCiSE composite Index is produced.
- **Chapter 8 – *Sensitivity analysis*** describes some of the uncertainties associated with the modelling process and the subjective choices, and the consequent impact on the Index results.
- **Chapter 9 – *Next steps*** highlights limitations of the pilot Index and areas for future consideration to develop the Index.

Chapter 2: Defining the civil service

.....

This chapter considers the object of assessment, the “civil service”, and establishes a definition of the unit of measurement that allows for international comparison.

.....

2.1 Alternative definitions of the civil service

The scope and responsibilities of the civil service varies across countries. Moreover, given that civil services across the world provide a wide-range of functions and are organised in a variety of ways, there is a need to establish exactly which parts are being assessed, and how.

In defining the civil service there are a number of approaches that can be taken. First, the civil service can be defined by function. Taking a narrow view this could refer to the central, “upstream” agencies which set policy direction and procedural regulation for the “downstream” agencies. This could include imposing regimes of transparency or accountability on them and marshalling budgetary and human resources for them. Those upstream agencies also produce policy outputs of their own for regulating economic behaviour and through changing tax and revenue, expenditure, state-owned-enterprises and investment policies. A broader functional perspective encompasses those agencies which are responsible for service delivery (although noting that services can be commissioned or funded, as well as provided by government).

In the 2001 version of the International Monetary Fund’s Government Financial Statistics Manual¹ and in Eurostat’s European System of Accounts², the classification of functions of government is used as an internationally-agreed statistical grouping of government expenses.

A second definition could take a national accounts perspective and again, taking a narrow view, identify civil service entities as those which are owned by government and whose financial reporting places them within the System of National Accounts (SNA) category of General Government. The OECD have undertaken work to establish a terminology and a new definition of what has been called the “public domain”.³ The new classification is now consistent with the system of national accounts (SNA).

1 IMF, 2001, Government Finance Statistics Manual (Washington: International Monetary Fund).

2 Eurostat, 2007, Manual on Sources and Methods for the Compilation of COFOG Statistics: Classification of the Functions of Government (COFOG), Eurostat Methodologies and Working Papers (Luxembourg: Office for Official Publications of the European Communities).

3 Pilichowski, E. and E. Turkisch (2008), Employment in Government in the Perspective of the Production Costs of Goods and Services in the Public Domain. OECD, Paris; OECD (2009), Measuring Government Activity, Paris, OECD.

The darkest shaded areas in Table 1 delineate those institutions which might be regarded as part of the civil service, namely the administrative units of central governments (and state governments, if a wider view of what constitutes the civil service is taken) and the typically very small (in terms of employment, not fiscal impact) social security funds at each level of government. Entities in government often lack a legal personality and are hence unable to own assets and incur liabilities which are reliant on central financial authorities. As Table 1 shows, for the SNA, General Government agencies include Ministries and departments in central government, along with state and local governments if taking a wider view. Since it is an accounting and not a management perspective, project implementation unit arrangements are covered to the extent that they are included in the budget (even if using consultants outside of any civil service regime). A less often noted but occasionally significant group of agencies within General Government encompasses those organisations that are largely funded and controlled by state or provincial government but not owned by government. In some countries, this includes schools and hospitals which are fully funded by government and where government determines the employment and financial regimes, but where there is an independent owner who is acting on the government's behalf, often a charity or other not-for-profit.

An even broader SNA perspective would pull in state-owned enterprises selling goods or services at an economically significant price, and financial and non-financial public corporations – to the extent that these are controlled by government as set out in the SNA.

Although this category would not be included as part of the civil service by practitioners or researchers, some have argued that the core public sector, broader than the civil service as usually understood, should include market producers and non-profits which are to all intents and purposes entirely reliant on the public sector as monopsony customer, and private enterprises which have been granted a distinctive and statutorily privileged market position.

This is on the basis that while there may be little direct public funding for these agencies there is often an assumption that government will meet any implicit contingent liability that arises, with the probability that government would, in the event of major operational failure, underwrite the debts of these entities. They are thus underpinned by an implicit guarantee.

Table 1: The public sector from a National Accounts perspective

Institutional domain				How transactions are recorded in the national accounts	Examples
The “public domain” ⁴	Public sector	General government ⁵	Central government	Administrative units in central government	Ministries and departments in central government
				All non-market non-profit institutions that are controlled and more than 50 percent financed by central government units	Schools, hospitals, etc. that are largely funded and controlled by central government but not owned by government
			State governments	Administrative units in state government	Departments in states, provinces
				All non-market non-profit institutions that are controlled and more than 50 percent financed by state government units	Schools, hospitals, etc. that are largely funded and controlled by state or provincial government but not owned by government
			Local governments	Administrative units in local government	Departments in counties, municipalities
				All non-market non-profit institutions that are controlled and more than 50 percent financed by local government units	Schools, hospitals, etc. that are largely funded and controlled by local government but not owned by government
			Social security funds	All social security funds at each level of government	Health fund, unemployment fund, pension fund
		Other public sector		Market producers, controlled by government, selling goods or services at an economically significant price (“public enterprises”): <ul style="list-style-type: none"> • Public financial (quasi-) corporations • Public non-financial (quasi-) corporations As defined by S.11 and S.12 in the SNA.	Publicly owned banks Publicly owned harbors, airports

Table 1: The public sector from a National Accounts perspective			
	Private sector in the public domain	Market producers, whose indirect public funding comprises more than 50 percent of total revenue: <ul style="list-style-type: none"> • Non-profit institutions • Profit institutions As defined by S.11, S.12 in the SNA	Profit or non-profit private hospitals accessible to publicly insured clients
		Non-profit institutions serving households, financed more than 50 percent by government, but not controlled by government: <ul style="list-style-type: none"> • Non-profit institutions serving households As defined by S.15 in the SNA	Schools, hospitals, etc. that are largely funded by government but not owned nor controlled by government
		Private enterprises with a distinctive and statutorily privileged market position: <ul style="list-style-type: none"> • Private sector utilities licensed to operate in very limited markets (water, energy, sewage, waste disposal, post, but not telecommunication) Legal monopolies As defined by S.11 in the SNA	Energy companies, local public transport companies National train company

Source: Developed largely from Pilichowski and Turkisch⁶ with OECD⁷

Note: The darkest shaded area highlights those institutions which might be regarded as part of the civil service, namely the administrative units of central governments. The medium shaded area covers those categories, in addition to the darkest shaded area, considered part of General Government under the SNA. The lighter shaded area covers state-owned enterprises selling goods or services at an economically significant price, and financial and non-financial public corporations – to the extent that these are controlled by government as set out in the SNA.

⁴ This description is devised for this purpose and is not a recognized SNA term.

⁵ As defined by section S.13 in the System of National Accounts).

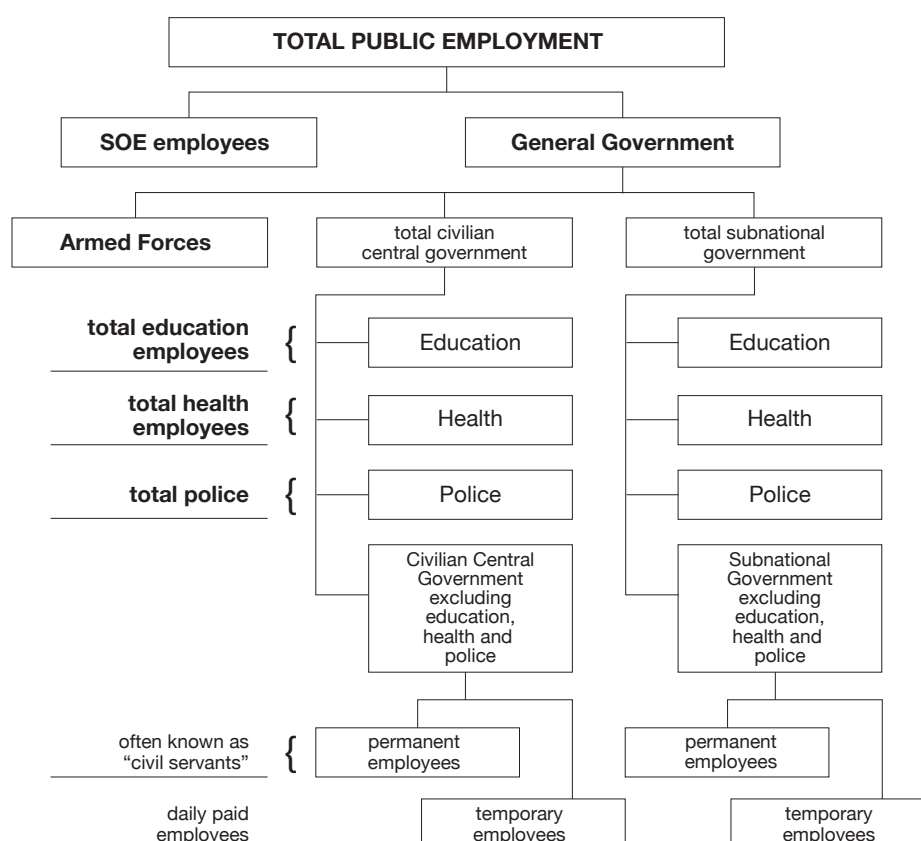
⁶ Pilichowski, E. and E. Turkisch (2008). Employment in Government in the Perspective of the Production Costs of Goods and Services in the Public Domain. OECD, Paris.

⁷ OECD (2009), Measuring Government Activity, Paris, OECD.

The final approach explored was of a civil service determined on the basis of employment regimes (Figure 1). The most stringent definition limits the selected entities to those which are required to hire most employees under the civil service law, excluding those using other legal employment regimes.

This can present practical difficulties as, although generally the civil service is understood to constitute a distinct body of staff within the public sector, staff that are commonly referred to as “civil servants” do not always have legally distinct employment contracts. When it exists formally, the essence of civil servant status is that the legal basis for employment – the laws and regulations that shape the nature of employment contracts – is different from that found elsewhere in the economy as defined by the general labour law. It also is generally different from that found elsewhere in the public sector, such as in the health or education sectors or in state-owned enterprises.

Figure 1: Typical legal employment regimes in general government



Source: World Bank 2007⁸

8 World Bank (2007). The World Bank's Administrative and Civil Service Reform Website. World Bank.

Historically, civil service employment was not a formal agreement between two equal parties, but rather a decision of the State, with appointments made by an authorised public institution in accordance with the civil service law, with tight regulations that shape the appointment process, and with many constraints on dismissal. Today, civil service employment tends to share some features that are typical of a voluntary arrangement between an employer and employee in the private sector. There can be other arrangements, particularly in the health, education, military and police sectors, that provide civil service-like protections and responsibilities for public employees who are not, in fact, civil servants. Subnational government employment is considered a separate, legally defined civil service in many countries.^{9,10} A number of organisations including the International Labour Organisations and the OECD have considered the legal basis for civil servant status.¹¹

The three methods offered here for defining the centre of government based on functions, national accounts or employment regimes omit other less useful criteria that could be applied, such as:

- A legal definition of the centre of government could comprise those entities that are created under the authority of the constitution or by public law.¹²
- Alternatively, the centre of government could be defined as the set of all entities that operate directly under the authority of the political executive.
- Ownership is a further set of ideas for defining the institutional units that belong to the public sector, drawing on accounting standards for criteria concerning which body has the power to govern the financial and operating policies of another entity.¹³

The problem with each of these alternative conceptions is that they result in a large and unwieldy group of agencies which, most importantly, is rarely consistent with professional or other working definitions.

9 Gow, J. I. and M. d. C. Pardo (1999), 'Comparing Different Civil Services', *International Review of Administrative Sciences*, 65 (4), 527-550.

10 Cardona, F. (2000). *Scope of Civil Services in European Countries: Trends and Developments*. Seminar at the European Institute of Public Administration. Maastricht, Sigma/OECD

11 SIGMA (1996a), *Civil Service Legislation Contents Checklist – Sigma Paper No. 5*, Paris, OECD.

12 Gill, D. (2002), 'Signposting the Zoo – from Agencification to a More Principled Choice of Government Organisational Forms', *OECD Journal on Budgeting*, 2 (1), 27-80

13 Lienert, I. (2009). *Where Does the Public Sector End and the Private Sector Begin?*. IMF, Washington DC.

2.2 Definition in the InCiSE Index

The approach taken for the InCiSE Index is to define the scope of ‘civil services’ by outlining and measuring performance on the core functions of civil services; the parts which can generally be classified as civil service in every country. This approach leads to a focus on (i) functions which deliver services or affect citizens directly and (ii) public management and policy functions carried out at the centre of government.

Table 2 gives more detail about what is included and excluded in the InCiSE Index:

Table 2: Public sector areas included in the InCiSE framework	
Parts/functions of the public sector	Degree of inclusion in measurement framework
Civil service functions which deliver services to citizens directly (e.g. tax and social security administration at the central/federal level).	A primary focus of the InCiSE Index with each function assessed in an individual indicator.
Central, public administration-type civil service functions (e.g. fiscal management, policy making, regulation).	A primary focus of the Index with each function assessed in an individual indicator.
The ‘mission support’ functions which support these core service delivery and central administration-type functions.	A primary focus of the Index.
The parts of civil services which direct and support the wider public sector on specific policy areas (e.g. Health Ministry, Education Ministry, Environment Ministry), but may not deliver public services to citizens directly.	Performance captured by indicators on the functions above which cut across most of the policy areas governments typically deal with. For example regulation, and policy making more generally, cut across a number of policy areas including health, education, environment, and so on. Performance on specific policy areas not individually assessed because: <ol style="list-style-type: none"> 1. Data is unlikely to exist which sufficiently isolates the ‘oversight’ specific ministries provide over different policy areas from the service delivery the wider public sector provides in these areas; and 2. Governments in all countries deal with a large number of policy areas and there is a need to limit the scope of the measurement framework.
The wider public sector itself (e.g. workers in public hospitals, schools and police departments).	Out of scope, although occasionally public sector data is used to proxy for civil service performance.
Local government.	Out of scope. Whilst public servants working in sub-national governments may in some countries be technically classified as ‘civil servants’, the scope of this Index is primarily concerned with the civil service at central government level.

The unit of analysis of interest for the InCiSE Index is the civil service, not the public sector more generally. Having said that, isolating civil service performance with currently available data is difficult, particularly given the varying sizes and shapes of civil services internationally. This issue is discussed further in Chapter 9 and Annex A which highlight some limitations of the Index.

Chapter 3: Defining the measurement framework

Chapter 3 considers the purpose of designing a framework for the Index, the principles specified, the framework derived and the justification for its components.

3.1 Purpose of the framework

The purpose of the framework is to define a common approach for assessing the effectiveness of a civil service, in a way which could realistically enable international data to be collected to measure against it. Whilst there are many alternative ways one could define civil service effectiveness, the framework outlined here, and the themes and dimensions therein, is informed by evidence such that if a civil service were to score highly against it, it is reasonable to conclude that the civil service would be high-performing relative to its international counterparts.

Our approach to deriving a common framework was to:

- Specify and adhere to a set of principles to inform the development of the framework (3.2);
- Draw on evidence (3.4.1) to identify key features of a draft framework (chapter 4), which was then extensively tested through consultation;
- Build on existing indicators and data where possible, while striving to develop a more comprehensive framework capturing all aspects of civil service effectiveness;
- Refine the framework through consultation with a number of experts, including academics, think-tanks, international organisations, civil servants (past and present) and subject experts.

This framework, and the resultant Index, provides civil services across the world with a powerful tool for (i) accountability, allowing citizens, commentators and ministers to establish how well their country's civil service is performing and (ii) performance improvement, for example by allowing senior decision makers to see which countries perform best in which areas and learn from them. This learning potential would be enhanced if the Index could connect with and inform existing learning and research initiatives in the field.

3.2 Principles for the framework

Prior to developing the framework we specified a list of general principles for our indicators. These principles are shown in Table 3.

Table 3: Principles for the InCiSE framework¹⁴

- **Coherent** framework of what are the key elements and drivers of an effective public administration.
- **Comprehensive**, meaning that it covers all relevant aspects and drivers of the performance of public administration.
- **Actionable**, offering genuine policy insights into what drives excellent public administration performance that can be implemented.
- **Transparent** in its methodology and assessment process to ensure credibility, robustness and replicability.
- **Feasible** to collect from a large group of countries at reasonable cost.
- **Replicable** to allow for both time and cross country comparison.

The pilot Index provides actionable insights, and is built on a coherent framework that is transparent, feasible and replicable:

- The indicator draws on literature to determine the key elements of an effective civil service and builds on the work of other indicators and data collection.
- All source data is specified and the methodology to derive the Index set out in this report, promoting transparency and replicability.
- The Index is feasible to produce for a large group of countries on a regular basis, largely due to its draw on existing data, use of imputation for missing scores, and normalisation of data to allow comparison.

Section 3.4 sets out the range of functions and attributes which the Index aims to measure. Work will be required post pilot publication to develop the framework further. To this end we encourage feedback as to how the indicator framework can be further strengthened.

3.3 Approach to assessing performance

The standard approach for assessing civil service effectiveness would be to think in terms of inputs, outputs and outcomes. This is the approach taken by some recent studies¹⁵. However, when looking specifically at civil services and the public administration-type functions they provide, this approach seems less attractive. While output and outcome measures may seem to cut through the conceptual uncertainty and simply ask what got done, in practice they are likely to be problematic for three reasons:

1. They can be affected by exogenous factors, making it difficult to isolate the contribution of the civil service.
2. Measuring output is itself problematic methodologically; difficulties include defining the units of output and obtaining information as to the value of this output, due to public sector output being provided for free or for an economically insignificant amount.
3. Normative and procedural concerns (i.e. how the output or outcome was achieved) are relevant to effectiveness.¹⁶

¹⁴ A number of these principles were also used to develop World Bank ISMPS project <http://www.worldbank.org/en/topic/governance/brief/indicators-of-the-strength-of-public-management-systems>

¹⁵ Goderis et al (2015), 'Public Sector achievement in 36 countries: a comparative assessment of inputs, outputs and outcomes', The Netherlands Institute for Social Research.

¹⁶ Wieland J. (2014) Governance Ethics: Global value creation, economic organization and normativity, Springer International Publishing, p205

Ascribing causality between public administration performance and outcomes is too difficult: “It is reasonable to conclude that because causality in achieving public policy outcomes is notoriously hard to assign, the quality of outputs is very hard to determine.¹⁷” Without the certainty that there is a causal relationship between a government output and the achievement of a stated policy objective, we cannot know the quality of the output because quality, in this context, is a measure of how well the output is contributing to the objective.

Given this difficulty, the preferred approach here is to focus on the effectiveness of the procedures within the civil service which (often indirectly) affect those outcomes. The approach deployed is therefore more process and output focused. Beyond just feasibility, one could argue an advantage of process-based indicators is that they are more instructive for potential performance improvements as it is processes which are ultimately changed to increase effectiveness. We acknowledge the problem with all procedural definitions of effectiveness that the procedures, however defined, may not actually correlate with the positive outcomes.

However, certain procedural measures remain at the core of any measure of effectiveness, both where there is evidence to support the relationship between such procedures and positive outcomes, and because procedures may have a benefit in themselves. For example, meritocracy of recruitment procedures in the civil service are important because there is a considerable evidence base to support the relationship between such procedures and outcomes associated with an effective civil service. However, the extent to which recruitment processes reward merit is also important for the principle of fairness which is valued in itself.

Aggregate inputs, such as the total human and financial resource put in, are not measured at this stage. It may be important to compare performance on the framework against those aggregate inputs (for example this could be done in the form of a civil service efficiency ratio; the ratio of the score for an indicator to the overall inputs). We have however examined the sensitivity of the Index when taking into account the relative wealth of countries as measured by GDP per capita (see Chapter 9 for more information).

17 Holt J. and Manning N. (2014), ‘Fukuyama is right about measuring state quality: now what?’, *Governance: An International Journal of Policy, Administration, and Institutions*, 27(4).

Figure 2 below shows what is captured in the indicator set.

The logic model highlights the focus on the effectiveness of the procedures within the civil service rather than outcomes, leading to the more process and output focused approach deployed.

3.4 The InCiSE framework

The InCiSE framework starts by defining the core characteristics of an effective civil service. To do this, it assesses effectiveness on the basis of two inter-related dimensions, 1) the delivery of its core functions and 2) an underlying set of attributes which are important drivers of effectiveness across all parts of the civil service.

Functions

Civil services across the world are of different shapes and sizes, but there are certain core functions which all of them deliver. On one side, they deliver a set of central executive functions for ministers. These may help to formulate policy for the country (the effects of which are borne by citizens). On the other side, civil services interact more directly with citizens through the delivery of services such as tax administration. As with

every organisation, mission support functions support these core external functions (on both sides). By looking across all 3 types of function, the aim is to measure how well civil services deliver the core elements of their roles.

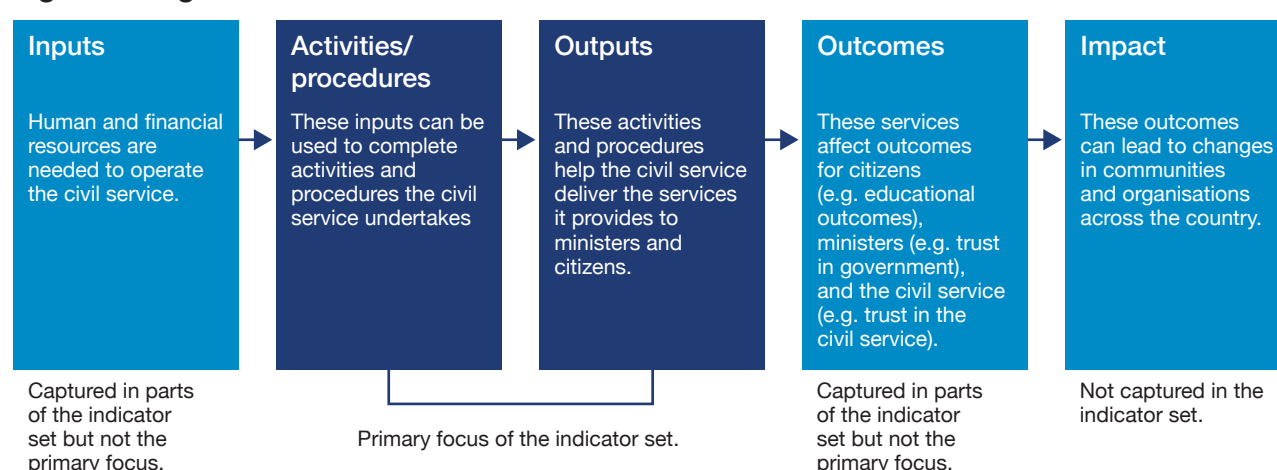
Figure 3 shows the functions currently included in the InCiSE framework.

Section 3.4.1 sets out the definition of each of these functions, along with a justification for their inclusion in the framework.

Attributes

Every civil service also has an underlying set of attributes which are important drivers of how effectively they deliver core functions. These attributes should apply to all parts of the civil service; they are not specific to particular parts or functions. The inclusion of attributes in the framework is based on both a normative and positive judgement: civil services should aim to cultivate and demonstrate these attributes, as they are commonly (but not necessarily universally) understood as aspects of best practice, and the included attributes should generally be determinants of performance across all functions. Figure 4 shows those attributes currently included in the InCiSE framework.

Figure 2: Logic model of the civil service



Section 3.4.1 sets out the definition of each of these attributes, along with a justification for their inclusion in the framework.

Figure 3: InCiSE Core Functions












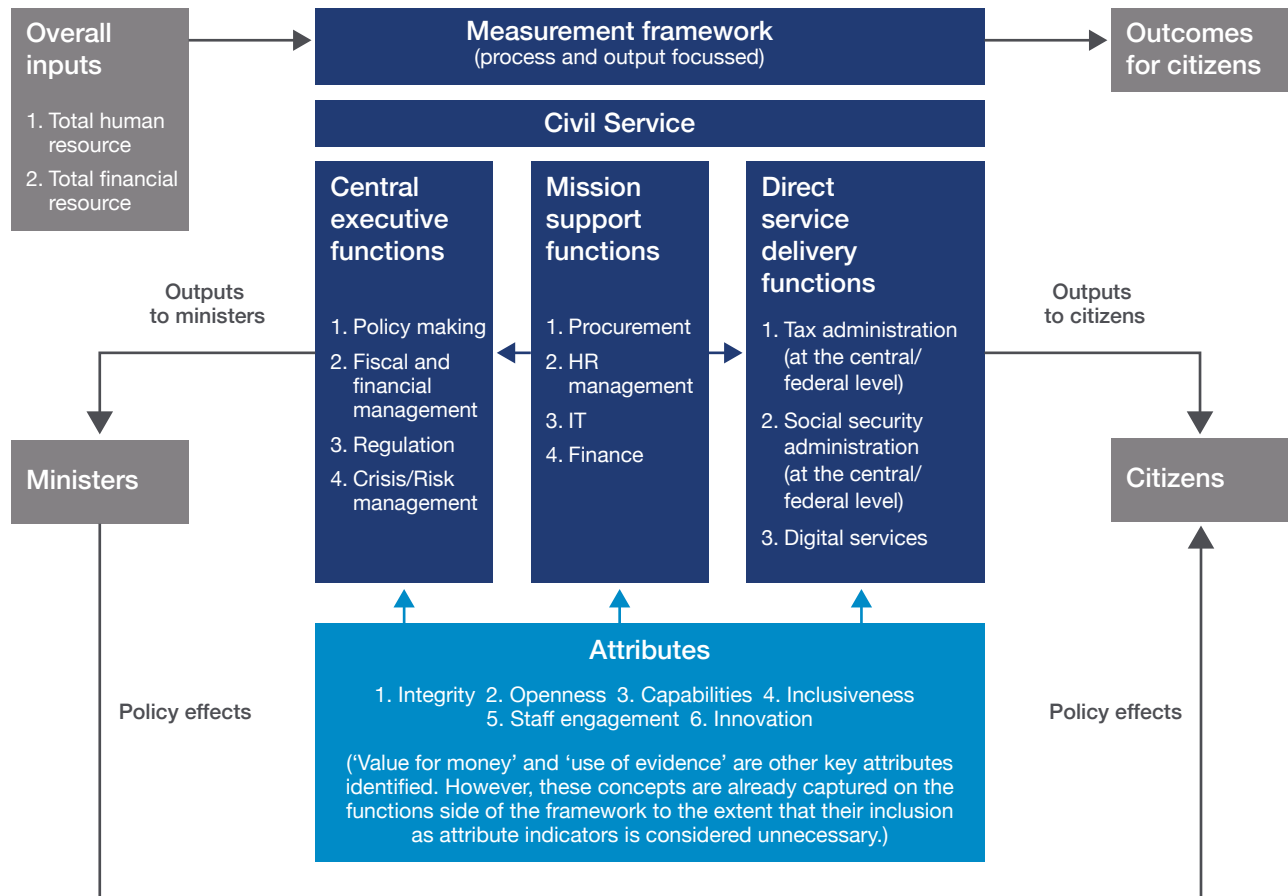
Central executive		Mission support		Direct service delivery	
	Policy making: The quality of the policy making process, including how policy is developed and coordinated across government and monitored during implementation.		Procurement: The extent to which the procurement process is efficient, competitive, fair, and pursues value for money.		Tax administration: The efficiency and effectiveness of tax collection (at the central/federal level).
	Fiscal and financial management: The quality of the budgeting process and the extent to which spending decisions are informed through economic appraisal and evaluation.		HR management: The meritocracy of recruitment and extent to which civil servants are effectively attracted, managed and developed.		Social security administration: The efficiency and effectiveness of social security administration (at the central/federal level).
	Regulation: The extent and quality of regulatory impact assessments and the degree of stakeholder engagement involved in them.		Information technology: The extent to which civil servants have the digital tools to work efficiently.		Digital services: The user-centricity, transparency and cross-border mobility of digitally-provided public services and the availability of 'key enablers'.
	Crisis/risk management: The effectiveness with which the government engages the whole of society to better assess, prevent, respond to and recover from the effects of extreme events.		Finance: The extent to which operations are supported by well-managed, efficient finance systems, particularly on the alignment of finance with the business strategy and the level of civil servant satisfaction with finance support.		

Figure 4: InCiSE Attributes

	1. Integrity: The extent to which civil servants behave with integrity, make decisions impartially and fairly, and strive to serve both citizens and ministers.		4. Inclusiveness: The extent to which the civil service is representative of the citizens it serves.
	2. Openness: The regular practise and degree of consultation with citizens to help guide the decisions we make and extent of transparency in our decision-making.		5. Staff engagement: Staff levels of pride, attachment and motivation to work for their organisation.
	3. Capabilities: The extent to which the workforce has the right mix of skills.		6. Innovation: The degree to which new ideas, policies, and ways of operating are able to freely develop.

The attributes and functions identified are brought together to form the InCiSE framework shown in Figure 5.

Figure 5: InCiSE Index framework



This framework shows the attributes which drive and support the successful delivery of the civil service across the three categories of functions. The ‘mission support’ functions (Procurement, HR Management, IT and Finance) underpin both the ‘central executive functions’ providing outputs to ministers and the ‘direct service delivery’ functions providing outputs to the public. For the reasons described in section 3.3, citizen outcomes are not currently included directly in the framework, although it is important to bear in mind that all procedures and outputs within the framework are delivered with a view to influence them.

3.4.1 Definitions and justification of the attributes and functions

In this section we set out the definition for each attribute and function specified in Figures 2 and 3, and the justification for its inclusion in the framework from our literature review. A bibliography of references in this section can be found in Annex C.

Policy Making

Definition: The quality of the policy making process, including how policy is developed and coordinated across government and monitored during implementation.

Justification: Policy making remains a central role of a civil service and the quality of evidence and appraisal are central to the success of policy. Kaufmann et al. outline three functions of good governance, including ‘the capacity of government to effectively formulate and implement sound policies’ (1999). Policymakers need to ‘receive rigorous analyses of comprehensive background information and evidence, and of the options for actions,’ according to the Office for Public Management (OPM) and the Chartered Institute of Public Finance and Accountability (CIPFA) (2004). This paper also advises that ‘good quality information and clear, objective advice can significantly reduce the risk of taking decisions that fail to achieve their objectives or have serious unintended consequences’. The necessity of sound evidence and reliability in policymaking is echoed by Bovaird and Löffler: ‘The choice of a particular policy direction should be informed by existing evidence on what has been tried elsewhere and whether it has been demonstrated to deliver the desired benefits,’ and ‘evidence can be used both to facilitate accountability and to promote improvement in policy-making, programme development and service delivery’ (2003).

Fiscal and Financial Management

Definition: The quality of the budgeting process and the extent to which spending decisions are informed through economic appraisal and evaluation.

Justification: Fiscal and financial management is an important measure of every system of public administration. The Indicators of the Strength of Public Management Systems (ISPMS) from the World Bank state 'Public sector management arrangements must also encourage fiscal and institutional sustainability as less tangible but equally critical outcomes' and 'Reforms of budgetary and financial management systems... are often crucial for development outcomes' (2012). Holt and Manning also consider that 'public administration practitioners break down the functioning of the central agencies into five management systems,' including fiscal and financial management which is made up of: 'planning and budgeting; financial management; and accounting, fiscal reporting and audit.' The OECD's recommendation paper on budgetary governance sets out ten principles for good budgetary governance which include 'ensur[ing] that performance, evaluation and value for money are integral to the budget process' and 'manag[ing] budgets within clear, credible and predictable limits for fiscal policy,' (OECD, 2015a).

Regulation

Definition: The extent and quality of regulatory impact assessments and the degree of stakeholder engagement involved in them.

Justification: The appropriate appraisal and evaluation of regulatory changes accompanied by sufficient stakeholder engagement is crucial to ensuring that any introductions are fully considered and fair, involving various stakeholders. This scrutiny is endorsed by many; the OECD, for instance, 'recognis[es] that regulations are one of the key levers by which governments act to promote economic prosperity, enhance welfare and pursue the public interest,' and that 'well-designed regulations can generate significant social and economic benefits which outweigh the costs of regulation, and contribute to social well-being,' (2012). The International Monetary Fund (IMF) acknowledges the importance of regulatory frameworks to successful governance: 'From the perspective of the IMF, countries with good governance have strong legal and regulatory frameworks in place,' (2016). Additionally, in promoting best practice, '[the] Regulatory Impact Analysis (RIA) is a multiple stakeholder assessment of the economic, environmental and social impact of regulations. The OECD and European Union have strongly promoted this evidence-based approach towards legislation' (Bovaird and Löffler, 2003).

Crisis/Risk Management

Definition: The effectiveness with which the government engages the whole of society to better assess, prevent, respond to and recover from the effects of extreme events.

Justification: The OECD Strategic Crisis Management report highlights crisis management as central to government's role and a 'fundamental element of good governance' (Baubion, 2013). New Zealand measures their public sector outputs against eight identified categories, of which two encompass contingent and emergency capabilities: 'so that an adequate response will be available in time to minimise loss, damage or injury,' cited by Van Dooren et al. (2006). Studies have shown that credibility and trust in governments to deal with crises is vital both to reassure and encourage support from the private sector and general public, as outlined by Christensen et al. (2011). In the OECD's recommendation of the council on the governance of critical risks, it is 'recognising that effective risk governance is a means of maintaining or achieving national competitive advantage against a backdrop of numerous geopolitical, environmental, societal and economic uncertainties as it represents an opportunity to invest in safer and better lives for the future' and 'recognising that citizens and businesses expect governments to be prepared for a wide range of possible crises and global shocks and to handle them effectively should they arise,' (OECD, 2014).

Procurement

Definition: The extent to which the procurement process is efficient, competitive, fair, and pursues value for money.

Justification: 'Government procurement accounts for an average of 15 per cent or more of a country's GDP,' (World Trade Organisation, 2015). As procurement makes up such a large proportion of countries' GDP, it must be managed appropriately. Effective procurement management can streamline contracts and reduce outgoings, contributing to improved efficiencies in civil services. Phillips et al. also observe that: 'effective procurement practices provide governments with a means of bringing about social, environmental and economic reform' (2007). On public procurement, the World Bank states it 'is a key variable in determining development outcomes and, when carried out in an efficient and transparent manner, it can play a strategic role in delivering more effective public services. It can also act as a powerful tool for development with profoundly positive repercussions for both good governance and more rapid and inclusive growth' (2016).

Human Resource Management

Definition: The meritocracy of recruitment and extent to which civil servants are effectively attracted, managed and developed.

Justification: ‘The public sector is very labour intensive – around 70 per cent of the budgets of most public organisations are spent on staff’ (Bovaird and Löffler, 2003), so good HR management is key to the successful functioning of an exemplary civil service. Performance management can help to create incentives for personal development in the civil service. Fukuyama (2013) recognises that recruitment and reward ‘remain at the core of any measure of quality of governance... whether bureaucrats are recruited and promoted on the basis of merit’. ‘If the HR policies are not right then public organisations will not attract the human resources they need to perform the functions of government and deliver the services that government has promised the electorate’ (Bovaird and Löffler, 2003).

Information Technology

Definition: The extent to which civil servants have the digital tools to work efficiently.

Justification: It is important for the public sector to keep up with IT developments in order to maintain optimum efficiency. Advanced IT can also allow for improvements in flexible working patterns for civil servants. Madzova et al. make the point that improved IT ‘can enhance the speed and efficiency of operations, by streamlining processes, lowering costs, improving research capacities and record keeping,’ (2013), highlighting the cost saving which could be achieved. Magno and Serafica go further to identify three ways in which IT promotes good governance: ‘(1) by increasing transparency, information, and accountability; (2) by facilitating accurate decision-making and public participation; and (3) by enhancing the efficient delivery of public goods and services,’ (2001). However, skills must be in place to harness this capability, and Fukuyama (2013) acknowledges the technical expertise of civil servants ‘remain[s] at the core of any measure of quality of governance’.

Finance

Definition: The extent to which operations are supported by well-managed, efficient finance systems, particularly on the alignment of finance with the business strategy and the level of civil servant satisfaction with finance support.

Justification: Managing government finance in the context of business operations contributes towards value for money for the taxpayer to the same extent as fiscal spending. ‘Taxpayers are entitled to receive assurance that Public Administrations take due care in managing funds,’ and ‘the issue of adequate internal control is at the heart of sound financial management of the national budget.’ (European Commission, 2006). Bouchard and McCrae point out that the ‘Control of money and budgets is a core role of any finance function,’ and they find that ‘previous work on improving decision making has highlighted the importance of financial leadership and strengthened performance management at the top of government departments’ (2013).

Tax Administration

Definition: The efficiency and effectiveness of tax collection (at the central/federal level).

Justification: Effective tax systems can be viewed as a critical building block for increased domestic resource mobilisation which is essential for civil service effectiveness and good governance. ‘Successful tax extraction provides resources that enable the government to operate in other domains,’ Fukuyama highlights ‘it is a necessary function of all states, and one for which considerable data exist’ (2013). The role of tax administration as the basis of government operations is made clear by the OECD, ‘Strong tax administrations and sound public financial management help maximise the domestic resources that are necessary for government to function, to sustain social safety nets, to maintain long-term fiscal sustainability, and to free up fiscal space for pursuing socio-economics objectives,’ (n.d.). Although priorities and circumstances vary widely across countries, the drive to elevate the collective standard of tax administration is of great importance. Holt and Manning highlight the importance of tax administration in measuring the effectiveness of public administration and it is one of the key functions highlighted by the World Bank Indicators of the Strength of Public Management Systems (2012).

Social Security Administration

Definition: The efficiency and effectiveness of social security administration (at the central/federal level).

Justification: Social security administration plays an important role in civil services. The OECD study *Towards Better Measurement of Government*, for example, highlights the importance of social security administration. Chalam says that ‘Social security is a human right as well as a social and economic necessity. All successful societies and economies have employed developmental strategies where social security systems played an important role to alleviate poverty and provide economic security that helps people to cope with life’s major risks or the need to quickly adapt to changing economic, political, demographic and societal circumstances’ (2014). McKinnon (2011) supports this view: ‘Several legal instruments adopted by the United Nations recognize social security as a basic human right. The State has the responsibility to create the enabling environment that would allow citizens to exercise this right,’ as well as pointing out ‘the role of social security as a requisite economic and social stabilizer’. In summary, ‘Social security systems play a central role in the efforts of every country to promote and ensure the social and economic well-being of its citizens’ (McKinnon, 2011).

Digital Services

Definition: The user-centricity, transparency and cross-border mobility of digitally-provided public services and the availability of ‘key enablers’.

Justification: A changing world and digital environment provide the impetus for a civil service to ensure modernity and remain user-centric for the public. In doing so, efficiencies should be achieved to enable cost savings in processes while also allowing for further accessibility of services. The OECD has supported this view of potential benefits: ‘ICT is increasingly used to support broader public sector development objectives... by changing service delivery approaches by creating personalised, high quality services to users, thereby increasing user satisfaction and effective service delivery; facilitating major work organisation and management changes creating back-office coherence and efficiency gains; increasing transparency of government activities, and increasing citizen engagement.’ (Lonti and Woods, 2008).

Integrity

Definition: The extent to which civil servants behave with integrity, make decisions impartially and fairly, and strive to serve both citizens and ministers.

Justification: Integrity is one of the core values associated with a civil service. The International Civil Service Commission highlights the importance of integrity to the work of United Nations (UN) common systems staff: ‘The concept of integrity... embraces all aspects of behaviour of an international civil servant, including... honesty, truthfulness, impartiality and incorruptibility. These qualities are as basic as those of competence and efficiency.’ (Civil Service Commission, 2002). The World Bank states: ‘A well-performing civil service resists petty corruption and provides the staff for many of the institutions that protect integrity in government’ (n.d.) while setting out the role of the World Bank in helping countries combat corruption. Indeed, the Charter of the United Nations lists integrity as key to the qualities employees must show: ‘The paramount consideration in the employment of the staff and in the determination of the conditions of service shall be the necessity of securing the highest standards of efficiency, competence, and integrity.’ (United Nations, 1945). This paper also outlines the need for impartiality as key to independence of a good civil service as well as Kaufmann et al. (1999) and Huther and Shah (1999). Numerous studies aiming to establish good governance have utilised similar metrics in their analyses, for instance Muiithi et al. at the LSE (2015). Therefore, the inclusion of integrity in the InCiSE is deemed necessary and crucial for the assessment of an effective civil service.

Openness

Definition: The regular practise and degree of consultation with citizens to help guide the decisions we make and extent of transparency in our decision-making.

Justification: The need for transparency within a civil service is imperative for the public to trust and feel empowered to hold the government accountable for their actions, whilst reducing corruption. The World Bank ‘supports efforts to encourage open and transparent government’ in their advice for helping governments to strengthen institutions against corruption (The World Bank, n.d.). The United Nations outlines the need for transparency and accountability in governance; ‘[this] implies a proactive effort to make information accessible to citizens’ and it is ‘one indicator of a government that is citizen-focussed and service-oriented’. (United Nations, 1999). Graham et al. also make reference to the United Nations Development Program’s five principles of good governance, in which transparency is identified as a key characteristic. The OECD summarises the importance of openness in that ‘access to information and proactive transparency help build citizens’ trust in government.’ (n.d.).

Capabilities

Definition: The extent to which the workforce has the right mix of skills.

Justification: The need for a variety of certain strong skills is vital for the successful operation of any organisation, civil services included. ‘Public service organisations need people with the right skills to direct and control them effectively’ (OPM and CIPFA, 2004). The standards for good governance set out by the Office for Public Management (OPM) and the Chartered Institute of Public Finance and Accountability (CIPFA) which outline seven principles for people in public life (known as the Nolan principles), include leadership as a core skill. It goes on to list necessary skills as ‘the ability to scrutinise and challenge information... including skills in financial management and the ability to recognise when outside expert advice is needed,’ (2004). Fukuyama acknowledges the importance of educational attainment of civil servants: ‘Another critical measure of capacity is the level of education and professionalization of government officials,’ along with the importance of digital capability: ‘what level of technical expertise they are required to possess’ (2013).

Inclusiveness

Definition: The extent to which the civil service is representative of the citizens it serves.

Justification: A modern civil service should be representative of the public it stands to serve. In order to do so, institutions must be inclusive in nature. ‘Governments are increasingly concerned about the importance of diversity in public institutions, to ensure that the needs, aspirations and experiences of a diverse range of citizens are reflected in the decision-making process,’ (OECD, 2015b). In their Post-2015 Millennium Development Goal reflections, the OECD outlines the greater success felt by inclusive public bodies: ‘Inclusive governments and an active civil society put forward more responsive, equitable policies’ and that these ‘build trust in government and help create... public services that are better suited to diverse needs,’ (n.d.). The guiding principles to the international civil service, set out by the International Civil Service Commission, support the claim that civil servants must ‘respect the dignity, worth and equality of all people’ and have: ‘a willingness to work without bias with persons of all nationalities, religions and cultures’ (2002). The OECD Government at a Glance report makes the point that opinion on the groups in need of representation in public administration has widened ‘and now includes a range of dimensions such as women; racial, ethnic and religious minorities; the poor; the elderly; the disabled; and other minority groups such as indigenous populations,’ and goes on to say that ‘a more representative public administration can better access previously overlooked knowledge, networks and perspectives for improved policy development and implementation,’ (2015b). A paper by the Office for Public Management (OPM) and the Chartered Institute of Public Finance and Accountability (CIPFA) highlights the potential benefits of this view; ‘To enrich governance deliberations by bringing together a group of people with different backgrounds – governing bodies need to recruit governors from different parts of society. Public trust and confidence in governance will increase if governance... [is] done by a diverse group of people who reflect the community.’ (2004).

Staff Engagement

Definition: Staff levels of pride, attachment and motivation to work for their organisation.

Justification: Staff who feel that their roles are making a real contribution, as well as a genuine interest and pride in their work will be more motivated and engaged. The OECD defines engagement as: ‘Engaged employees are those who are “committed to their organisation’s goals and values, motivated to contribute to organisational success, and are able at the same time to enhance their own sense of wellbeing”’ (2015d). The paper goes on to state that ‘engaged employees are critical to successfully manage change in the public administration, to enhance service orientation and to ensure customer satisfaction,’ (OECD, 2015d). A report for the UK Government by MacLeod and Clarke points out that ‘Employee engagement strategies enable people to be the best they can at work, recognising that this can only happen if they feel respected, involved, heard, well led and valued by those they work for and with’ (2009). The paper also highlights the wider benefits associated: ‘they are motivated and able to give of their best to help it succeed – and from that flows a series of tangible benefits for organisation and individual alike.’ (MacLeod and Clarke, 2009).

Innovation

Definition: The degree to which new ideas, policies, and ways of operating are able to freely develop.

Justification: The opportunities for government innovation are vast and numerous; ‘Governments are operating in a new landscape. The public sector faces economic, social and environmental challenges; technology is revolutionizing how citizens interact with government; individuals and organisations across society are forming new kinds of partnerships; and citizens are more informed and connected than ever. Together these factors create opportunities for new ways of thinking about government and how it works’ (OECD, 2015c). The necessity for innovation has been highlighted by Nesta: ‘For public sectors to become more adept at innovation they need to treat it with the same seriousness they deal with handling risk, financial controls or regulatory enforcement’ (Mulgan, 2014).

Classification of value for money and use of evidence within the framework

There are concepts, such as the pursuit of value for money (VFM) and use of evidence, which could be considered attributes but which are particularly relevant to some functions and are therefore captured through this side of the framework.

By utilising the deep knowledge and experience of the founder organisations and their networks, and by undertaking extensive consultation, we have identified a range of themes relevant to each indicator, and metrics to measure each theme. These themes and metrics are described in Chapter 4. There are a number of indicators and themes for which appropriate metrics could not be identified and as such measurement of these indicators and themes has not been possible; they are thus omitted from this edition. There are naturally a number of potential limiting factors relevant to the metrics included in the Index and these are described in further detail in Annex A.

Chapter 4: Measuring against the framework

Chapter 4 explores the data used to measure against the framework specified in Chapter 3, and the weighting given to data to produce the score for each indicator.

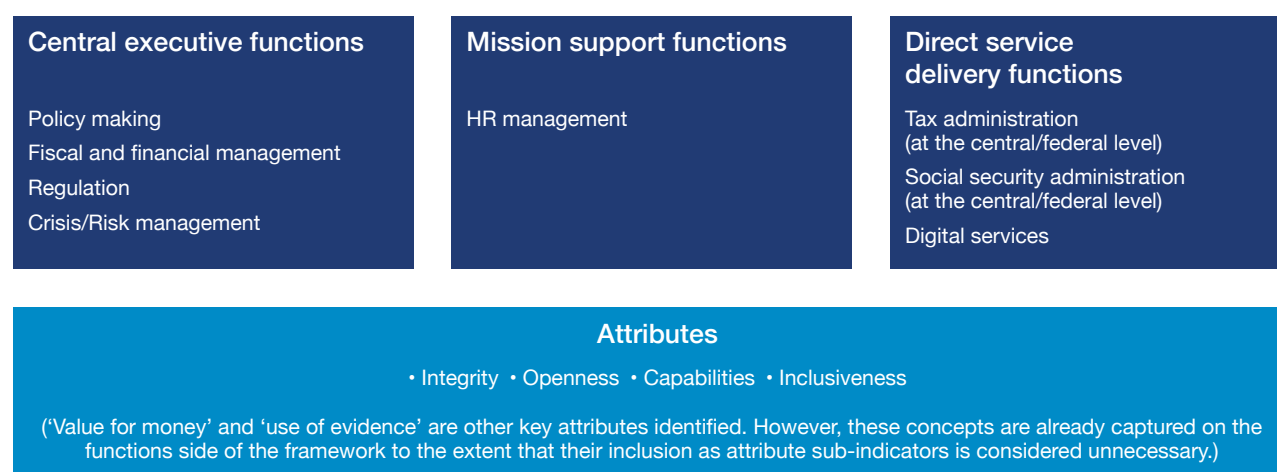
4.1 Data availability

The InCiSE measurement framework, as outlined in Chapter 3, is such that if a civil service were to score highly against it, it is reasonable to conclude that the civil service would be high-performing relative to its international counterparts. This chapter of the report focuses on what is measured under each indicator, and the data used to do so. However, it should first be noted that existing data does not enable measurement against all of the InCiSE framework. Nevertheless, the Index results provide valuable insights for accountability and performance improvement and we will strive to close gaps in data coverage as it is developed further. The metrics included in the Index represent a wide range of data sources and are derived from a combination of administrative data, survey findings and expert assessments. Although every effort has been made to ensure data is current and measures the subject of civil service effectiveness closely, due to limited data availability, some metrics may measure wider public sector, rather than civil service performance. While all data used is the most recent available, not all of the data used is updated every year. Several countries were excluded from the pilot Index because they had fewer than 75% of the metrics available.

Of the eleven functions proposed in the InCiSE framework, eight are measured and included in the pilot InCiSE composite Index (explained in Chapter 7). Of the six attributes in the InCiSE framework, four are measured and included in the composite.

Table 4: Indicators included in the InCiSE composite	
Measured and included in the InCiSE composite.	Not yet measured or included in the InCiSE composite due to data availability.
<p>Central executive functions</p> <ul style="list-style-type: none"> Policy making Fiscal and Financial Management (FFM) Regulation Crisis/risk management <p>Mission support functions</p> <ul style="list-style-type: none"> Human Resource Management (HRM) <p>Direct service delivery functions</p> <ul style="list-style-type: none"> Tax administration Social security administration Digital services <p>Attributes</p> <ul style="list-style-type: none"> Integrity Openness Capabilities Inclusiveness 	<p>Mission support functions</p> <ul style="list-style-type: none"> Procurement Information Technology (IT) Finance <p>Attributes</p> <ul style="list-style-type: none"> Staff engagement Innovation

In some of the areas where data was not available we are aware of new data collection that may help to fill gaps over time. In other areas, new data collection may need to be initiated over time to fulfill the measurement of this indicator.

Figure 6: Functions and attributes included in the pilot InCiSE composite

A wealth of data, from survey respondents, expert opinion and administrative data underlies the 12 indicators measured in the Index. In total, 76 metrics are spread across these indicators – although some metrics are themselves indices meaning that the actual number of ‘total metrics’ is far higher. 34 metrics underlie the attributes, while 42 metrics underlie the core functions. The quality of this data varies – in Figure 7 we set out an indicative assessment of the quality of data supporting each indicator.

Figure 7: Indicative data quality for framework indicators

			Rag rating summary of data quality
Attributes Weight: 1/3 Metrics: 34		Integrity	Green
		Openness	Green
		Capabilities	Red
		Inclusiveness	Red
		Staff Engagement	N/A
		Innovation	N/A
Core functions Weight: 2/3 Metrics: 42	Central Executive	Policymaking	Yellow
		Fiscal and financial	Yellow
		Regulation	Green
		Risk/Crisis Management	Yellow
	Mission Support	Procurement	N/A
		HR	Yellow
		IT	N/A
		Finance	N/A
	Direct Service Delivery	Tax Administration	Red
		Social Security Administration	Red
		Digital Services	Yellow

The data quality was assessed according to the following criteria:

- Green** – The indicator contains a large number of metrics, which seem to give a detailed, and relatively (but not necessarily completely) comprehensive picture of performance on the indicator. The metrics generally have few limitations. For example they get close to measuring civil service, rather than wider public sector, performance, and they are gathered from a regularly updated data source, so are up-to-date. Whilst metrics based on subjective expert assessments have potential drawbacks (see Chapter 8 covering sensitivity analysis for further details), indicators which utilise these where measurement by other means is particularly difficult are still given green ratings.
- Yellow** – The indicator contains a number of metrics, which seem to give a fairly detailed picture of performance on the indicator. However there are themes which should be measured under the indicator, which currently are not. The metrics included generally have limitations, but these are not prohibitive. Some may be based on subjective expert assessment. Some may measure wider public sector, rather than civil service performance, and some may be out of date.
- Red** – the indicator contains a small number of metrics, which do not generally give a detailed or comprehensive picture of performance on the indicator, but do give some partial information about the effectiveness of the civil service in this area.

- **N/A** denotes those indicators for which severe data limitations have prevented their inclusion in the pilot Index.

4.2 Data underpinning the indicators

Many of the datasets from which the metrics are drawn are updated annually, enabling the Index to be revised on a regular basis to reflect country developments. This iteration of the index includes data up until January 2017.

In the following tables we set out the metrics underpinning each of the indicators and the weighting given to each to make up the indicator score. Metrics were not identified for all the themes we would like to measure – we identify where this is the case in the tables by use of ticks (✓), crosses (X) and tildes (~).

✓	A tick indicates the theme is closely represented by one or more metrics utilised in the framework
X	A cross indicates that data has not yet been identified to represent this theme of an indicator in the framework
~	A tilde indicates that the theme is only approximately represented by one or more metrics in the framework

In each table the shorthand for the metrics, used in the Index interactive graphics, which are available online, are shown in italics.

4.2.1 Functions

Central Executive

Policy Making

The score for Policy Making is calculated based on eight metrics from the Bertelsmann Sustainable Governance Indicators which mainly assess the role civil servants play in setting strategic policy direction, coordination of policy across government, and the monitoring of policy implementation. The Bertelsmann Sustainable Governance Indicators are updated annually and cover 41 countries. Data included is from 2016.

Bertelsmann Sustainable Governance Indicators:

“As a cross-national comparative survey designed to identify and foster successes in effective policymaking, the SGI explores how governments target sustainable development. We advocate for more sustainable governance, which is built on three pillars:

- Policy Performance
- Democracy
- Governance

Driven by evidence-based analyses, the SGI helps a variety of stakeholders throughout the OECD and EU navigate the complexity of effective governance.¹⁸”

Further information about the Bertelsmann Sustainable Governance Indicators can be found at <http://www.sgi-network.org/2016/>.

18 Sustainable Governance Indicators. Bertelsmann Stiftung. [Online]. Available at: <http://www.sgi-network.org/2016/About>. [Accessed on 27 April 2017].

Table 5 shows the metrics measured in the policy making indicator, the source of data, and the weighting given within the indicator.

Table 5: Policy making indicator			
Theme to be measured	Currently captured in indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets).
The quality of policy advice	~	25	<ul style="list-style-type: none"> Bertelsmann ‘<i>Scholarly Advice</i>’ sub-indicator is used as a proxy for the evidence base to policy decisions by assessing the degree of civil servant to academic coordination: ‘How influential are non-governmental academic experts for government decision-making?’ (50%). Bertelsmann ‘<i>Government Office Expertise</i>’ sub-indicator assesses: ‘Does the government office/prime minister’s office (GO/PMO) have the expertise to evaluate ministerial draft bills substantively?’ (50%).
Degree of strategic policy direction	✓	25	<ul style="list-style-type: none"> Bertelsmann ‘<i>Strategic Planning</i>’ sub-indicator assesses: ‘How much influence do strategic planning units and bodies have on government decision-making?’ (100%).
Coordination of policy proposals	✓	25	<ul style="list-style-type: none"> Bertelsmann ‘<i>Ministerial Bureaucracy</i>’ sub-indicator assesses: ‘How effectively do ministry officials/civil servants coordinate policy proposals?’ (50%). Bertelsmann ‘<i>Line Ministries</i>’ sub-indicator assesses ‘To what extent do line ministries involve the government office/prime minister’s office in the preparation of policy proposals?’ (50%).
Timeliness and accuracy of policy delivery	X	–	No data identified
Degree of policy monitoring	✓	25	<ul style="list-style-type: none"> Bertelsmann ‘<i>Monitoring Ministries</i>’ sub-indicator assesses: ‘How effectively does the GO/PMO monitor line ministry activities with regard to implementation?’ (33%). Bertelsmann sub-indicator ‘<i>Monitoring Agencies</i>’, assesses: ‘How effectively do federal and subnational ministries monitor the activities of bureaucracies/executive agencies with regard to implementation?’ (33%). Bertelsmann ‘<i>National Standards</i>’ sub-indicator assesses: ‘To what extent does central government ensure that subnational self-governments realise national standards of public services?’ (33%).

The weighting within the policy making indicator is equally split between the four themes where metrics were identified. Within the themes, weighting was equally split between the metrics.

Fiscal and financial management

The score for Fiscal and Financial Management is calculated based on two sources:

- The World Economic Forum (WEF) Global Competitiveness Index (GCI)
- OECD medium-term and performance budgeting data

The WEF GCI covers 150 countries, and is updated annually. Survey data included is from 2016. The OECD data covers 32 countries. The survey data included here is from 2011 (performance budgeting) and 2012 (medium-term budgeting), and was published in the 2013 Government at a Glance report.

Global Competitiveness Index:

“The GCI combines 114 indicators that capture concepts that matter for productivity and long-term prosperity.... These indicators are grouped into 12 pillars...: institutions, infrastructure, macroeconomic environment, health and primary education, higher education and training, goods market efficiency, labor market efficiency, financial market development, technological readiness, market size, business sophistication, and innovation. These pillars are in turn organised into three subindexes: basic requirements, efficiency enhancers, and innovation and sophistication factors. The three subindexes are given different weights in the calculation of the overall Index, depending on each economy's stage of development, as proxied by its GDP per capita and the share of exports represented by raw materials¹⁹.”

OECD medium-term budgeting:

“Data... draw upon country responses to questions from the 2012 OECD Survey on Budgeting Practises and Procedures. Responses represent the countries' own assessments of current practices and procedures. The composite index... contains 10 variables that cover information on the existence of medium-term perspective in the budget process, the number of years the estimate covers, the types of expenditures included in the frameworks, the possibility of carry over unused funds from one year to another and how they are monitored²⁰.”

OECD performance budgeting:

“Data refer to 2011 and draw upon country responses to questions from the 2011 OECD Survey on Performance Budgeting. Responses represent countries' own assessments of current practices and procedures. For EU member countries, results exclude any EU funding. The composite index... contains 11 variables that cover information on the availability and type of performance information developed, processes for monitoring and reporting on results and whether (and how) performance information is used on budget negotiations and decision making by the central budget authorities, line ministries and politicians²¹.”

19 Global Competitiveness Index. World Economic Forum. [Online]. Available at: <http://reports.weforum.org/global-competitiveness-report-2015-2016/methodology/>. [Accessed on 27 April 2017].

20 Government at a Glance, Medium-term expenditure frameworks. OECD. [Online]. Available at: <http://www.oecd-ilibrary.org/docserver/download/4213201ec027.pdf?expires=1493288249&id=id&accname=guest&checksum=F1942DF8725DA66F41F53532D47200D1>. [Accessed on 27 April 2017].

21 Government at a Glance 2013, Performance Budgeting. OECD. [Online]. Available at: <http://www.oecd-ilibrary.org/docserver/download/4213201ec029.pdf?expires=1493288108&id=id&accname=guest&checksum=C9DAFD442CFD772363B3F79047EE4F65>. [Accessed on 27 April 2017].

Table 6 shows the metrics measured in the fiscal and financial management indicator, the source of data, and the weighting given within the indicator.

Table 6: Fiscal and financial management indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Economic appraisal (use, quality, guidance)	~	33	<ul style="list-style-type: none"> • <i>'Public Spending'</i>: WEF GCI business executive opinion: 'How would you rate the composition of public spending in your country?' (1) extremely wasteful; (7) highly efficient in providing necessary goods and services (100%).
Economic evaluation (use, quality, guidance)	~		
Medium-term budgeting	✓	33	<ul style="list-style-type: none"> • OECD <i>'Medium-Term Budgeting'</i> index (100%).
Performance budgeting	✓	33	<ul style="list-style-type: none"> • OECD <i>'Performance Budgeting'</i> index (100%).

The weighting within the Fiscal and Financial Management indicator gives a one third share to the combined 'economic appraisal'/'economic evaluation' themes, as between these themes there is one metric to approximately capture them, and one third each to 'medium-term budgeting' and 'performance budgeting'.

Regulation

The score for Regulation is calculated based on six metrics from the OECD's 'Indicators of Regulatory Policy and Governance' to assess 3 parts of the process behind creating regulation. The OECD data covers 34 countries and is updated every 3-4 years. Survey data included is from 2014.

Indicators of Regulatory Policy and Governance:

"Indicators of Regulatory Policy and Governance (IREG) present up-to-date evidence of OECD member countries' regulatory policy and governance practices advocated in the 2012 Recommendation of the Council on Regulatory Policy and Governance. They cover in detail three principles of the 2012 Recommendation: stakeholder engagement, Regulatory Impact Assessment (RIA) and ex post evaluation, and provide a baseline measurement to track countries' progress over time and identify areas for reform."²²

22 Indicators of Regulatory Policy and Governance 2015. OECD. [Online]. Available at: <http://www.oecd.org/gov/regulatory-policy/indicators-regulatory-policy-and-governance.htm>. [Accessed 27 April 2017].

Table 7 shows the metrics measured in the regulation indicator, the source of data, and the weighting given within the indicator.

Table 7: Regulation indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Ex ante appraisal: application, quality, sustainability, transparency, oversight.	✓	33	<ul style="list-style-type: none"> • OECD i-reg indicators – Regulatory ‘<i>Impact Assessment – Primary Laws</i>’ (50%); • OECD i-reg indicators – Regulatory ‘<i>Impact Assessment – Secondary Laws</i>’ (50%). <p>Stakeholder engagement and ex post evaluation assessments are made on four areas of performance: methodology, systematic adoption, transparency and oversight/quality control.</p>
Stakeholder engagement: application, quality, transparency, oversight	✓	33	<ul style="list-style-type: none"> • OECD i-reg indicators – ‘<i>Stakeholder Engagement – Primary Laws</i>’ (50%); • OECD i-reg indicators – ‘<i>Stakeholder Engagement – Secondary Laws</i>’ (50%).
Ex post evaluation: application, quality, sustainability, transparency, oversight	✓	33	<ul style="list-style-type: none"> • OECD i-reg indicators – Ex-post ‘<i>Evaluation – Primary Laws</i>’ (50%); • OECD i-reg indicators – Ex-post ‘<i>Evaluation – Secondary Laws</i>’ (50%).

The weighting within the Regulation indicator is equally split between the three themes. Within the themes, the weighting was equally split between metrics.

For each of the 3 parts of the process, assessments are made for both primary and secondary laws, giving 6 separate composite indicators in total. Data was collected through surveys of government officials. Countries were asked to support responses with evidence.

Crisis/risk management

This indicator covers nine areas related to disaster risk management. All metrics are based on the parts of the Hyogo Framework for Action (HFA) country progress reports, which track each country's progress towards implementing the UN priorities for disaster risk reduction action, most relevant to the civil service. These progress reports contain scores for each country, but these are based on self-assessment so are seen as less robust and are not included in our indicator. Scores on each of the nine metrics here are instead calculated by 'counting yeses' to a number of binary questions which are factual in nature (only examples are given for each area in the table but forty-three questions are used in total). Survey data used here is from 2015.

Hyogo Framework for Action:

"The World Conference on Disaster Reduction was held from 18 to 22 January 2005 in Kobe, Hyogo, Japan, and adopted the present Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters (hereafter referred to as the "Framework for Action"). The Conference provided a unique opportunity to promote a strategic and systematic approach to reducing vulnerabilities and risks to hazards. It underscored the need for, and identified ways of, building the resilience of nations and communities to disasters.²³"

²³ Hyogo Framework for Action 2005 - 2015. World Conference on Disaster Reduction. [Online]. Available at: <http://www.unisdr.org/2005/wcdr/intergover/official-doc/L-docs/Hyogo-framework-for-action-english.pdf>. [Accessed 27 April 2017].

Table 8 shows the metrics measured in the crisis/risk management indicator, the source of data, and the weighting given within the indicator.

Table 8: Crisis/risk management indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Integrated risk planning	✓	22 (2/9) ²⁴	<ul style="list-style-type: none"> • <i>'Risk Planning Extent'</i>: The extent to which disaster risk is integrated into national policy planning is measured by counting the types of policy planning into which risk is integrated (out of 7, e.g. national development plans) (50%). • <i>'Disaster Spending Appraisal'</i>: A score for economic appraisal looks at whether the costs and benefits of spending related to disaster spending are considered (50%).
Risk monitoring	✓	22 (2/9)	<ul style="list-style-type: none"> • <i>'Risk Assessment Quality'</i>: The quality of multi-hazard risk assessment is measured by assessing the processes behind risk assessments (out of 5, e.g. whether risk assessments have agreed national standards) (50%); • The <i>'Degree of Risk Monitoring'</i> is measured by assessing the extent and use of reports and databases of this type (50%).
Public information dissemination and public awareness strategies	✓	33 (3/9)	<ul style="list-style-type: none"> • <i>'Early Warning Systems'</i>: The quality of early warning systems is assessed by looking at the processes in place (out of 3, e.g. whether protocols are used and applied) (33%). • <i>'Public Information'</i> dissemination is scored by looking at the types of communications in place (out of 4, e.g. is a national disaster information system publicly available?) (33%). • <i>'Public Awareness Strategy'</i> looks at the relevant workstreams in place (out of 5, e.g. whether public education campaigns for risk awareness exist) (33%).
International cooperation and risk coordination	✓	11 (1/9)	<ul style="list-style-type: none"> • <i>'International Cooperation'</i> is measured by assessing the number of processes/activities in place for international cooperation and international risk management (out of 6, e.g. whether transboundary protocols are in place) (100%).
Preparedness for disaster response	X	–	No data identified
Post disaster assessment methodology	✓	11 (1/9)	<ul style="list-style-type: none"> • <i>'Post Disaster Assessment'</i>: A score is given for post disaster damage and loss assessment by considering, for example, whether a specified methodology for this exists (out of 3) (100%).

Note – percentages do not sum due to rounding

²⁴ For this indicator, weightings between themes are rounded in the table to the nearest percentage point and are based on multiples of 1/9 fractions.

The weighting within the Crisis/risk management indicator relates to the availability of metrics under each theme; there are nine metrics in total underlying this indicator across its six themes. For the 'integrated risk planning' and 'risk monitoring' themes, both have 2/9 of the overall indicator weighting, with weighting split equally across the two metrics for each theme. For 'public information dissemination and public awareness strategies', this theme has 3/9 (1/3) of the weight, split equally across its three metrics. Both 'international cooperation and risk coordination' and 'post disaster assessment methodology' are measured by a single metric and are given 1/9 of the total weighting. No suitable data to measure 'preparedness for disaster response' could be found so it is not currently captured.

Mission support

Human resource management

The score for Human Resource Management is calculated based on Quality of Government (QoG) expert assessments to determine two important parts of HR management: the meritocracy of recruitment and attracting talent. Data included is from 2015.

Quality of Government:

The Quality of Government Institute at the University of Gothenburg have developed the QoG Expert Survey Data.

"The QoG Survey is a unique data set with information on the structure and behavior of public administration in a range of different countries. The data covers 159 countries and is based on a web survey of 1294 experts.

The dataset covers different dimensions of Quality of Government, such as, politicisation, professionalisation, openness, and impartiality.²⁵"

²⁵ Expert Survey Data. The Quality of Government Institute, University of Gothenburg. [Online]. Available at: <http://qog.pol.gu.se/data/datadownloads/qogexpertsurveydata>. [Accessed 27 April 2017].

Table 9 shows the metrics measured in the human resource management indicator, the source of data, and the weighting given within the indicator.

Table 9: Human resource management indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Meritocracy of recruitment	✓	75	<ul style="list-style-type: none"> • <i>'Applicant Skills'</i>: QoG expert assessment of extent to which the skills and merits of the applicants decide who gets the job when recruiting public sector employees (25%); • <i>'Connections Bias in Recruitment: Political'</i>: QoG expert assessment of extent to which political connections of the applicants decide who gets the job (25%); • <i>'Connections Bias in Recruitment: Personal'</i>: QoG expert assessment of extent to which personal connections of the applicants decide who gets the job (25%); • <i>'Recruitment via Formal Exam System'</i>: QoG expert assessment of extent to which public sector employees are hired using a formal examination system (25%).
Attracting and retaining talent	✓	25	<ul style="list-style-type: none"> • <i>'Comparable Salaries'</i>: QoG expert assessment of extent to which senior officials have salaries that are comparable with the salaries of private sector managers with roughly similar training and responsibilities (100%).
Talent deployment (i.e. minimising skills gaps)	X	–	No data identified
Performance management	X	–	No data identified
Quality of learning and development	X	–	No data identified
Level of customer (i.e. civil servant) satisfaction	X	–	No data identified

The weighting within the Human Resource Management indicator awards 75% of the total weight to the 'meritocracy of recruitment' theme. This is split equally among the four metrics which underlie this theme. The 'attracting and retaining talent' theme has the remaining 25% with a single metric capturing this theme. The remaining four themes are not currently captured in the indicator set because of a lack of suitable data.

Direct service delivery

Tax Administration

The score for Tax Administration is calculated based on metrics mostly taken from 'OECD Tax Administration 2015' which is a fairly comprehensive assessment of OECD tax administration systems. Metrics included cover efficiency and the extent of digital use in tax collection made up of 2013 data. In addition, some metrics are taken from the World Bank's Doing Business Index with data from 2016.

OECD Tax Administration 2015:

"Tax Administration 2015, produced under the auspices of the Forum on Tax Administration, is a unique and comprehensive survey of tax administration systems, practices and performance across 56 advanced and emerging economies (including all OECD, EU, and G20 members). Its starting point is the premise that revenue bodies can be better informed and work more effectively together given a broad understanding of the administrative context in which each operates. However, its information content is also likely to be of interest to many external parties (e.g. academics, external audit agencies, regional tax bodies, and international bodies providing technical assistance).

The series identifies some of the fundamental elements of national tax system administration and uses data, analyses and country examples to identify key trends, comparative levels of performance, recent and planned developments, and good practices.²⁶"

Doing Business index:

"Doing Business presents quantitative indicators on business regulations and the protection of property rights that can be compared across 190 economies – from Afghanistan to Zimbabwe – and over time.

Doing Business measures regulations affecting 11 areas of the life of a business. Ten of these areas are included in this year's ranking on the ease of doing business: starting a business, dealing with construction permits, getting electricity, registering property, getting credit, protecting minority investors, paying taxes, trading across borders, enforcing contracts and resolving insolvency.²⁷"

²⁶ Tax Administration 2015. OECD. [Online]. Available at: <http://www.oecd.org/ctp/administration/tax-administration-23077727.htm>. [Accessed 27 April 2017].

²⁷ Doing Business 2017. The World Bank. [Online]. Available at: <http://www.doingbusiness.org/reports/global-reports/doing-business-2017>. [Accessed 27 April 2017].

Table 10 shows the metrics measured in the Tax Administration indicator, the source of data, and the weighting given within the indicator.

Table 10: Tax administrator indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Overall efficiency of collection	✓	33	<ul style="list-style-type: none"> • <i>'Collection Cost'</i>: Data from OECD Tax Administration on: cost of collection ratios (administrative costs/net revenue collected) (50%); • <i>'Tax Debt'</i> ratios (total year end tax debt (excl. disputed)/net revenue collected) (50%).
User centricity of services	✓	33	<ul style="list-style-type: none"> • <i>'Time to Pay Taxes – Business'</i>: Data from the World Bank's 'Doing Business' Index on the time it takes businesses to pay taxes (100%).
Extent and quality of digital provision	✓	33	<ul style="list-style-type: none"> • <i>'E-Filed Tax Returns – Personal'</i>: Data from OECD Tax Administration on: the percentage of tax returns e-filed during the last fiscal year for personal taxes (33%); • <i>'E-Filed Tax Returns – Corporate'</i>: Data from OECD Tax Administration on: the percentage of tax returns e-filed during the last fiscal year for corporate taxes (33%); • <i>'E-Filed Tax Returns – VAT'</i>: Data from OECD Tax Administration on: the percentage of tax returns e-filed during the last fiscal year for VAT taxes (33%).
Prevention of tax evasion	X	–	No data identified
Level of tax gap measurement	X	–	No data identified

The weighting within the Tax Administration indicator is currently equally split between the three themes where relevant data was found. Within the themes, weighting is equally split between the metrics.

Social Security Administration

The score for Social Security Administration is calculated based on metrics mostly taken from data from the European Commission (via Eurostat) on administration costs as a proportion of total expenditure on social security, capturing the overall efficiency of the social security administration system. Total expenditure includes: social protection benefits, admin costs, sickness/health care payments, disability payments, pensions, child benefits, unemployment benefits, housing benefits, social exclusion benefits and other expenditure. The comparison of heterogeneous systems is difficult so the conclusions which can be drawn from the indicator may need further investigation. Only a single metric is included at this point.

Table 11 shows the metrics measured in the social security administration indicator, the source of data, and the weighting given within the indicator.

Table 11: Social security administration indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Overall efficiency of distribution	✓	100	<ul style="list-style-type: none"> Administration costs as a percentage of total expenditure on social security (100%).
User centricity of services	X	–	No data identified
Extent and quality of digital provision	X	–	No data identified
Prevention of fraud	X	–	No data identified

The Social Security indicator currently wholly relies on one data source, given 100% of the indicator weighting, which relates to the theme of overall efficiency of distribution.

Digital services

The score for digital services is calculated based on metrics taken from the European Commission's E-Government Benchmarks to assess the quality of digital services provided across four main dimensions. A detailed assessment of the quality of services across seven life events, according to the four dimensions, is made. The benchmark is usually published annually and covers 33 European countries. Data is an average of Mystery Shopping exercises conducted in 2014 and 2015.

E-Government benchmarks:

"E-Government services were assessed in 34 participating countries, including all of the EU28. The benchmark makes use of Mystery Shopping, where the quality and quantity of online public services is measured by assessors acting as a user. The subject of the benchmark is a set of seven life events. Together, these life events represent virtually all domains of government. Each life event is reviewed once every two years.²⁸"

Table 12 shows the metrics measured in the digital services indicator, the source of data, and the weighting given within the indicator.

Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
User centricity of services	✓	25	<ul style="list-style-type: none"> European Commission E-Government benchmark: '<i>User Centricity</i>' (100%), indicates to what extent (information about) a service is provided online.
Transparency of service	✓	25	<ul style="list-style-type: none"> European Commission E-Government benchmark: '<i>Transparency</i>' (100%), indicates to what extent governments are transparent as regards a) their own responsibilities and performance, b) the process of service delivery and c) personal data involved.
Cross-border mobility of services	✓	25	<ul style="list-style-type: none"> European Commission E-Government benchmark: '<i>Cross Border Mobility</i>' (100%), indicates to what extent European users can use online services in another country;
The availability of 'key enablers'	✓	25	<ul style="list-style-type: none"> European Commission E-Government benchmark: '<i>Key Enablers</i>' (100%), indicates the extent to which five technical pre-conditions for eGovernment are used.

The weighting within the Digital Services indicator is equally split between the four themes. As only one metric was used within each theme, this was given 100% of the theme weighting.

28 E-Government Benchmark 2016. European Commission. [Online]. Available at: https://www.egovement.ch/index.php/download_file/force/991/3343/. [Accessed on 27 April 2017].

4.2.2 Attributes

Integrity

The score for Integrity is calculated based on metrics covering six main themes related to the integrity and other key values of civil servants. Metrics capturing perceptions of these values from country experts, citizens and businesses make up 80% of the indicator. The remaining 20% assesses the laws and procedures in place to preserve integrity and prevent conflicts of interest, such as the degree of whistleblower protection. The indicator uses a range of data sources:

- Transparency International's Global Corruption Barometer, which covers 100 countries. Data from 2013.
- World Economic Forum's (WEF) Global Competitiveness Index (GCI), which covers over 150 countries. Survey data from 2016.
- Quality of Government (QoG), which covers over 100 countries. Expert survey data from 2015.
- OECD private interest (2013), conflict of interest and whistleblower data (2014), which covers 36, 26 and 30 countries respectively. Survey data from 2013 and 2014.

Global Corruption Barometer:

"In the Global Corruption Barometer we ask people to tell us how corrupt they think different powerful groups in their country are – that is whether they think "none", "some", "most" or "all" of them are corrupt. The nine different groups that we ask about include high-level political actors (the president/prime minister's office, members of parliament and government officials); key public sector employees who interact with citizens (tax officials, the police, judges/magistrates and local government councillors); and those who are not part of the public sector, but are influential in political life (business executives and religious leaders).²⁹"

Global Competitiveness Index:

See Policy Making above.

Quality of Government:

See Human Resource Management above.

OECD conflict of interest, private interest, and whistleblower data:

"Government at a Glance provides readers with a dashboard of key indicators assembled with the goal of contributing to the analysis and international comparison of public sector performance. Indicators on government revenues, expenditures, and employment are provided alongside key output and outcome data in the sectors of education, health and justice.

Government at a Glance also includes indicators on key governance and public management issues, such as transparency in governance, regulatory governance, public procurement and the implementation of employment and remuneration reforms since 2009.³⁰"

²⁹ Global Corruption Barometer. Transparency International. [Online]. Available at: http://transparencia.org.es/wp-content/uploads/2016/11/informe_barometro-2016-ingles.pdf. [Accessed 27 April 2017].

³⁰ Government at a Glance 2015. OECD. [Online]. Available at: http://www.oecd-ilibrary.org/governance/government-at-a-glance-2015_gov_glance-2015-en. [Accessed 27 April 2017].

Table 13 shows the metrics measured in the integrity indicator, the source of data, and the weighting given within the indicator.

Table 13: Integrity indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Levels/ perceptions of corruption	✓	30	<ul style="list-style-type: none"> • <i>'Corruption Perceptions'</i>: Global Corruption Barometer – % of people viewing public officials/civil servants as corrupt (25%); • <i>'Public Officials Stealing'</i>: QoG – expert opinion of how often public sector officials steal or embezzle public funds (25%); • <i>'Public Officials Favours for Bribes'</i>: QoG – expert opinion of how often public sector officials grant favours for bribes (25%); • <i>'Government Favouritism of Business'</i>: WEF GCI – business executive opinion of extent government officials show favouritism to well-connected firms (25%).
Fairness and impartiality	✓	12.5	<ul style="list-style-type: none"> • <i>'Fair Treatment by Public Officials'</i>: QoG expert assessment of extent to which public sector employees treat some groups in society unfairly (50%); • <i>'Public Officials Act Impartially'</i>: QoG expert assessment of extent to which public sector employees act impartially when deciding how to implement policy (50%).
Adherence to rules and procedures	✓	12.5	<ul style="list-style-type: none"> • <i>'Public Officials Follow Rules'</i>: QoG expert assessment of extent to which public sector employees strive to follow rules (100%).
Striving to serve citizens and ministers	✓	12.5	<ul style="list-style-type: none"> • <i>'Public Officials Strive to Help Citizens'</i>: QoG expert assessment of extent to which public sector employees strive to help citizens (33%); • <i>'Public Officials Strive to Implement Policies'</i>: QoG expert assessment of extent to which public sector employees strive to implement policies decided by political leaders (33%); • <i>'Public Officials Strive to Fulfill Ideology'</i>: QoG expert assessment of extent to which public sector employees strive to fulfil ideology of party in government (33%).
Work ethic	✓	12.5	<ul style="list-style-type: none"> • <i>'Employee Absences'</i>: QoG expert assessment of extent to which public sector employees are absent without permission (50%); • <i>'Employee Efficiency'</i>: QoG expert assessment of extent to which public sector employees strive to be efficient (50%).
Processes in place to preserve integrity and prevent conflicts of interest	✓	20	<ul style="list-style-type: none"> • <i>'Post Employment Cooling Off'</i>: OECD data on whether a post-employment cooling off period exists and whether it's paid (25%); • <i>'Lobbyist Protection'</i>: OECD data on degree of protection against lobbyists and other private interests influencing advisory groups (25%); • <i>'Whistleblower Protection: Coverage'</i>: OECD data on number of groups who receive whistleblower protection (25%); • <i>'Whistleblower Protection: Degree'</i>: QoG expert assessment of degree of whistleblower protection (25%).

The weighting within the Integrity indicator is split between the six themes as follows: Levels/perception of corruption – 30%; fairness and impartiality – 12.5%; adherence to rules and procedures – 12.5%; striving to serve citizens and ministers – 12.5%; work ethic – 12.5%; and processes in place to preserve integrity and prevent conflicts of interest – 20%. Levels/perceptions of corruption and Processes in place to preserve integrity and prevent conflicts of interest themes contain metrics from a wider variety of data sources and were therefore weighted more heavily.

Within each theme, the weight is further split between the metrics which measure that theme. For all the Integrity themes, this weight is split equally among the metrics, although the number of metrics under each theme varies: 1 metric (adherence to rules and procedures); 2 metrics (fairness and impartiality and work ethic); 3 metrics (striving to serve citizens and ministers); and 4 metrics (levels/perceptions of corruption and processes in place to preserve integrity and prevent conflicts of interest).

Openness

The score for Openness assumes there are two distinct themes embedded within the concept of open government: societal consultation and transparency. Each get 50% weight within the indicator.

The data on both sides consists of existing composite indicators, for example composites designed to measure open government as a whole and others looking more specifically at open government data. There are four metrics on the consultation side (the top two themes in the table) and five metrics on the transparency side (the bottom four themes in the table).

The indicator uses a range of data sources:

- The World Justice Project's Open Government Index, which covers 102 countries. Survey data from 2015.
- The UN's E-participation Index, covering 193 countries and updated biennially. Survey data from 2016.
- Bertelsmann's Sustainable Governance Indicators (SGIs), covering 41 countries and updated annually, expert survey data from 2016.
- The World Wide Web Foundation's Open Data Barometer, covering 92 countries and updated annually. Survey data from 2015.
- The Open Knowledge Foundation's Open Data Index, covering 122 countries and updated annually. Survey data used from 2015.
- OECD's OURdata index, covering 29 countries. Survey data from 2014.

Open Government Index:

“The WJP Open Government Index is the first effort to measure government openness based on the general public’s experiences and perceptions in 102 countries. The Open Government Index is composed of four dimensions: publicised laws and government data, right to information, civic participation, and complaint mechanisms.³¹”

E-participation Index:

“The e-participation index (EPI) is derived as a supplementary index to the UN E-Government Survey. It extends the dimension of the Survey by focusing on the use of online services to facilitate provision of information by governments to citizens (“e-information sharing”), interaction with stakeholders (“e-consultation”), and engagement in decision-making processes (“e-decision making”).

E-Participation Framework:

- E-information: Enabling participation by providing citizens with public information and access to information without or upon demand
- E-consultation: Engaging citizens in contributions to and deliberation on public policies and services
- E-decision-making: Empowering citizens through co-design of policy option and co-production of service components and delivery modalities.³²”

Sustainable Governance Indicators:

“The SGI is a platform built on a cross-national survey of governance that identifies reform needs in 41 EU and OECD countries.

The SGI brings together a broad network of experts and practitioners aiming to understand what works best in sustainable governance.³³”

Open Data Barometer:

“Covering 92 countries in the present edition, the Barometer ranks nations on:

- Readiness: How prepared are governments for open data initiatives? What policies are in place?
- Implementation: Are governments putting their commitments into practice?
- Impact: Is open government data being used in ways that bring practical benefit?”

Open Data Index:

“The Global Open Data Index is an annual effort to measure the state of open government data around the world. The crowdsourced survey is designed to assess the openness of specific government datasets according to the Open Definition.³⁴”

OURdata index:

“The OECD OURdata Index measures government efforts to implement the G8 Open Data charter based on the availability, accessibility and government support to promote the reuse of data, focusing on the central OGD portal in each country.³⁵”

31 Open Government Index. World Justice Project. [Online]. Available at: <https://worldjusticeproject.org/our-work/wjp-rule-law-index/wjp-open-government-index-2015>. [Accessed 27 April 2017].

32 E-Participation Index. United Nations. [Online]. Available at: <https://publicadministration.un.org/egovkb/en-us/About/Overview/E-Participation>. [Accessed on 27 April 2017].

33 Sustainable Governance Indicators. Bertelsmann Stiftung. [Online]. Available at: <http://www.sgi-network.org/2016/>. [Accessed 27 April 2017].

34 Open Data Index. Open Knowledge International. [Online]. Available at: <http://index.okfn.org/about/>. [Accessed 27 April 2017].

35 Open, Useful, Reusable Data Index. OECD. [Online]. Available at: <http://www.oecd-ilibrary.org/docserver/download/4215081e.pdf?expires=1493300444&id=id&accname=guest&checksum=665A4CCF1038B2E655ECB2D663FA9D34>. [Accessed on 27 April 2017].

Table 14 shows the metrics measured in the openness indicator, the source of data, and the weighting given within the indicator.

Table 14: Openness indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
The degree and quality of societal consultation	✓	37.5	<ul style="list-style-type: none"> • The '<i>Civic Participation</i>' component of the Open Government Index (33%), which "measures the effectiveness of civic participation mechanisms, including the protection of the freedoms of opinion and expression, and assembly and association, and the right to petition the government. It also measures whether people can voice concerns to various government officers and members of the legislature, and whether government officials provide sufficient information and notice about decisions affecting the community, including opportunities for citizen feedback"; • '<i>E-Government Engagement</i>': The UN's E-Participation Index, which reviews the quality and usefulness of e-government programs for the purpose of engaging people in public policy-making and implementation (33%); • '<i>Negotiating Public Support</i>' Bertelsmann sub-indicator (33%), which "assesses how successfully the government consults with societal actors such as trade unions, employers' associations, leading business associations, religious communities, and social and environmental interest groups in preparing its policy."
The existence and quality of complaint mechanisms	✓	12.5	<ul style="list-style-type: none"> • The '<i>Complaint Mechanisms</i>' component of the Open Government Index (100%), which "measures whether people are able to bring specific complaints to the government about the provision of public services or the performance of government officers in carrying out their legal duties in practice, and how government officials respond to such complaints. It also measures whether people can challenge government decisions before another government agency or a judge."
Government data availability and accessibility	✓	30	<ul style="list-style-type: none"> • '<i>Open Data Practice and Impact</i>': The Open Data Barometer (ODB) measures the implementation of open data practice and is the only index to also measure the impact of open data (e.g. how many use it) (33%). • '<i>Government Datasets Openness</i>': The Open Data Index (ODI) measures whether publicly held data across 13 areas is defined as open, with results crowdsourced from volunteers reviewing websites (33%). • '<i>Data Availability and Government Support</i>': The OURdata index also aims to capture the availability and accessibility of data but uniquely it also attempts to measure the level of pro-active support governments provide to foster innovative re-use of the data (33%).
Government data impact and support for re-use	✓		

Table 14: Openness indicator

Right to information (e.g. FOIs)	✓	10	<ul style="list-style-type: none"> The '<i>Rights to Information</i>' component of the Open Government Index (100%) which "measures whether requests for information held by a government agency are granted. It also measures whether these requests are granted within a reasonable time period, if the information provided is pertinent and complete, and if requests for information are granted at a reasonable cost and without having to pay a bribe. This dimension also measures whether people are aware of their right to information, and whether relevant records – such as budget figures of government officials, ombudsman reports, and information relative to community projects – are accessible to the public upon request."
Publicised laws	✓	10	<ul style="list-style-type: none"> The '<i>Publicised Laws</i>' component of the Open Government Index (100%), which "measures whether basic laws and information on legal rights are publicly available, presented in plain language, and are made accessible in all languages used by significant segments of the population. This dimension also measures the quality and accessibility of information published by the government in print or online (i.e. active transparency), and whether administrative regulations, drafts of legislation, administrative decisions, and high court decisions are made accessible to the public in a timely manner."

The weighting within the Openness indicator is split between six themes as follows: Levels/perception of corruption – 30%; fairness and impartiality – 12.5%; adherence to rules and procedures – 12.5%; striving to serve citizens and ministers – 12.5%; work ethic – 12.5%; and processes in place to preserve integrity and prevent conflicts of interest – 20%. This choice of weighting relates to the number of metrics available under each theme.

Within each theme the weight is split by the number of metrics.

Capabilities

The score for Capabilities is based on OECD PIAAC data, which was collected between August 2011 and March 2012 in most participating countries. Around 166,000 adults, representing 724 million adults aged 16 to 65, were surveyed in 24 countries. Given that not all included countries are covered, some countries have all data imputed for this indicator (see table 18).

Programme for the International Assessment of Adult Competencies:

“The Programme for the International Assessment of Adult Competencies (PIAAC) developed and conducts the Survey of Adult Skills. The survey measures adults’ proficiency in key information-processing skills – literacy, numeracy and problem solving in technology-rich environments – and gathers information and data on how adults use their skills at home, at work and in the wider community.³⁶”

Table 15 shows the metrics measured in the capabilities indicator, the source of data, and the weighting given within the indicator.

Table 15: Capabilities indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Leadership capability	X	–	No data identified
Commercial capability	X	–	No data identified
Analytical capability	X	–	No data identified
Digital capability	X	–	No data identified
Core capability (e.g. problem-solving, numeracy, literacy skills)	✓	75	<ul style="list-style-type: none"> • <i>‘Literacy Skills’</i>: Taken from the OECD’s PIAAC (Survey of Adult Skills) data (analysis of the microdata): the proportion of the public sector getting level 4 or 5 for literacy skills (33%); • <i>‘Numeracy Skills’</i>: The proportion of the public sector getting level 4 or 5 for numeracy skills (33%); • <i>‘Problem Solving Skills’</i>: The proportion of the public sector getting level 3 for problem skills (33%).
Educational attainment of the workforce	✓	25	<ul style="list-style-type: none"> • <i>‘Educational Attainment’</i>: The proportion of the public sector with tertiary education, taken from the OECD’s PIAAC (Survey of Adult Skills) data (analysis of the microdata) (100%).

The weighting within the Capabilities indicator is split between the core capability and educational attainment of the workforce themes; the two areas for which usable data has been found. For the core capability theme, the weighting is equally split between the three metrics. Thus, although the core capability theme has three quarters of the overall weighting and the educational attainment theme the remaining quarter, the four metrics underpinning the indicator are equally weighted.

36 Survey of Adult Skills. OECD, Programme for the Assessment of Adult Competencies. [Online]. Available at: <http://www.oecd.org/skills/piaac/>. [Accessed 27 April 2017].

Inclusiveness

The score for Inclusiveness is based on the difference in demographics within central government compared to the country's workforce overall. A high score is given for having a similar demographic profile. The indicator used survey data published in Government at a Glance 2013 (GaaG) and collected in 2010, and Quality of Government survey data from 2015.

Government at a Glance

"Government at a Glance provides readers with a dashboard of key indicators assembled with the goal of contributing to the analysis and international comparison of public sector performance. Indicators on government revenues, expenditures, and employment are provided alongside key output and outcome data in the sectors of education, health and justice. Government at a Glance also includes indicators on key governance and public management issues, such as transparency in governance, regulatory governance, public procurement and the implementation of employment and remuneration reforms since 2009. While measuring government performance has long been recognised as playing an important role in increasing the effectiveness and efficiency of the public administration, following the economic crisis and fiscal tightening in many member countries, good indicators are needed more than ever to help governments make informed decisions regarding tough choices and help restore confidence in government institutions.³⁷"

³⁷ Government at a Glance. OECD. [Online]. Available at: http://www.oecd-ilibrary.org/governance/government-at-a-glance-2015_gov_glance-2015-en. [Accessed 27 April 2017].

Table 16 shows the metrics measured in the inclusiveness indicator, the source of data, and the weighting given within the indicator.

Table 16: Inclusiveness indicator			
Themes to be measured	Currently captured in Indicator?	Weighting in indicator (%)	<ul style="list-style-type: none"> Metrics to capture theme and the weighting each receives within the metric (shown in brackets)
Proportionate gender representation	✓	50	<ul style="list-style-type: none"> 'Gender: Central Government Share': OECD data – the absolute difference between the share of total central government employment filled by women and the share of women in the labour force (25%); 'Gender: Public Sector Share': QoG data – the absolute difference between proportion of women in public sector and in the labour force (25%); 'Gender: Management Share': OECD data – the absolute difference between the share of employment in top management positions within central government filled by women and the share of women in the labour force (25%); 'Gender: Senior Management Share': QoG data – the absolute difference between proportion of women in central government senior positions and in the labour force (25%)
Proportionate ethnic minority representation	✓	50	<ul style="list-style-type: none"> 'Ethnic and Religious Group Representation': QoG data – expert opinion as to whether 'Key ethnic and religious groups in society are proportionally represented among public sector employees (1=hardly ever, 7=almost always)' (100%)
Proportionate disability representation	X	–	No data identified
Proportionate socio-economic representation	X	–	No data identified
Proportionate Lesbian Gay Bi-sexual Transgender Other sexuality representation	X	–	No data identified

The weighting within the inclusiveness indicator is equally split between the gender representation and proportionate ethnic minority representation theme; the two areas for which usable data has been found. For the gender representation themes, the weighting is equally split between the four metrics.

For transparency, we set out limitations associated with the above data in Annex A. InCiSE is a long term project and we expect to include further data and refine the data currently included in the Index as we receive feedback about how the indicators could be further developed.

Chapter 5: Index country coverage

.....

This chapter examines which countries have been included and excluded from the pilot Index, and explains our approach where a country is included but some of their data is missing.

.....

5.1 Countries included in the Index

The ambition of the InCiSE Index is that the indicators may serve two purposes: firstly, to serve as an accountability tool, and secondly, to provide a performance improvement tool. To meet this ambition and achieve the greatest impact the Index aims to cover a wide and growing range of countries as it is developed further.

Covering the widest range of countries in the Index is tempered by availability of data. There is considerable variation in country coverage for the data used in this index (described in chapter 4). Expanding the range of countries would lead to a greater degree of data imputation, a reduction in the scope of the framework or a greater reliance on proxies for the civil service. Table 17 shows the proportion of metrics available for each country. Only two countries, Norway and the United Kingdom, had no missing metrics.

Table 17: Proportion of metrics available for each country

(Note: dark red area represents those countries not included in the pilot release of the Index)

Country	Total number of metrics	Proportion of metrics available (%)
Norway	76	100
United Kingdom	76	100
Austria	75	99
France	75	99
Italy	75	99
Netherlands	75	99
Finland	74	97
Germany	74	97
Denmark	73	96
Poland	73	96
Sweden	73	96
Czech Republic	72	95
Spain	72	95
Slovenia	70	92
Portugal	69	91
Slovakia	69	91
Hungary	68	89
Turkey	68	89
Belgium	67	88
Australia	66	87
New Zealand	66	87
Estonia	65	86
Mexico	65	86
Switzerland	65	86
Greece	64	84
Canada	62	82
Ireland	60	79
Japan	59	78
Korea, Republic of	59	78
Chile	58	76
United States of America	57	75
Bulgaria	57	75
Croatia	54	71
Romania	53	70
Iceland	51	67
Israel	49	64
Colombia	46	61

To balance the issue of country coverage and data availability, it was decided that any country with less than 75% of the data available would be excluded at this stage from the Index. Countries below this point have larger amounts of missing data and a greater proportion of the total data would need to be imputed. Bulgaria is excluded to keep the Index limited to OECD countries, for simplicity at this stage. As a result, we have 31 countries currently included in the pilot Index. Our hope is that further countries will be included as data is identified for use or new data collection is initiated.

Across the metrics underlying each indicator there is some variation in the countries covered by the data. Table 18 gives an overview of where data was available for each indicator and where data imputation was required.

Green indicates data was available for all metrics within the indicator for that country; amber indicates that data was available for some metrics within the indicator; and red indicates that no data was available for the indicator such that all metrics were imputed.

Table 18: Indicator data availability by country[illegible]

Policy making, fiscal and financial management and human resources are the only indicators where data is available for all metrics across all countries currently included in the Index. However, this does not necessarily mean there are no limitations with the available data and we aspire to continually improve data across all indicators. There may be drawbacks to data quality stemming, for example, from data which is less current than is desirable, data measuring concepts which are not a perfect fit with the indicator in question, data which is based on subjective assessment and data which uses the wider public sector as a proxy for the civil service. More detail on potential data limitations we have identified are set out in Annex A.

For all other indicators, some degree of data imputation was required. A statistical methodology was adopted to estimate/ 'impute' missing data. Section 5.2 describes the approach to imputing missing data.

5.2 Imputation method

Researchers modeling data often encounter the problem of missing data regarding one or more of the variables under investigation. The most common approach is to remove those observations with missing values, allowing for a complete analysis of those for which full data is available. However, this approach wastes data and reduces power, and also produces biased estimates when the values are not missing completely at random³⁸. One alternative is to use one of the many methods available for imputing the missing values. Of the available methods, multiple imputation (MI) is attractive with theoretical and simulation studies showing that it yields estimates with good statistical properties,

such as efficiency and validity, when a correct model is specified for the imputation³⁹. An imputation represents one set of plausible values for missing data, and so multiple imputations represent multiple sets of plausible values. For studies with roughly 10-60% missing values, past research suggests that multiple imputation is the method of choice.⁴⁰

A number of methods of imputation (including MI) use other variables to predict the missing values. MI requires assumptions to be made about the distribution of the variable and its predictor variables. Each missing value is imputed several times, therefore generating several independent, completed data sets. Each completed data set is analysed and then combined. This imputation method assumes that the data are missing at random, a hypothesis that cannot be verified, since we have no knowledge of the unobserved data.⁴¹ Nevertheless, the more predictors included in the imputation model, the assumption that data is missing at random is more likely to hold because the uncertainty associated with missingness is reduced.⁴²

When a continuous variable contains missing values, a linear regression imputation method can be used to fill in missing values.⁴³ The method of multiple imputation utilised in developing the Index fills in missing values in variables iteratively by using chained equations, a sequence of univariate imputation methods. This requires that a predictive model of the variables with missing values can be specified; the predictive model can include other variables with missing values, while taking into account the problems associated with predictors that have a high degree of missing data. Variables correlated to the variable with missing data need to be utilised in the predictive model.

38 Little R. and Rubin D. (1987), Statistical analysis with missing data.

39 Little R. and Rubin D. (1987), Statistical analysis with missing data.

40 Barzi F. and Woodward M. (2004), Imputations of Missing Values in Practice: Results from Imputations of Serum Cholesterol in 28 Cohort Studies, *American Journal of Epidemiology* 160 (1): 34-45.

41 Ibid.

42 Shafer J.L. (1997), Analysis of incomplete multivariate data.

43 Rubin D. B. (1987), Multiple Imputation for Nonresponse in Surveys; Schenker, N., and Taylor J.M.G. (1996) Partially parametric techniques for multiple imputation. *Computational Statistics & Data Analysis* 22: 425-446

The technique used for undertaking multiple imputation in the Index is multivariate imputation using chained equations (MICE). MICE is very popular in practice. Its popularity stems from the flexibility it offers for imputing the different types of data in observational studies. The variable-by-variable specification of MICE allows practitioners to simultaneously impute variables of different types by choosing from different univariate imputation methods appropriate for each variable. Being able to specify a separate model for each variable provides flexibility in incorporating certain characteristics specific to each variable.^{44,45}

Data from a wider range of countries than those included in the pilot edition of the Index has been utilised for multiple imputation. While the overall availability of data for some countries led to our judgement to exclude them from the first release of the Index, for particular datasets which covered these countries, this data was utilised for imputing values for the 31 countries included in the Index.

For variables where strong correlations with other datasets are not observed, median imputation is used, with values imputed independently of any predictor. With unconditional median imputation, the median value of the variable in question for the countries where data is available is substituted for each missing value. Since all imputations are the same, this method will underestimate the variance for the variable. For those metrics where median imputation has been used, and for the functions and attributes these metrics feed into, caution should be taken when interpreting the results and less weight should be allotted to the relative position of countries. However for interpreting the overall results of the Index we are confident the use of median imputation has not led to significant bias, due to the small proportion of missing data overall (10%) and the small proportion of missing data imputed using this technique (4%).

A mixture of multiple and median imputation is used in the estimation of missing data for the InCiSE Index. Further details and discussion on the implementation of MI are available from a number of sources⁴⁶ and Section 5.3 sets out the method of imputation used to estimate missing data for each metric that feeds into the Index, documenting the correlations observed between variables and explaining the predictive models used for multiple imputation.

44 For more information about multivariate imputation using chained equations, see van Buuren, Boshuizen, and Knook (1999); Raghunathan et al. (2001); van Buuren et al. (2006); van Buuren (2007); White, Royston, and Wood (2011); and Royston (2004, 2005a, 2005b, 2007, 2009), among others.

45 The number of iterations necessary for MICE to converge depends on, among other things, the fractions of missing information and initial values. The higher the fractions of missing information and the farther the initial values are from the mode of the posterior predictive distribution of missing data, the slower the convergence, and thus the larger the number of iterations required. Current literature suggests that in many practical applications a low number of burn-in iterations, somewhere between 5 and 20 iterations, is usually sufficient for convergence. In any case, examination of the data and missing-data patterns is highly recommended when investigating convergence of MICE.

46 StataCorp LP, Stata Multiple-Imputation Reference Manual, Release 13, 2013. For more information about multivariate imputation using chained equations, see van Buuren, Boshuizen, and Knook (1999); Raghunathan et al. (2001); van Buuren et al. (2006); van Buuren (2007); White, Royston, and Wood (2011); and Royston (2004, 2005a, 2005b, 2007, 2009), among others. For more information about the compatibility of conditional specifications, see Arnold, Castillo, and Sarabia (2001); van Buuren (2007); and Arnold, Castillo, and Sarabia (1999) and references therein.

5.3 Imputation approach for indicators

As discussed in section 5.2, the dataset utilised for generating imputed data contains data for countries which are not included in the Index. Data was imputed for all countries in this larger dataset, and as such the approach includes imputation for some metrics where no imputed data is present in the Index, because this covers a more restricted set of countries. Separate imputation models were developed for each indicator and for some indicators a number of imputation models were devised for the different themes. In a small number of cases, metrics that are not featured directly in the indicator framework are used as predictors in imputation models for metrics that do feature in the Index, due to their correlation with these metrics; this is highlighted where relevant.

5.3.1 Functions

Polymaking

For the polymaking indicator, for all the metrics which contribute to the score, data was available for every country in the Index and no imputation was required.

Fiscal and financial management

For the fiscal and financial management indicator, for all the metrics which contribute to the score, data was available for every country in the Index and no imputation was required.

Regulation

Data was available for the metrics included in this indicator across every country in the Index but one. The score for Regulation is calculated based on six metrics from the OECD's 'Indicators of Regulatory Policy and Governance' (iREG). Data was missing for two metrics; Regulatory Impact Assessment – Primary laws and Stakeholder Engagement – Primary Laws.

For this indicator, multiple imputation was used. Our analysis found that a number of Bertelsmann Sustainable Governance Indicators metrics are suitable predictors for OECD iREG metrics. Our imputation model utilised these metrics as complete predictors in addition to the iREG imputation variables.

Crisis/risk management

For the crisis/risk management indicator, all metrics are missing for 8 countries in the Index and have been imputed. No obvious predictor variables were discovered, so simple median imputation was used. This assigns the median value for countries where data was available as the estimate of data for countries where it is missing.

Human resources management

For the human resources management indicator, for all the metrics which contribute to the score, data was available for every country in the Index and no imputation was required.

Tax administration

For tax administration, data is missing for one or more metrics for 11 countries in the Index, although there are no countries where all data is missing.

We found significant correlations between the cost of collection ratios metric and the World Bank's 'Doing Business' Index on the time it takes businesses to pay taxes. We also found significant correlation between the three digital metrics on e-filing of tax returns and the cost of collection ratios. Correlation was weaker between the digital metrics and the time it takes to pay business taxes metric.

The tax debt ratio appears to be correlated with the time it takes businesses to pay taxes but the correlation is not statistically significant. With these observed relationships we proceeded with a multiple imputation approach with all metrics included as predictors, except for the tax debt ratio. The tax debt ratio metric was imputed using simple median imputation.

Social security administration

For social security administration only one metric is currently included in the indicator. Data is missing for 8 countries in the Index and data has been imputed for these countries using multiple imputation. We considered and examined a wide range of metrics outside of this indicator to obtain a suitable predictor and observed a correlation with one of the dimensions of the Quality of Government Expert Survey. We acknowledge that the observed correlation does not identify or imply a causal relationship between these metrics and we wish to identify and develop data to strengthen this indicator as a priority.

Digital services

The score for Digital Services is calculated based on metrics taken from the European Commission's E-Government Benchmarks. All metrics are missing for 8 countries in the Index and have been imputed by multiple imputation. We found that the Online Service Index metric of the UN e-Government Development Index⁴⁷ was generally correlated with all the E-Government Benchmarks metrics and so this metric was utilised as a complete predictor.

5.3.2 Attributes

Integrity

For integrity sixteen metrics are included in the indicator and data is missing for four of the metrics. The Global Corruption Barometer metric measuring the percentage of people viewing public officials/civil servants as corrupt is missing data for 5 countries in the Index. OECD data on whether a post-employment cooling off period exists and whether it's paid, data on the degree of protection against lobbyists and other private interests influencing advisory groups, and data on the number of groups who receive whistleblower protection are also missing for 1 country for the first metric and 7 countries for the latter two metrics.

The imputation approach taken considered these metrics in two groups. Excluding the three OECD metrics with missing data, we found all integrity metrics are highly correlated, noting that the QoG expert assessment of the extent to which public sector employees strive to fulfil ideology of party in government is negatively correlated. On this basis we proceeded with multiple imputation for missing data in this group of metrics with all as predictors for each other.

For the second group, our examination of the data showed that among the three OECD metrics in this group with missing data, data for each metric is fairly clustered around one point and simple median imputation is used.

47 UN E-Government Development Index <https://publicadministration.un.org/egovkb/en-us/Data-Center>

Openness

The approach to imputation for openness treated metrics concerning societal consultation separately from those concerning transparency. There are four metrics on the consultation side with five metrics on the transparency side.

For societal consultation, we observed correlation between the metrics with missing data (the ‘complaint mechanisms’ and ‘civic participation’ components of the Open Government Index) and the metrics with non-missing data (UN’s E-Participation Index and ‘Negotiating public support’ Bertelsmann metric). Multiple imputation was undertaken with these metrics utilised as complete predictors.

For the transparency group of metrics, we found all metrics are highly correlated. On this basis we proceeded with multiple imputation for missing data in this group, with all metrics utilised as predictors for each other.

Capabilities

For capabilities, four metrics are included in the indicator and data is missing for all of these. For 3 of the metrics 10 countries were missing data and for one of the metrics 13 countries were missing.

We examined the relationship of these metrics to others outside of the capabilities indicator. The most suitable predictors observed were the QoG expert assessment of the extent to which the skills and merits of the applicants decide who gets the job when recruiting public sector employees and EU membership; both metrics are positively associated with skills variables. Our imputation model utilised these metrics as complete predictors.

Inclusiveness

For capabilities, five metrics are included in the indicator and data is missing for two of these; the OECD data covering the share of total central government employment filled by women and the share of employment in top management positions within central government filled by women as absolute differences in comparison to each of those equivalent shares of women in the labour force as a whole. For these two metrics data is missing for 10 and 12 countries in the Index respectively. We proceeded with multiple imputation for missing data with all metrics utilised as predictors for each other.

Chapter 6: Index results

.....
In this chapter we describe the approach to the normalisation and reporting of Index results and advise on their interpretation.
.....

6.1 Normalisation

The data sets which feed into the Index capture a range of aspects of the civil service through survey respondents, expert opinion and administrative data. Additionally, the datasets utilise a variety of different measurement scales. To enable comparison of scores and for the purposes of aggregating the data to provide scores for the different functions and attributes, and for the composite Index, we transform the data into comparable units with the same scale – a process known as normalisation.

The normalisation method we employ in our pilot InCiSE Index is the Min-Max method. The Min-Max⁴⁸ process of normalisation preserves the distribution of the data and scales all numeric variables in the range [0,1]. Scoring 1 means that country has the highest score on that metric, 0 the lowest. An example of normalisation using the Min-Max methodology is as follows:

Worked Example

The normalisation method is illustrated below for a metric underpinning the Integrity indicator; the Quality of Government Institute expert assessment of the extent to which public sector employees strive to help citizens.

The first column of the table presents the country scores prior to normalisation.

For each country score, the minimum score observed for the metric (2.32) is subtracted from this score and the result is then divided by the difference between the maximum (6.29) and minimum scores observed. This provides the normalised score for a country. For example Australia's score before normalisation is 5.48. Subtracting the minimum score of 2.32 from this score gives 3.16. The difference between the maximum and the minimum scores is 3.97. Dividing 3.16 by 3.97 gives Australia's normalised score of 0.80 to 2 decimal places.

48 Handbook on Constructing Composite Indicators, OECD

Table 19: Normalisation example

Country	QoG Institute expert assessment of the extent to which public sector employees strive to help citizens	Normalised score
AUS	5.48	0.80
AUT	5.43	0.78
BEL	5.43	0.78
CAN	6.24	0.99
CHL	5.40	0.78
CZE	4.00	0.42
DNK	5.53	0.81
EST	4.63	0.58
FIN	5.33	0.76
FRA	5.09	0.70
DEU	5.03	0.68
GRC	3.60	0.32
HUN	3.73	0.36
IRL	5.00	0.68
ITA	3.70	0.35
JPN	5.75	0.87
KOR	4.74	0.61
MEX	2.32	0.00
NLD	5.21	0.73
NZL	6.29	1.00
NOR	5.57	0.82
POL	5.33	0.76
PRT	4.53	0.56
SVK	3.89	0.40
SVN	4.75	0.61
ESP	4.91	0.65
SWE	5.92	0.91
CHE	5.60	0.83
TUR	3.92	0.40
GBR	5.18	0.72
USA	5.25	0.74

In mathematical terms, each data point x_{qc}^t for a country c and a time t is transformed according to the formula given below:

$$I_{qc}^t = \frac{x_{qc}^t - \min_c(x_q^t)}{\max_c(x_q^t) - \min_c(x_q^t)}$$

where $\min_c(x_q^t)$ and $\max_c(x_q^t)$ are the minimum maximum value of x_{qc}^t across all countries c at time t .⁴⁹

One issue with standardising the range but not the variance can be that, if you have outliers in your data set, normalising your data will scale the ‘normal’ data to a very small interval. Variables with higher variation will have more effect on the final outcome than variables with lower variance. However, analysis of our data has shown the standard deviations of the normalised variables are all between 0.18 and 0.35 (the maximum value for this standardised range would be 0.5).

Looking ahead to future editions of the Index, another variant of the Min-Max method is,

$$I_{qc}^t = \frac{x_{qc}^t - \min_{t \in T} \min_c(x_q^t)}{\max_{t \in T} \max_c(x_q^t) - \min_{t \in T} \min_c(x_q^t)}$$

where the minimum and maximum for each variable are calculated across countries and time, in order to take into account the evolution of the indicators.⁵⁰ In future editions of the index the distinction between developments in absolute and relative performance of countries may be a pertinent analytical question. For example, it is possible for absolute and relative performance to move in opposite directions over time. Adopting such a method may help to address the issue of comparability between years.

This transformation is not stable when data for a new time point becomes available. This implies an adjustment of the time period analysed, T which may affect the minimum and maximum for some variables and hence the values of I_{qc}^t . Maintaining comparability between the existing and new data would require the Index for the previous years to be recalculated.⁵¹

6.2 Reporting of results

For the pilot edition of the Index, results are reported for each metric online, normalised between 0 and 1 according to the method set out in Section 6.1, following imputation where required. Results for each indicator, the 8 functions and 4 attributes, are also reported online and within the Main Report. These indicator results are aggregated from the contributing metrics according to the weightings described in Section 4.2.

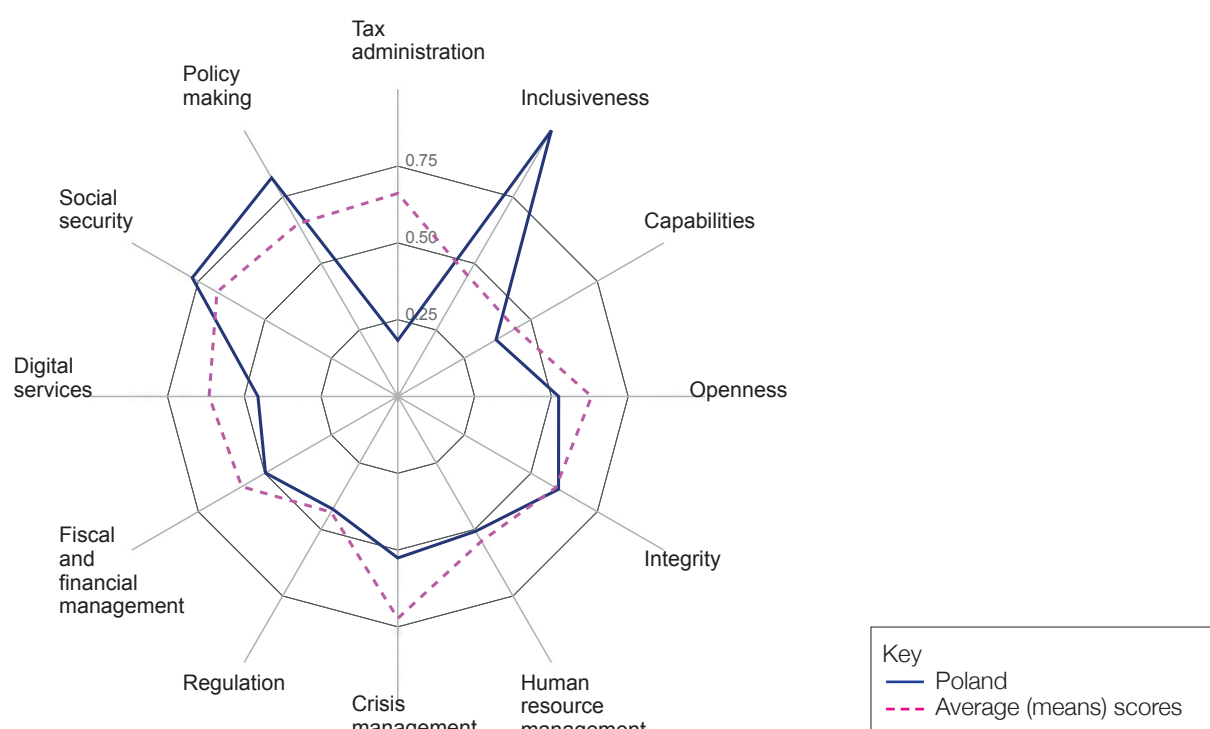
The scores for the composite (the overall) Index are built from the indicator scores according to the weighting approach set out and described in more detail in Section 7.2. These are provided both online and within the Main Report.

The normalised results are presented by way of radar graphs, an example is shown in Figure 8.

49 Ibid.

50 Ibid.

51 Ibid.

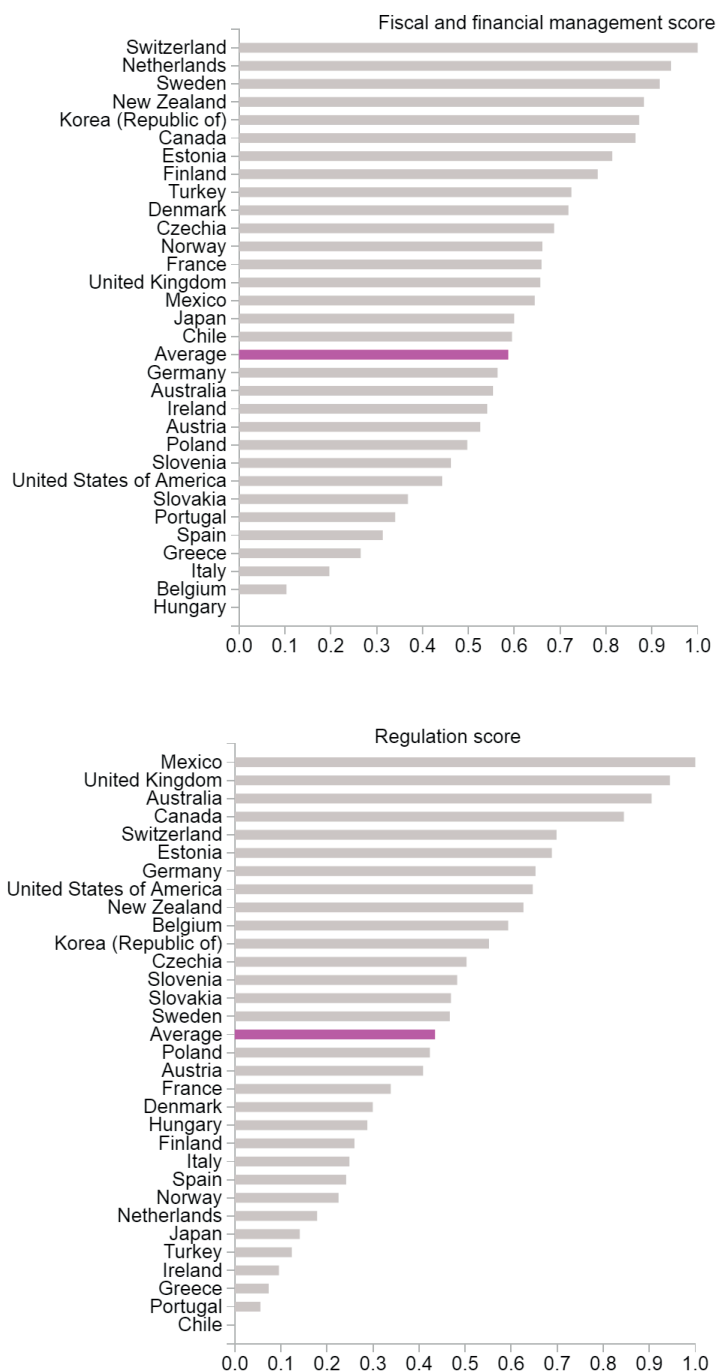
Figure 8 Radar Graph Example – Poland

The normalised data, and radar diagrams drawn from this, allow for an assessment of relative performance of the group of countries included in the Index. So for example, it may allow statements such as 'Country X is assessed to perform more strongly on the Inclusiveness indicator than country Y' or 'Country X is assessed to perform more strongly than the average for all countries in the Index'.

The indicator and composite Index results should be interpreted in the context of the quality and depth of the underlying metrics, and the degree and method of imputation utilised to estimate missing data; these are all described within this Technical report.

A low score does not necessarily mean that a country is performing badly, or that such improvements are more pressing, than for an indicator which has a higher score. The nature of a normalised comparative index means that it is showing a comparative assessment against other countries, not an absolute assessment. As such, relatively weaker scores may aid identification of areas for improvement or learning from more strongly performing countries, but will not indicate priorities for reform and attention on their own. In Figure 8, while the relative score for Crisis Management may be higher than that for Regulation, this does not imply that a country will rank more highly for this indicator. The average relative performance (score) for Crisis Management is higher than for Regulation, such that Poland's score for Crisis Management is somewhat below the average, whereas its score for Regulation is much closer to the average. The higher average score for Crisis Management reflects the grouping of country results which are closer to the leading country than for Regulation.

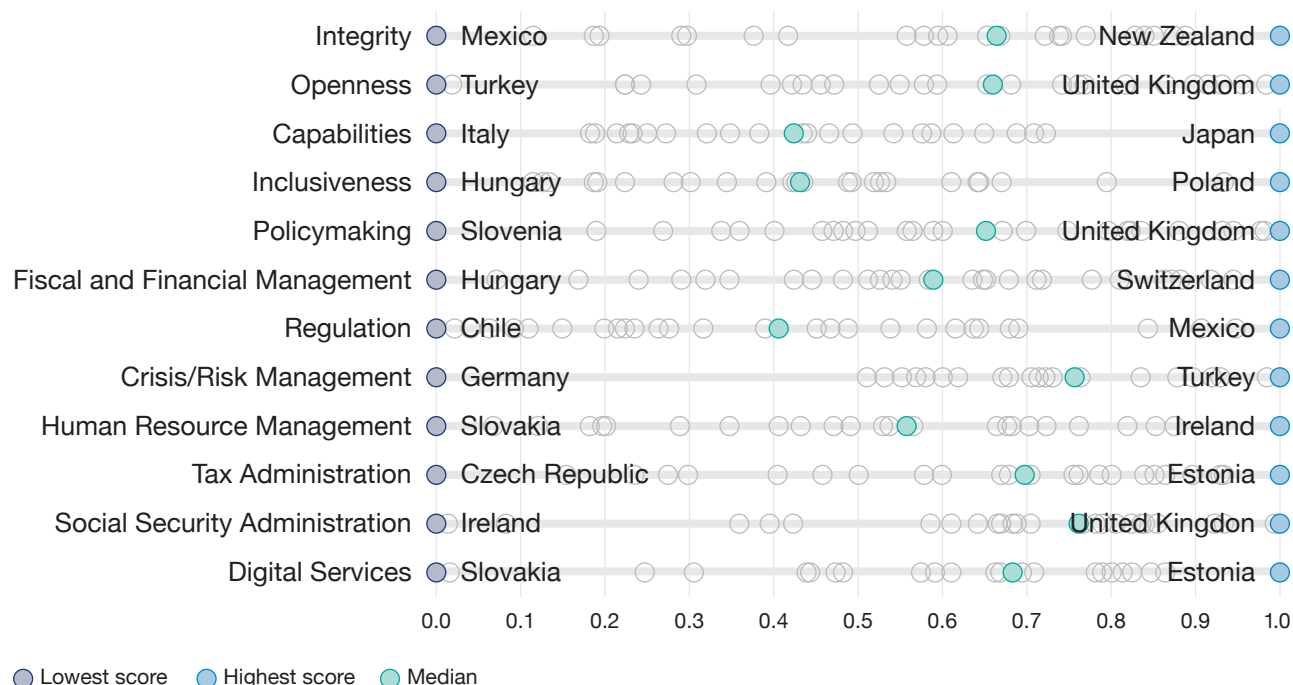
The distribution of results within the Index, indicators and underlying metrics shows the degree of closeness of results in various parts of the ranking, which may indicate whether a particular country is the clear leader (or follower) in performance or whether there is a wider group of relatively strong (weak) countries with similar scores. Information covering the distribution of results is provided in the Main Report, with further detail available in figures 9 and 10, which show the ranking and distribution of indicator scores.

Figure 9 Bar Chart Examples

The bar chart examples demonstrate the variation observed in the distribution of results across the function and attribute indicators. For example, the 10 strongest performing countries for Fiscal and Financial Management are more tightly grouped in their relative performance than the strongest performing countries for Regulation. At the other end of the distribution, scores among the weakest performing countries for Regulation are in fact more tightly grouped than the weakest performing group for Fiscal and Financial Management; this can be observed from the range of scores occupied by the bottom 10 countries for each indicator. The composite Index weights the scores for indicators rather than the ranks.

Figure 10 Distribution of Indicator scores

Figure 10 provides an alternative, summary view, of the distribution of indicator scores. Clustering of scores can be seen for some indicators, while others are more equally spread.



Chapter 7: Composite

.....
In chapter 7 we describe the advantages and disadvantages associated with the use of composite measures. We then describe the construction and weighting of the InCiSE composite and how our approach mitigates against some of the issues with composite indicators that have been identified.
.....

7.1 Strengths and weaknesses of composites

A composite index is the result of compiling individual indicators into a single index in accordance with an underlying framework. The composite can measure multi-dimensional concepts which cannot be captured by a single indicator.

Composite indicators which compare country performance are increasingly recognised as a useful tool in policy analysis and public communication. However, composite indicators can send misleading policy messages if they are poorly constructed or misinterpreted.

The main strengths and weaknesses of using composite indicators are included in Table 20⁵².

Table 20: Strengths and weaknesses of composite indicators.	
Pros:	Cons:
<ul style="list-style-type: none">• Can summarise complex, multi-dimensional realities with a view to supporting decision makers.• Are easier to interpret than a battery of many separate indicators.• Can assess progress of countries over time.• Reduce the visible size of a set of indicators without dropping the underlying information base• Thus make it possible to include more information within the existing size limit; the number of underpinning measures can be much higher than the number of composite measures.• By providing summary measures and visualisations, can place issues of country performance and progress at the centre of the policy arena.• Facilitate communication with general public (i.e. citizens, media, etc.) and promote accountability.• Help to construct/underpin narratives for lay and literate audiences.• Enable users to compare complex dimensions effectively	<ul style="list-style-type: none">• May send misleading policy messages if poorly constructed or misinterpreted.• May reduce complex things to simple rankings which could lead to overly simplistic interpretations.• May be misused, e.g. to support a desired policy, if the construction process is not transparent and/or lacks sound statistical or conceptual principles.• The selection of indicators and weights could be the subject of political dispute.• May disguise serious failings in some dimensions and increase the difficulty of identifying proper remedial action, if the construction process is not transparent.• May lead to inappropriate policies if dimensions of performance that are difficult to measure are ignored.

52 Saisana M. and Tarantola S. (2002), State-of-the-art report on current methodologies and practices for composite indicator development, EUR 20408 EN, European Commission-JRC: Italy

The table sets out some of the most important, and connected, benefits and drawbacks of developing and utilising composite indicators. Our reasoning for utilising a composite indicator can be summarised as follows:

“The key strength of aggregate indicators is their ability to convey information on many parameters succinctly (Booyesen, 2002; Hahn, 2008; Zhou & Ang, 2009; Balica, 2012b). Therefore, composite indices are powerful and communicative tools because they present clear and concise results to non-technical audiences such as scores or rankings (Kenney et al., 2012). That helps to promote a multi-stakeholder dialogue in establishing common understanding of supranational concerns and overcoming socio-political barriers of decision making (Preston et al., 2011: 183). The two main advantages of aggregate measures are:

1. Variables that cannot be directly observed may be inferred by integrating multiple indicators as part of a composite indicator.
2. Composite indices usage helps to overcome the problems of precision, reliability and accuracy by reducing the influence of measurement error as the number of observations from multiple sources increase (Kaufmann & Kraay, 2007; Maggino & Zumbo, 2012).⁵³”

Bearing these potential benefits in mind, we remain aware that even the most carefully constructed of these indices can lack transparency and comparability over time, suffer from selection bias and be of limited use in helping countries to identify how effectively to improve the quality of the civil service.⁵⁴

To account for these limitations and to promote transparency in deploying a composite index, we have taken the following steps:

- i) We describe our favoured approach to constructing a composite Index in Section 7.2.
- ii) We have tested and consulted on the chosen approach widely, refining and adjusting the method over time; seeking the views of a range of multilateral organisations, academic institutions and other experts in the field of public governance, including those most familiar with the usage and drawbacks of composite measures.
- iii) To promote a fuller understanding of the ramifications of different approaches to weighting of indicators, we have also tested the impact of a range of alternative weightings on the Index results. More information about the sensitivity analysis is provided in Chapter 8.
- v) Alongside the composite Index scores and rankings, we present ranks across all indicators in our Main Report. This provides a fuller picture of how the Index is built up from the indicators.

7.2 The InCiSE composite

The InCiSE Index brings together the indicators, each measuring a different dimension of effectiveness into an overall composite indicator of civil service effectiveness.

⁵³ Muriithi et al (2015), Quantifying Governance: An indicator-based approach, Report for DFID by MPA students at the LSE

⁵⁴ Arndt C. and Oman C. (2006), Uses and Abuses of Governance Indicators, OECD Development Centre Studies

Weighting

The measurement framework outlined in Section 3 gives a comprehensive overview of civil service effectiveness. However, existing data does not enable one to measure against all of it. Of the 6 attributes proposed for measurement, so far 4 are included with an individual indicator. Of the 11 core functions proposed, 8 are included with an individual indicator. There are 12 indicators in total. A number of data sources, most with multiple metrics contained within them, underlie the 12 indicators in the Index. In total, 76 variables are spread across these indicators.

A weighted average of scores on each indicator then gives the Index; an overall composite score. Functions make up two-thirds of the overall weight and attributes one-third.

Our preference would be for each attribute and function indicator to be given equal weight within their respective shares of the Index weighting. Due to poorer data quality on two of the attributes, Capabilities and Inclusiveness, their weighting is reduced. A similar adjustment could be made for Tax and Social Security Administration, but given their overwhelming importance as civil service functions it was judged that reducing their weight would not be appropriate. Figure 11 sets out the Index weighting of scores. The Index is more heavily weighted towards functions as a greater number of metrics were available, and these were generally judged to be of better data quality for the purpose of our Index. As a result, core functions were given two-thirds of the overall composite weight, and attributes one-third.

Figure 11: Data weightings in the InCiSE Index

			A	B	C	D
Attributes Weight: 1/3 Metrics: 34		Integrity	1/18	16	1/12	1/9
		Openness	1/18	9	1/12	1/9
		Capabilities	1/18	4	1/12	1/9
		Inclusiveness	1/18	5	1/12	1/9
		Staff Engagement	1/18	–	–	–
		Innovation	1/18	–	–	–
Core functions Weight: 2/3 Metrics: 42	Central Executive	Policymaking	2/33	8	1/12	1/12
		Fiscal and financial	2/33	3	1/12	1/12
		Regulation	2/33	6	1/12	1/12
		Risk/Crisis Management	2/33	9	1/12	1/12
	Mission Support	Procurement	2/33	–	–	–
		HR	2/33	5	1/12	1/12
		IT	2/33	–	–	–
		Finance	2/33	–	–	–
	Direct Service Delivery	Tax Administration	2/33	6	1/12	1/12
		Social Security Administration	2/33	1	1/12	1/12
		Digital Services	2/33	4	1/12	1/12

- A Preferred weight if data on all-indicators was available (equal weight within attributes and functions)
 B Number of data metrics found and included
 C Preferred weight given missing data for some indicators (equal weight within attributes and functions)
 D Data quality-adjusted weight (used in this Index)

At a further level of disaggregation, Section 4 describes how the individual metrics are weighted in each of the indicators they support.

Chapter 8:

Sensitivity Analysis

.....

In chapter 8 we describe the sensitivity analysis undertaken, and the resultant effect on the pilot Index results.

.....

8.1 Introduction to sensitivity analysis

The development of the Index involves stages where subjective judgements have to be made: the selection of individual data sets, the treatment of missing values, the choice of aggregation model and the weights of the indicators, for example.

The quality of the Index depends on the soundness of its assumptions. Good modelling practice requires that we develop an evaluation of the confidence in the model, assessing the uncertainties associated with the modelling process and the subjective choices taken. This role is fulfilled by sensitivity analysis, scrutinising aspects of the relationships between inputs into the Index and its final output.

The approach taken to sensitivity analysis shows how variation in the output can be apportioned, qualitatively or quantitatively, to different sources of variation in the assumptions, and of how the Index depends upon the information fed into it. Sensitivity analysis can help to gauge the robustness of the composite indicator ranking and increases transparency, to identify which countries are assessed more or less favourably under certain assumptions and to help develop a debate around the Index.

In the sensitivity analysis undertaken for the InCiSE Index we focus on four main assumptions: adjusting for GDP per capita, the method of aggregation, exclusion of one limiting dataset at a time and imputation of missing data.

8.2 Sensitivity analysis 1: Adjusting for GDP per Capita

Despite its acknowledged shortcomings, GDP per capita is still the most commonly used summary indicator of living standards. Much of the policy advice provided by international organisations is framed on cross-country comparisons of per capita GDP.⁵⁵ The framework established by the InCiSE Index sets out to measure and compare effectiveness of civil services through an approach that is comparable across countries. However, the economic resources available to develop and maintain the effective institutions and practices of civil services, and governments' ability to raise revenues through taxation, are dependent on the level of economic development and income levels of the economy, as measured by GDP per capita.

⁵⁵ World Economic Outlook, April 2003 – Chapter 3: Growth and Institutions

Research on growth and institutions has sought to identify the deep structural determinants of countries' level of development. The various measures of institutional quality reflect the generally high correlations among themselves and measures of economic performance (Table 21). All appear closely related to cross-country differences in GDP per capita, as well as to other measures of economic performance such as growth rates and the volatility of growth. Building on the close correlation between institutional quality and development, recent analyses attempt to address the possibility of reverse causality from development to institutions, and the relative significance of institutions compared with other influences on development, such as trade openness, geographical factors, and economic policies.⁵⁶

Those countries with higher GDP per capita may also be historically more stable, with greater social capital and more established institutions. While these are recognised as aspects which may support a highly effective civil service, in considering the reverse causality from development to institutions, and the impact of other influences – it is helpful to test the sensitivity of adjusting the Index for GDP per capita for several reasons.

Firstly, it allows us to measure and estimate the importance of income levels for Index scores by assessing how Index scores are affected when the following adjustments are made:

- Scores and relative position are revised to reflect income levels, the extent of the revision depends on the difference between a country's GDP per capita and the average, and also its original score in the Index.

- The method used here is as follows: the country's Index score (prior to normalisation) is divided by its GDP per capita, scores are then normalised between 0 and 1. For example; if two countries had obtained identical scores in the Index but one country (A) had a GDP per capita twice that of the other (B), then after adjusting for GDP, B's score (and position) would be higher than A's.

Secondly, investigating this sensitivity allows for comparison of performance across smaller groups of countries with similar income/GDP per capita characteristics. In future, this may enhance the possibility of the Index serving as a tool to drive performance improvement, for example by allowing decision makers to see which countries perform best at a similar income level, this may support adopting arrangements and improvements of practices that are more transferable between countries of comparable income levels.

Thirdly, this adjustment allows for what might be considered a 'fairer' approach to assessment; assessing effectiveness in light of varying income levels and resources among countries, and estimating whether countries are stronger or weaker performers considering their relative resource constraints. This adjustment may provide insights into capacity for effectiveness or potential for improved effectiveness, and may enhance the impact and take up of the Index among those countries with lower incomes and perhaps lower scores in the core Index.

⁵⁶ Kaufmann, Kraay, and Zoido-Lobaton (1999b); Heritage Foundation (2003); Gurr and Marshall (2000); and World Development Indicators, World Bank (2002)

Table 21: Index rankings comparison after adjustment for GDP per capita				
Country	Index Score	Rank	GDP Adjusted Index Score	Rank after adjustment
Canada	1.00	1	0.76	5
New Zealand	0.95	2	0.87	3
Australia	0.91	3	0.67	9
United Kingdom	0.91	4	0.74	6
Finland	0.87	5	0.71	7
Sweden	0.86	6	0.62	11
Estonia	0.81	7	1.00	1
Norway	0.81	8	0.43	18
Korea, Republic of	0.78	9	0.76	4
United States of America	0.74	10	0.43	19
Denmark	0.73	11	0.52	13
Switzerland	0.61	12	0.32	24
Belgium	0.60	13	0.45	17
Netherlands	0.58	14	0.40	21
Japan	0.57	15	0.52	14
Austria	0.56	16	0.38	22
France	0.50	17	0.42	20
Spain	0.49	18	0.48	15
Mexico	0.47	19	0.95	2
Ireland	0.46	20	0.22	26
Poland	0.44	21	0.58	12
Chile	0.44	22	0.68	8
Slovenia	0.44	23	0.48	16
Germany	0.40	24	0.27	25
Turkey	0.37	25	0.66	10
Portugal	0.31	26	0.36	23
Italy	0.21	27	0.19	27
Czech Republic	0.16	28	0.16	28
Greece	0.04	29	0.05	29
Hungary	0.00	30	0.01	30
Slovakia	0.00	31	0.00	31

8.3 Sensitivity analysis 2: Aggregation Method

The InCiSE index brings together the underlying data sets into ‘indicators’; attributes and functions, each measuring a different dimension of effectiveness. These attributes and functions are then drawn into an overall composite indicator of civil service effectiveness. Attributes contribute one third of the composite weighting and functions make-up the remaining two thirds (see Chapter 7 for further detail and justification for the aggregation approach adopted).

While our approach to aggregation is reasonable and justified, we acknowledge there are a range of possible approaches and weightings which could be followed in building up the composite from these component indicators, with corresponding arguments in favour. Indeed, there are also a range of possibilities for building up the indicators from their constituent datasets (Chapter 4 provides further detail as how the indicators are weighted across the contributing datasets). As part of the sensitivity analysis undertaken, we investigated the impact of alternative aggregation approaches to test the impact on the composite results obtained from the particular approach adopted.

We first looked at the impact of adjusting the weighting between attributes and functions such that each contributes half of the composite weighting. Results are shown in Table 22.

Table 22: Comparison of scores and ranking under adjusted aggregation method. Attributes and Functions each provide half the weighting.

Country	Index score	Rank	Index score adjusted for equal weighting	Rank
Canada	1.00	1	1.00	1
New Zealand	0.95	2	0.98	2
Australia	0.91	3	0.90	3
United Kingdom	0.91	4	0.89	5
Finland	0.87	5	0.89	4
Sweden	0.86	6	0.85	6
Estonia	0.81	7	0.75	8
Norway	0.81	8	0.83	7
Korea, Republic of	0.78	9	0.73	11
United States of America	0.74	10	0.75	9
Denmark	0.73	11	0.74	10
Switzerland	0.61	12	0.64	12
Belgium	0.60	13	0.61	15
Netherlands	0.58	14	0.63	13
Japan	0.57	15	0.61	14
Austria	0.56	16	0.59	16
France	0.50	17	0.51	17
Spain	0.49	18	0.48	19
Mexico	0.47	19	0.37	24
Ireland	0.46	20	0.47	21
Poland	0.44	21	0.49	18
Chile	0.44	22	0.46	22
Slovenia	0.44	23	0.46	23
Germany	0.40	24	0.48	20
Turkey	0.37	25	0.30	26
Portugal	0.31	26	0.31	25
Italy	0.21	27	0.19	27
Czech Republic	0.16	28	0.17	28
Greece	0.04	29	0.08	29
Hungary	0.00	30	0.00	31
Slovakia	0.00	31	0.02	30

Also considered is an alternative weighting where attributes contribute two thirds of the composite, with functions making up the remaining third. Results for this sensitivity are shown in Table 23.

Table 23: Comparison of scores and ranking under adjusted aggregation method. Attributes two thirds weighting and Functions one third.

Country	Index score	Rank	Index score adjusted for two thirds attributes weighting	Rank
Canada	1.00	1	0.99	2
New Zealand	0.95	2	1.00	1
Australia	0.91	3	0.89	4
United Kingdom	0.91	4	0.86	5
Finland	0.87	5	0.92	3
Sweden	0.86	6	0.84	7
Estonia	0.81	7	0.68	11
Norway	0.81	8	0.85	6
Korea, Republic of	0.78	9	0.68	12
United States of America	0.74	10	0.75	9
Denmark	0.73	11	0.76	8
Switzerland	0.61	12	0.67	13
Belgium	0.60	13	0.62	15
Netherlands	0.58	14	0.68	10
Japan	0.57	15	0.64	14
Austria	0.56	16	0.61	16
France	0.50	17	0.53	19
Spain	0.49	18	0.47	23
Mexico	0.47	19	0.27	25
Ireland	0.46	20	0.47	22
Poland	0.44	21	0.54	18
Chile	0.44	22	0.49	20
Slovenia	0.44	23	0.48	21
Germany	0.40	24	0.55	17
Turkey	0.37	25	0.24	26
Portugal	0.31	26	0.31	24
Italy	0.21	27	0.17	28
Czech Republic	0.16	28	0.19	27
Greece	0.04	29	0.11	29
Hungary	0.00	30	0.00	31
Slovakia	0.00	31	0.05	30

From the results of this sensitivity analysis it appears while there is an impact on the Index scores and rankings obtained under the two alternative aggregation methods, this impact is small in relation to the overall Index; most countries obtain very similar scores and rankings under the alternative aggregation approaches when compared to the approach followed in the Index. This is reassuring, and is indicative of the high degree of correlation between the indicators. However, naturally if the weighting possibilities were pushed further towards their limits or extended to the level of the data sets underlying the indicators, more significant impacts on the scores and rankings might be observed; although the merits of such an alternative weighting would likely be in question. Our favoured approach to aggregation appropriately takes into account a broad range of attributes and functions to allow for a reasonable assessment of overall civil service effectiveness.

8.4 Sensitivity analysis 3: Exclusion of Input Data

In Chapter 4 we set out the data underpinning each of the indicators and the weighting given to each to make up the indicator score. Where we have identified limitations in the data used, these are specified. It is important to examine the results of the Index and investigate whether utilising data which has limitations has any skewed effect on the results obtained. We therefore examine the exclusion of three categories of data that have general limiting qualities; data resulting from subjective assessment, data deemed to be out of date, and data where the wider public sector is used to proxy the civil service.

8.4.1 Subjective assessment

As noted earlier, the InCiSE Index is based on a mixture of administrative data and subjective or perceptions-based measures, taken from surveys of households and firms as well as expert assessments produced by various organisations. Our view is that perceptions data have particular value in the measurement of governance. First, perceptions matter because agents base their actions on their perceptions. If citizens believe that the courts are inefficient or the police are corrupt, they are unlikely to avail themselves of their services. Similarly, enterprises base their investment decisions - and citizens their voting decisions - on their perceived view of the investment climate and the government's performance. Secondly, in many areas of the Index framework, there are few alternatives to relying on perceptions data. For instance, this has been particularly the case for Integrity, an attribute in the framework, which almost by definition leaves few records that can be captured by purely objective measures. Additionally even when objective or fact-based data are available, this may capture the notion of laws and arrangements "in theory", which often differs substantially from the reality that exists "in practice". For example, in every one of the 70 countries covered in the 2007 and 2008 waves of the Global Integrity Index, it is formally illegal for a public official to accept a bribe. Yet, despite them being identical when measured in theory, there are large differences across these countries in perceptions of the frequency with which bribes are in fact accepted by public officials.

Despite these advantages, one might nevertheless reasonably be concerned about various potential problems in the interpretation of the subjective data we rely upon in the Index. Broadly such concerns question the extent to which perceptions data adequately capture the relevant reality. A first basic issue is simply that perceptions data on civil service effectiveness are imprecise. This by itself is not surprising – as we have argued above, all measures of effectiveness are necessarily imprecise proxies for the outcomes they are intended to measure, but imprecision alone does not disqualify the use of perceptions-based data.

A potentially more serious concern is that there are various systematic biases in perceptions data on civil service effectiveness. One possibility is that different types of respondents differ systematically in their perceptions of the same underlying reality. For example, it could be the case that business people, represented by owners of the businesses covered in a survey, or the expert assessments provided by commercial business information providers, have different views of what constitutes an effective civil service than other types of respondents, such as households or public sector agencies.

Another possibility is that biases are introduced by the ideological orientation of the organisation providing the subjective assessments of governance. An additional type of bias might be the possibility that subjective assessments of governance are driven by factors other than governance itself, such as the level of development or recent economic performance of a country.

Yet another potential source of bias comes from the possibility that different providers of governance perceptions data rely on each other's assessments, and as a result make correlated perceptions errors. This would undermine the information content in such indicators. Assessing this concern is difficult because the high correlation between governance perceptions rankings from different sources could be due either to perception errors, or due to the fact that these sources are in fact accurately measuring cross-country corruption differences and so necessarily agree with each other.

Concern about these potential sources of bias will be moderated to the extent that excluding underlying data based on subjective assessment affects the resulting Index scores and rankings. The results of this aspect of the sensitivity analysis are demonstrated in Table 24.

Table 24: Sensitivity analysis – Removal of subjective data

Country	Index score	Rank	Adjusted Index Score	Adjusted Rank
Canada	1.00	1	0.94	4
New Zealand	0.95	2	0.95	3
Australia	0.91	3	0.93	5
United Kingdom	0.91	4	0.92	6
Finland	0.87	5	0.87	7
Sweden	0.86	6	0.98	2
Estonia	0.81	7	1.00	1
Norway	0.81	8	0.83	8
Korea, Republic of	0.78	9	0.76	9
United States of America	0.74	10	0.66	11
Denmark	0.73	11	0.70	10
Switzerland	0.61	12	0.62	13
Belgium	0.60	13	0.46	20
Netherlands	0.58	14	0.61	14
Japan	0.57	15	0.48	18
Austria	0.56	16	0.63	12
France	0.50	17	0.48	19
Spain	0.49	18	0.51	17
Mexico	0.47	19	0.57	15
Ireland	0.46	20	0.36	23
Poland	0.44	21	0.24	26
Chile	0.44	22	0.32	25
Slovenia	0.44	23	0.54	16
Germany	0.40	24	0.33	24
Turkey	0.37	25	0.44	21
Portugal	0.31	26	0.37	22
Italy	0.21	27	0.14	28
Czech Republic	0.16	28	0.21	27
Greece	0.04	29	0.02	30
Hungary	0.00	30	0.00	31
Slovakia	0.00	31	0.05	29

8.4.2 Out of date data

The project aims to define a framework comprising the necessary aspects constituting an effective civil service, in a way that can realistically enable international data to be collected to measure against it. The framework developed takes into account a range of functions and attributes of civil services and utilises data from a range of sources. Many of these data sources are released and updated annually, such that on release of the Index, we can be confident that the data utilised presents a recent and highly relevant assessment of the aspect of effectiveness this data underpins. However, some of the data sources are updated less frequently, were released with the scheduled update process still to be confirmed, or were published with no scheduled updates envisioned.

The project aspires to utilise the best available data and where possible, this data should be recent. Where data is not released annually and is less recent, the inclusion of data sets has been based on a judgement of its relevance to the framework despite this drawback. However, inclusion of these datasets may reduce the extent to which the composite reflects recent developments towards greater or reduced civil service effectiveness among the countries included in the Index. A country that has made rapid progress along a number of fronts may be dragged back by the inclusion of a measure which is somewhat out of date and drawn from a period before the effects of such initiatives had been felt. Table 25 shows the impact on the rankings of excluding data released prior to 2013, the year judged as a cut-off point in this sensitivity analysis.

Table 25: Sensitivity analysis – Removal of data pre-2013

Country	Index Score	Rank	Adjusted Index Score	Adjusted Rank
Canada	1.00	1	1.00	1
New Zealand	0.95	2	0.99	2
Australia	0.91	3	0.95	3
United Kingdom	0.91	4	0.92	5
Finland	0.87	5	0.94	4
Sweden	0.86	6	0.88	6
Estonia	0.81	7	0.80	9
Norway	0.81	8	0.88	7
Korea, Republic of	0.78	9	0.75	12
United States of America	0.74	10	0.80	10
Denmark	0.73	11	0.81	8
Switzerland	0.61	12	0.78	11
Belgium	0.60	13	0.73	13
Netherlands	0.58	14	0.71	14
Japan	0.57	15	0.63	16
Austria	0.56	16	0.61	17
France	0.50	17	0.57	18
Spain	0.49	18	0.53	20
Mexico	0.47	19	0.40	23
Ireland	0.46	20	0.69	15
Poland	0.44	21	0.48	22
Chile	0.44	22	0.56	19
Slovenia	0.44	23	0.38	24
Germany	0.40	24	0.52	21
Turkey	0.37	25	0.35	25
Portugal	0.31	26	0.32	26
Italy	0.21	27	0.24	28
Czech Republic	0.16	28	0.26	27
Greece	0.04	29	0.05	29
Hungary	0.00	30	0.04	30
Slovakia	0.00	31	0.00	31

8.4.3 Public sector proxy

The Index is intended to cover the effectiveness of the civil service as defined in Chapter 2, rather than the wider public sector. Occasionally, however, due to availability of data, data covering the wider public sector data is used to proxy for the civil service where this is deemed to be a reasonable proxy. To ensure that the use of such proxies does not significantly distort the Index results, we developed a version of the Index where these proxy measures are excluded. Table 26 compares the results and rankings for this sensitivity.

Table 26: Sensitivity analysis – Removal of public sector proxy data

Country	Index Score	Rank	Adjusted Index Score	Adjusted Rank
Canada	1.00	1	1.00	1
New Zealand	0.95	2	0.87	5
Australia	0.91	3	0.98	2
United Kingdom	0.91	4	0.92	3
Finland	0.87	5	0.84	7
Sweden	0.86	6	0.86	6
Estonia	0.81	7	0.83	8
Norway	0.81	8	0.78	9
Korea, Republic of	0.78	9	0.92	4
United States of America	0.74	10	0.70	11
Denmark	0.73	11	0.72	10
Switzerland	0.61	12	0.55	15
Belgium	0.60	13	0.37	19
Netherlands	0.58	14	0.34	21
Japan	0.57	15	0.25	24
Austria	0.56	16	0.50	16
France	0.50	17	0.60	13
Spain	0.49	18	0.63	12
Mexico	0.47	19	0.59	14
Ireland	0.46	20	0.33	22
Poland	0.44	21	0.22	25
Chile	0.44	22	0.31	23
Slovenia	0.44	23	0.48	17
Germany	0.40	24	0.18	27
Turkey	0.37	25	0.37	20
Portugal	0.31	26	0.46	18
Italy	0.21	27	0.18	26
Czech Republic	0.16	28	0.01	30
Greece	0.04	29	0.06	29
Hungary	0.00	30	0.12	28
Slovakia	0.00	31	0.00	31

8.5 Sensitivity analysis 4: Imputation Method

Chapter 5 describes the statistical methodology adopted to estimate or ‘impute’ missing data. With imputed data included, a complete data set is estimated and the Index can be determined. While the approach to imputation adopted is sensible, based on demonstrated relationships between variables and the proportion of missing data, it is important to consider the impacts of alternative imputation methods.

A range of alternative approaches to imputation are possible, and of course one response to missing data is list-deletion of those countries with missing data, although this would not meet the objective of our chosen imputation approach by allowing the Index to represent a range of countries despite missing data. To examine the impact of the favoured imputation approach we compared the Index calculated to an approach using simple mean imputation to estimate missing data. Mean imputation is a method in which the missing value on a certain variable is replaced by the mean of the available cases. This method maintains the sample size and is easy to use, but the variability in the data is reduced, so the standard deviations and the variance estimates tend to be underestimated. Table 27 documents the comparison of the Index generated under the two imputation methods.

Table 27: Sensitivity analysis – Differing imputation methods

Country	Multiple Imputation Index Score	Multiple Imputation Rank	Mean Imputation Score	Mean Imputation Rank
Canada	1.00	1	1.00	1
New Zealand	0.95	2	0.94	4
Australia	0.91	3	0.91	6
United Kingdom	0.91	4	0.98	2
Finland	0.87	5	0.97	3
Sweden	0.86	6	0.91	5
Estonia	0.81	7	0.87	8
Norway	0.81	8	0.88	7
Korea, Republic of	0.78	9	0.75	10
United States of America	0.74	10	0.69	11
Denmark	0.73	11	0.79	9
Switzerland	0.61	12	0.65	12
Belgium	0.60	13	0.61	16
Netherlands	0.58	14	0.63	13
Japan	0.57	15	0.62	15
Austria	0.56	16	0.63	14
France	0.50	17	0.53	18
Spain	0.49	18	0.55	17
Mexico	0.47	19	0.42	22
Ireland	0.46	20	0.50	19
Poland	0.44	21	0.46	21
Chile	0.44	22	0.42	23
Slovenia	0.44	23	0.49	20
Germany	0.40	24	0.42	24
Turkey	0.37	25	0.37	25
Portugal	0.31	26	0.34	26
Italy	0.21	27	0.20	27
Czech Republic	0.16	28	0.18	28
Greece	0.04	29	0.07	29
Hungary	0.00	30	0.02	30
Slovakia	0.00	31	0.00	31

Chapter 9: Next steps

.....
In chapter 9 we note some limitations of the pilot Index, and planned next steps.

9.1 Identified limitations of the Index

This paper has offered a summary of key aspects of the methodology to construct the pilot Index. When interpreting the results of the Index the limitations listed, which are common to many other international indicators, and planned next steps should be noted. The development of the pilot Index is the culmination of the first stage of an ambitious and innovative effort to develop and measure against a framework of civil service effectiveness. Wide-ranging and sustained engagement from inception to publication has offered challenge, creativity and expertise to support improvement of the Index up to the point of the pilot publication and beyond. As this Index develops, we hope to overcome or reduce the impact of these limitations as well as strengthening the indicator in other areas.

The general limitations of the Index are set out below. Limitations associated with particular datasets included in the Index are identified in Annex A.

Missing data

Only two countries have non-missing values across all 76 metrics. This is because different included data sets cover different sets of countries. However, a number of countries have close to the full complement of data, and missing data for these countries has been imputed. The imputation methods utilised are well supported for the degree of missing data observed and are informed by observed relationships between metrics where possible.

Data Updates

Whilst every effort has been made to make use of the most up-to-date data, each iteration of InCiSE will, by necessity, have a cut-off for data being included in that years Index. InCiSE reports will specify the date up to which data released has been included. For instance, the inaugural pilot report includes data released up to January 2017. The InCiSE Index will be updated annually with the most recent data available.

Varying depth of data

With some data metrics only representing, for example, one survey question, but others representing scores on an entire index relevant to the indicator (e.g. Open Data Barometer score measuring open data), the depth of data in indicators varies.

Public sector performance being used as a proxy for civil service performance

Some of the metrics used in the Index explicitly try to measure performance of the public sector, not just civil service bodies. Therefore until similar data collections are completed for the civil service, these measures can only be considered proxies. The OECD handbook on constructing composite indicators notes that: “Proxy measures can be used when the desired data are unavailable or when cross-country comparability is limited.”

Mixture of quantitative and qualitative/subjective data

The Index mixes quantitative and qualitative data. The OECD handbook on constructing composite indicators notes that “Given a scarcity of internationally comparable quantitative (hard) data, composite indicators often include qualitative (soft) data from surveys or policy reviews.”

Possible ‘spill over’ across indicators

The types of things measured under some of the indicators will occasionally be relevant to others. Attempts have been made to minimise occurrences of this by adapting the framework to ensure each indicator is a unique concept.

Proxy measures of effectiveness

The nature of the true level of civil service effectiveness in a country is inherently unobservable, and therefore any observed empirical measure of the aptitude or capacity will be a proxy for the broader dimensions of effectiveness it reflects. One consequence of this is that our estimates of effectiveness are subject to uncertainty. Users should not over-interpret small differences in performance (across countries or, in future editions, over time) in the aggregate Index and the underlying indicators. The presence of uncertainty does not imply that the Index cannot be used to make meaningful comparisons of effectiveness across countries or over time.

9.2 Priorities for the next 12 months

InCiSE is a long term project and the founding organisations have committed to supporting its development for a further four years. This will include publishing an annual report and developing an interactive website. An International Advisory Panel has also been established to guide this work.

As described earlier, this first edition of the Index is a pilot- further work is required to refine the methodology and make the data more robust. The main issues that the project will focus on over the next 12 months include:

a) Strengthening data collection

Tasks will include: determining how best to collect data on the 5 omitted indicators (staff engagement, innovation, IT, procurement, finance); exploring how data collection for existing indicators can be strengthened, including thematic gaps; expanding data sources and looking at ways of avoiding over reliance on a single survey; reducing instances of public sector performance being used as a proxy for civil service performance; and cross-country applicability issues.

Complementary work, particularly by the OECD, may help to fill some data gaps over time. Additional data collection, for example through the commissioning of specific country surveys, may also help to provide valuable insights.

b) Refining the InCiSE framework

Framework tasks will include: exploring the potential to add new civil service functions (for example, security, foreign affairs, defence and justice); and ensuring each indicator in the framework is a unique concept, to avoid duplication or overlap between the themes being measured.

An issue that has emerged through the pilot is the need to consider how best to adapt the InCiSE framework to enable it to measure and compare core functions within federal government systems where some aspects of this work are often carried out by a lower tier.

Changes to the InCiSE framework will mainly be determined through user feedback and discussions with data owners to learn lessons from their own experiences, as well as through guidance from an International Advisory Panel.

c) Expanding country coverage

InCiSE will explore the scope to expand the Index’s current country coverage over time, including the potential to include some non-OECD countries. Country coverage in future will largely be determined by feedback from countries on the usefulness of having their own set of indicators, and the availability of data to produce results.

Several countries were excluded from the pilot Index because they had fewer than 75% of the metrics available. However, three countries (Bulgaria, Romania and Croatia) had over 70% of data available. A small increase in data collection may help them meet the threshold for inclusion in the next Index.

d) Increasing InCiSE partners and network

The founding institutions will actively encourage more partners to join them and support the development of the InCiSE Index over time, through advice, expertise and funding. The project's International Advisory Panel will also play a key role in promoting InCiSE and encouraging more partners.

The Blavatnik School of Government will host an international conference in Autumn 2017 to discuss the pilot Index with a wide range of interested players. This event will additionally provide a key opportunity to build support for InCiSE and increase its network.

The InCiSE pilot Index has already brought together a rich volume of data and insights. Its launch provides an important opportunity to stimulate wide-ranging discussions at country level, as well as globally, about civil service effectiveness issues generally, as well as the relevance and usefulness of some data.

Feedback is vitally important to help improve the InCiSE Index and the project team would welcome responses from country governments, as well as policy and learning networks, think tanks and academia.

The InCiSE project team will continue to coordinate closely with other institutions who are engaged in similar efforts to measure civil service effectiveness, as well as with the many organisations who have made a vital contribution to InCiSE so far by permitting use of their research or survey data. Continued collaboration will be essential to help strengthen InCiSE in the coming years.

Appendices

- a) Annex A: Data limitations
- b) Annex B: Reference list
- c) Annex C: Reader's Guide – Abbreviations and Country Codes

Annex A: Data limitations

Table 28 sets out the limitations of the data used within the Index. The data is described in chapter 4 of this report.

Table 28: Data limitations within the Index	
Indicator	Limitations
Policy making	<ul style="list-style-type: none"> The data may be capturing the wider public sector rather than the civil service, as the indices are designed to measure government performance. For example, 'Strategic planning' may include the performance of policy-related ministerial decisions and/or academics. The 'Monitoring agencies, bureaucracies' metric includes subnational governments, so is only a proxy for the civil service. A large number of the metrics included are subjective. However, the standardisation process embedded in the Bertelsmann methodology helps to regulate scores across countries. The survey questions give only a partial picture of the quality of policy advice. For example, they do not capture the quality of written and oral briefing, the quality of draft legislation, or the extent to which policy advice is based on evidence.
Fiscal and financial management	<ul style="list-style-type: none"> Some of the data used in this indicator is from 2012. More detailed data on the use and quality of economic appraisal and evaluation would improve the robustness of this indicator. The current metrics only partially measure what we would like to include in this indicator.
Regulation	<ul style="list-style-type: none"> The metrics are mainly based on information provided to the OECD by countries directly. Whilst countries are asked to support the information they provide with evidence, it may be possible that countries exaggerate claims about the quality of their processes.
Crisis/risk management	<ul style="list-style-type: none"> The Hyogo framework gives an overview of best practice in crisis/risk management, but many of the countries in this indicator set appear to already be adhering to large parts of it. This leads to a lack of variation in scores. Ideally this indicator would also be able to measure the quality of response from civil servants after risks have occurred. All metrics included so far only contain details on what risk management strategies and activities exist, with less detail on their quality.
Human Resources Management	<ul style="list-style-type: none"> Data on meritocracy is based on expert opinion, which may not necessarily accurately reflect reality. Data on meritocracy also refers to 'public sector employees'. Attracting talent is measured using expert assessments, but hard quantitative data would be preferable for assessing financial incentives provided. Important themes of Human Resources Management do not have data available, for example the quality of learning and development opportunities. Financial incentives do not give a full picture of how 'attractive' an organisation is to work for.
Tax administration	<ul style="list-style-type: none"> Cost of collection ratios are problematic because amount of revenue collected depends to a large extent on tax rates. Adjusting revenue collected for tax rates may be a potential solution, but how feasible this might be is unknown. Further investigation may be needed into the validity of comparing across different countries' tax systems (e.g. into how it's been decided which taxes should be included in each country).

Table 28: Data limitations within the Index	
Social security	<ul style="list-style-type: none"> • Whilst the included metric may give a good overall picture of efficiency, only having one metric does not give a comprehensive picture of system performance. • Further investigation may be needed into the validity of comparing across different countries' welfare systems.
Digital services	<ul style="list-style-type: none"> • The data only covers mystery-shopper assessments of seven life events so does not assess all of the services which governments typically provide digitally.
Integrity	<ul style="list-style-type: none"> • QoG data on the core values and principles civil servants adhere to is for public sector, not civil servants specifically. OECD data on whistleblower protection also covers the public sector. • A large amount of the data included is subjective. A lot of it relies on expert opinion, which is subject to the choice of experts. Citizen and business perceptions data have the usual limitations which surround perception-based measurement.
Openness	<ul style="list-style-type: none"> • The indices are designed to measure open government in general, so may capture elements of ministerial and wider public sector performance.
Capabilities	<ul style="list-style-type: none"> • Leadership capabilities within the organisation are important, but not currently fully captured. At present this indicator does not cover more specific measures of skills, for example capabilities in specific areas such as economics and statistics. Measures of whether these skills are deployed in the right areas may also be desirable but difficult to obtain. • There is currently no data on the educational attainment of civil services across the world. OECD PIAAC data gives an estimate of public sector educational attainment. Industry codes can be used to get closer to civil service but sample sizes at that level were too small to be reliable.
Inclusiveness	<ul style="list-style-type: none"> • Data is not available on most demographics of interest. • Most OECD data is from 2010. However for a small number of countries data is for 2011 and in one case is for 2009. • The QoG gender data asks experts to estimate gender representation (%), which may not always be accurate. For ethnicity metrics they are asked to score ethnic/religious inclusiveness on a scale of 1 to 7; if management information data were available it would give a more accurate assessment. • This indicator penalises countries for having over-representation of 'disadvantaged' groups. An alternative approach would be to give the same, 'perfect' score if a disadvantaged group is overrepresented.

Annex B: Reference list

- Baubion, C. (2013). OECD Risk Management: Strategic Crisis Management. OECD. [Online]. Available from: <http://www.mmc.com/content/dam/mmc-web/Files/Strategic-Crisis-Management-paper-July-2013.pdf>. [Accessed on 21 February 2017].
- Bouchard, P. and McCrae, J. (2013). Financial Leadership for Government. [Online]. Institute for Government and The Chartered Institute of Management Accountants (CIMA). Available from: [https://www.instituteforgovernment.org.uk/sites/default/files/publications/2739%20IFG%20-%20Financial%20Leadership%20v10%20\(online-version\).pdf](https://www.instituteforgovernment.org.uk/sites/default/files/publications/2739%20IFG%20-%20Financial%20Leadership%20v10%20(online-version).pdf). [Accessed 10 March 2017].
- Bovaird, T. and Löffler, E. ed. (2003). Public Management and Governance. 2nd edn. Routledge.
- Centre for the Study of Governance Innovation (GovInn). (2014). What is Governance Innovation? [Online]. Available from: <http://governanceinnovation.org/what-is-governance/>. [Accessed 7 March 2017].
- Chalam, K.S. ed. (2014). Governance in South Asia: State of the Civil Services. Pp. 59-72. SAGE Publications Ltd.
- Christensen, T., Fimreite, A.L. and Lægreid, P. (2011). Crisis Management: The Perceptions of Citizens and Civil Servants in Norway. *Administration & Society*, 43(5), pp.561-594.
- European Commission. (2006). Welcome to the world of Public Internal Financial Control. [Online]. Available from: http://ec.europa.eu/budget/library/biblio/documents/control/brochure_pifc_en.pdf. [Accessed 10 March 2017].
- Fukuyama, F. (2013). What Is Governance? *Governance: An International Journal of Policy, Administration, and Institutions*, 26(3), pp. 347-368.
- Graham, J., Amos, B. and Plumptre, T. (2003). Principles for Good Governance in the 21st Century. [Online]. Institute on Governance, Policy Brief No.15. Available from: <http://unpan1.un.org/intradoc/groups/public/documents/UNPAN/UNPAN011842.pdf>. [Accessed 17 March 2017].
- Herbert, N. and Healey, J. (2015). Time to Govern Up. [Online]. Govern Up. Available from: <http://www.governup.org/latest/>. [Accessed 22 February 2017].
- Holt, J. and Manning, N. (2014). Fukuyama is Right about Measuring State Quality: Now what? *Governance: An International Journal of Policy, Administration and Institution*, 27(4), pp.717-728.
- Huther, J. and Shah, A. (1999). Applying a Simple Measure of Good Governance to the Debate on Fiscal Decentralization. [Online]. World Bank. Available from: <http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/wps1894.pdf>. [Accessed 1 March 2017].
- International Civil Service Commission. (2002). Standards of conduct for the international civil service. [Online]. Available from: <http://icsc.un.org/resources/pdfs/general/standards.pdf>. [Accessed 15 February 2017].
- International Monetary Fund (IMF). (2016). Consultation on the IMF Guideline on Governance Issues. [Online]. Available from: <https://www.imf.org/external/np/exr/consult/2016/governance/>. [Accessed 9 March 2017].

Kaufmann, D., Kraay, A. and Zoido-Lobaton, P. (1999). Governance Matters. The World Bank Development Research Group Macroeconomics and Growth and World Bank Institute, Governance, Regulation and Finance. Policy Research Working Paper 2196. [Online]. Available from: <https://poseidon01.ssrn.com/delivery.php?ID=492099070121007123122084103071079102064038020065064007072124071027074112085110025086006103020004074093027019087120122087004002065107067088079116012120070113003011075024126024116123112090078&EXT=pdf>. [Accessed 20 February 2017].

Lonti, Z. and M. Woods (2008), Towards Government at a Glance: Identification of Core Data and Issues related to Public Sector Efficiency, OECD Working Papers on Public Governance, No. 7, OECD Publishing. doi:10.1787/245570167540

MacLeod, D. and Clarke, N. (2009). Engaging for Success: enhancing performance through employee engagement. A report to Government. [Online]. Available from: <http://dera.ioe.ac.uk/1810/1/file52215.pdf>. [Accessed 8 March 2017].

Madzova, V., Sajnoski, K. and Davcev, L. (2013). E-Government as an Efficient Tool towards Good Governance (Trends and Comparative Analysis throughout Worldwide Regions and within West Balkan Countries). *Balkan Social Science Review*, (1), pp.157-174.

Magno, F. and Serafica, R. (2001). Information Technology for Good Governance. De La Salle University. [Online]. Available from: <http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan002708.pdf>. [Accessed on 3 March 2017].

McKinnon, R. (2011). Good Governance in Social Security Administration. *International Social Security Review*, 64(4), pp.3-8.

Mulgan, G. (2014). Innovation in the Public Sector – How can public organisations better create, improve and adapt? [Online]. Nesta, version 1. Available from: http://www.nesta.org.uk/sites/default/files/innovation_in_the_public_sector-how_can_public_organisations_better_create_improve_and_adapt.pdf. [Accessed 7 March 2017].

Muriithi, K., Jimenez, M., Jannin, N., Sajid, N., Singh, S. and Sharma, S. (2015). Quantifying Governance: An indicator-based approach. The London School of Economics and Department for International Development Capstone Team. [Online]. Available from: https://assets.publishing.service.gov.uk/media/57a08971ed915d622c000209/61506_Capstone_Report_DFID_FINAL_Quantifying_Governance_Indicators.pdf. [Accessed 21 February 2017].

OECD. (2007). Towards Better Measurement of Government. [Online]. OECD Working Papers on Public Governance, 2007(1). OECD Publishing. Available from: <http://www.oecd.org/governance/digital-government/38134037.pdf>. [Accessed 10 March 2017].

OECD. (2012). Recommendation of the Council on Regulatory Policy and Governance. The OECD Regulatory Policy Committee. [Online]. Available from: <http://www.oecd.org/governance/regulatory-policy/49990817.pdf>. [Accessed 9 March 2017].

OECD. (2014). Recommendations of the Council on the Governance of Critical Risks. Meeting of The OECD Council at Ministerial Level. Paris 6-7 May 2014. [Online]. Available from: <http://www.oecd.org/gov/risk/Critical-Risks-Recommendation.pdf>. [Accessed on 2 March 2017].

OECD. (2015a). Recommendation of the council on Budgetary Governance. Public Governance & Territorial Development Directorate. [Online]. Available from: <http://www.oecd.org/gov/budgeting/Recommendation-of-the-Council-on-Budgetary-Governance.pdf>. [Accessed on 3 March 2017].

OECD. (2015b). Inclusive government for a more inclusive society. Government at a Glance 2015, OECD Publishing, Paris. [Online]. Available from: http://dx.doi.org/10.1787/gov_glance-2015-6-en. [Accessed on 21 February 2017].

OECD. (2015c). The Innovation Imperative in the Public Sector: Setting an Agenda for Action, OECD Publishing, Paris. [Online]. Available from: <http://dx.doi.org/10.1787/9789264236561-en>. [Accessed 7 March 2017].

OECD. (2015d). Public Sector Leadership for Improved Employee Engagement and Organisational Success. Lead-Engage Perform. OECD expert meeting, 21-22 January 2015. [Online]. Available from: <https://www.oecd.org/gov/pem/agenda.pdf>. [Accessed on 3 March 2017].

OECD. (n.d.). Building more effective, accountable, and inclusive institutions for all. [Online]. Post 2015 Millennium Development Goals: Element 6, Paper 1. Available from: https://www.oecd.org/dac/_POST-2015%20effective%20and%20accountable%20institutions.pdf. [Accessed 1 March 2017].

Office for Public Management Ltd (OPM) and The Chartered Institute of Public Finance and Accountability (CIPFA). (2004). The Good Governance Standard for Public Services. The Independent Commission on Good Governance in Public Services. Hackney Press Ltd.

Phillips, W., Caldwell, N. and Callender, G. (2007). Public Procurement – a pillar of good governance? [Online]. Available from: https://www.researchgate.net/publication/272487779_Public_procurement_A_pillar_of_good_governance. [Accessed on 22 February 2017].

The Committee on Standards in Public Life. (1996). Summary of the Nolan Committee's First Report on Standards in Public Life. [Online]. The National Archives. Available from: <http://webarchive.nationalarchives.gov.uk/20140131031506/http://www.archive.official-documents.co.uk/document/parliament/nolan/nolan.htm>. [Accessed 17 March 2017].

The World Bank. (2012). Indicators of the Strength of Public Management Systems: A key part of the Public Sector Management results story. PRMPS Discussion Paper: Draft. [Online]. Available from: <http://siteresources.worldbank.org/PUBLICSECTORANDGOVERNANCE/Resources/285741-1354024300711/ISPMS.pdf>. [Accessed 21 February 2017].

The World Bank. (2016). Benchmarking Public Procurement 2016. [Online]. Available from: <http://bpp.worldbank.org/~media/WBG/BPP/Documents/Reports/Benchmarking-Public-Procurement-2016.pdf>. [Accessed 22 February 2017].

The World Bank (n.d.). [Online]. Available from: <http://www1.worldbank.org/publicsector/anticorrupt/corruptn/cor02.htm>. [Accessed 1 March 2017].

United Nations. (1945.) Charter of the United Nations. Article 101. [Online]. Available from: <http://www.un.org/en/sections/un-charter/un-charter-full-text/>. [Accessed 15th February 2017].

United Nations Public Administration. (1999). Transparency in Government. Presentation for ILEA Seminar. [Online]. Available from: <http://unpan1.un.org/intradoc/groups/public/documents/un/unpan012062.pdf>. [Accessed 20 February 2017].

Van Dooren, W., Manning, N., Malinska, J., Kraan, D. J., Sterck, M. and Bouckaert, G. (2006). OECD project on Management in Government: Comparative Country Data. Issues in Output Measurement for "Government at a Glance". OECD, Technical paper 2 (Second draft).

World Trade Organisation. (2015). Government Procurement Agreement: opening markets and promoting good governance. [Online]. Available from: https://www.wto.org/english/thewto_e/20y_e/gpa_brochure2015_e.pdf. [Accessed 1 March 2017].

Annex C – Reader’s Guide

C.1 Abbreviations

InCiSE	International Civil Service Effectiveness
SNA	System of National accounts
OPM	Office for Public Management
CIPFA	Chartered Institute of Public Finance and Accountability
ISPMS	Indicators of the Strength of Public Management Systems
RIA	Regulatory Impact Assessment
OECD	Organisation for Economic Co-operation and Development
IMF	International Monetary Fund
UN	United Nations
VFM	Value for money
GCI	Global Competitiveness Index
WEF	World Economic Forum
iREG	Indicators of Regulatory Policy and Governance
QoG	Quality of Government
SGI	Sustainable Governance Indicators
EPI	E-participation index
ODI	Open Data Index
ODB	Open Data Barometer
PIAAC	Programme for the International Assessment of Adult Competencies
GaaG	Government at a Glance
MI	Multiple imputation
MICE	Multivariate imputation using chained equations
OURdata	Open, Useful, Reusable data

C.2 Country codes

The following ISO country codes are used in some tables of this report⁵⁷:

Country	ISO alpha – 3 digit country code	Country	ISO alpha – 3 digit country code
Australia	AUS	Korea, Republic of	KOR
Austria	AUT	Mexico	MEX
Belgium	BEL	Netherlands	NLD
Canada	CAN	New Zealand	NZL
Chile	CHL	Norway	NOR
Czech Republic	CZE	Poland	POL
Denmark	DNK	Portugal	PRT
Estonia	EST	Slovakia	SVK
Finland	FIN	Slovenia	SVN
France	FRA	Spain	ESP
Germany	DEU	Sweden	SWE
Greece	GRC	Switzerland	CHE
Hungary	HUN	Turkey	TUR
Ireland	IRL	United Kingdom	GBR
Italy	ITA	United States of America	USA
Japan	JPN		

⁵⁷ <https://www.iso.org/obp/ui/#home>

