



DRAFT



# Measuring and improving state sector productivity

Draft report



# Measuring and improving state sector productivity

Draft report

December 2017

## The New Zealand Productivity Commission

Te Kōmihana Whai Hua o Aotearoa<sup>1</sup>

**How to cite this document:** New Zealand Productivity Commission. (2017). *Measuring and improving state sector productivity: Draft Report*. Available from [www.productivity.govt.nz/current-inquiries](http://www.productivity.govt.nz/current-inquiries)

Date: December 2017

The Commission – an independent Crown entity – completes in-depth inquiry reports on topics selected by the Government, carries out productivity-related research and promotes understanding of productivity issues. The Commission aims to provide insightful, well-informed and accessible advice that leads to the best possible improvement in the wellbeing of New Zealanders. The New Zealand Productivity Commission Act 2010 guides and binds the Commission.

You can find information on the Commission at [www.productivity.govt.nz](http://www.productivity.govt.nz), or by calling +64 4 903 5150.

### Disclaimer

The contents of this report must not be construed as legal advice. The Commission does not accept any responsibility or liability for an action taken as a result of reading, or reliance placed because of having read any part, or all, of the information in this report. The Commission does not accept any responsibility or liability for any error, inadequacy, deficiency, flaw in or omission from this report.

ISBN: 978-1-98-851901-2 (print)

ISBN: 978-1-98-851902-9 (online)

### Inquiry contacts

#### Administration

Robyn Sadlier

T: (04) 903 5167

E: [info@productivity.govt.nz](mailto:info@productivity.govt.nz)

**Website** [www.productivity.govt.nz](http://www.productivity.govt.nz)

**Twitter** @nzprocom

**LinkedIn** NZ Productivity Commission

#### Other matters

Judy Kavanagh

Inquiry Director

T: (04) 903 5165

E: [judy.kavanagh@productivity.govt.nz](mailto:judy.kavanagh@productivity.govt.nz)

---

<sup>1</sup> The Commission that pursues abundance for New Zealand

# Terms of reference

Issued by the Minister of Finance (the “referring Minister”). Pursuant to sections 9 and 11 of the New Zealand Productivity Commission Act 2010, I hereby request that the New Zealand Productivity Commission (“the Commission”) undertake an inquiry into how the New Zealand State sector can effectively measure and improve productivity in core public services, with a particular focus on health, education, justice and social support.

## Context

Improving the productivity of the state sector, the value we are realising from our resources, helps improve the prosperity of the country, and allows for better outcomes to be achieved from scarce tax payer resource.

Recent progress has been made in improving value across the different dimensions of value for money performance. The Better Public Services Results determine priority areas for improvement. Social Investment and other effectiveness work is getting better at identifying where to invest and tracking what the impact of investment is.

A third dimension of performance is efficiency/productivity. For many of the core public services that constitute a large proportion of existing expenditure, there are still opportunities to better understand efficiency and how to optimise inputs/resources in delivering quality products and services. Current gaps in good measures of productivity limit assurance Ministers have on performance and innovation of current delivery models, and Chief Executives ability to understand and improve their business. It also suggests an opportunity to achieve more from current resources, and better engage the State sector workforces on opportunities to do things better.

Public services are often complex covering a range of services, clients, and different mechanisms to achieve a range of desired outcomes. This can make it more difficult than private sector industries to capture performance, and to take actions to improve it. Internationally, there are few common productivity measures that capture quality dimensions in key sectors like education and health. But, there are lessons on how to better understand dimensions like quality in inputs and outputs, leverage innovation and economies of scale, and improve productivity and efficiency in the public sector.

## Scope

The Productivity Commission (the Commission) is to consider New Zealand and international public and private sector best practice in understanding and improving productivity. This should focus on the narrower definition of productivity as how efficiently inputs/resources are being utilised to generate quality outputs/services.

The Commission should take account of broader definitions of performance and productivity, in considering how to capture elements like quality, and how efficiency measures can complement dimensions like effectiveness. However, the Commission should not focus advice on the contribution of services to longer-term outcomes, prioritisation of interventions, or other performance dimensions already being developed through social investment or other work programmes.

The inquiry should focus on developing practical guidance and recommendations that consider perspectives and roles of different state sector decision-makers such as Ministers, Chief Executives, and managers, and how these different needs can be balanced.

The inquiry is to focus on guidance that is relevant to decision-makers across the “core” services in the health, education, justice and social development sectors, such as: teaching, hospitals and primary healthcare, policing, courts, corrections, and work and income services.

Having regard to the above, the Commission should undertake an inquiry that considers and provides advice on:

- a) How to measure efficiency/productivity in each of the identified core public service sectors: health, education, justice, social support. This should focus on meso (sector) and micro (function or service) level measures. Guidance should consider key measurement and accuracy issues, and how imperfect measures are most appropriately and usefully employed.
- b) The appropriate role of identified efficiency/productivity measures in public sector performance frameworks, with the goal of improving assurance to Ministers and incentives on agencies for improvement. This should draw on theory and evidence of incentive and disincentive effects of measurement and other performance approaches on different workforces.
- c) Developing the capability, culture and systems that can support agencies to better measure, understand and improve productivity.

The Commission should prioritise its effort by using its judgement as to the degree of depth and sophistication of analysis it applies to satisfy each part of the Terms of Reference; and to the degree of depth in each specific sector, while providing advice on best measures in the identified sectors.

### **Exclusions**

The Commission should not carry out in depth analysis or provide detailed recommendations on specific policies relating to service access or provision in sectors.

The Commission should not duplicate work on issues like where to invest, or service effectiveness, being developed as part of the social investment approach.

### **Consultation requirements**

In undertaking this inquiry the Commission should consult with key interest groups and affected parties relevant to the identified sectors and particular services where efficiency measures are identified. Consultation should include public sector agencies, those in receipt of public services, and private sector agents who may have relevant insights.

### **Timeframe**

The Commission must publish a draft report and/or discussion document, for public comment, followed by a final report that must be presented to the Minister of Finance as Referring Minister by 30 August 2018.

HON STEVEN JOYCE, MINISTER OF FINANCE

# About the draft report

This draft report aims to assist individuals and organisations to participate in the inquiry. It outlines the background to the inquiry, the Commission's intended approach, and the matters about which the Commission is seeking comment and information.

This draft report contains the Commission's draft findings and a limited number of questions to which responses are invited but not required. The Commission welcomes information and comment on any part of this report and on any issues that participants consider relevant to the inquiry's terms of reference.

## Key inquiry dates

Submissions due on the draft report	1 March 2018
Engagement with interested parties on the draft report	February 2018
Final report to the Government	30 August 2018

## Why you should register your interest

The Commission seeks your help in gathering ideas, opinions and information to ensure this inquiry is well informed and relevant. The Commission will keep registered participants informed as the inquiry progresses.

You can register for updates at [www.productivity.govt.nz/subscribe-to-updates](http://www.productivity.govt.nz/subscribe-to-updates), or by emailing your contact details to [info@productivity.govt.nz](mailto:info@productivity.govt.nz).

## Why you should make a submission

Submissions provide information to the inquiry and help shape the Commission's recommendations in the final report to the Government. Inquiry reports will quote or refer to relevant information from submissions.

## How to make a submission

The due date for submissions in response to this report is **1 March 2018**. Late submissions will be accepted, but lateness may limit the Commission's ability to consider them fully.

Anyone can make a submission. Your submission may be written or in electronic or audio format. A submission may range from a short letter on one issue to a substantial response covering multiple issues. Please provide relevant facts, figures, data, examples and documents where possible to support your views. The Commission welcomes all submissions, but multiple, identical submissions will not carry more weight than the merits of your arguments. Your submission may incorporate relevant material provided to other reviews or inquiries.

Your submission should include your name and contact details and the details of any organisation you represent. The Commission will not accept submissions that, in its opinion, contain inappropriate or defamatory content.

### Sending in your submission

Web: [www.productivity.govt.nz/make-a-submission](http://www.productivity.govt.nz/make-a-submission)

Email: [info@productivity.govt.nz](mailto:info@productivity.govt.nz)

Post: State Sector Productivity Inquiry  
New Zealand Productivity Commission  
PO Box 8036  
The Terrace  
WELLINGTON 6143  
New Zealand

The Commission appreciates receiving an electronic copy of posted submissions, preferably in Microsoft Word or searchable PDF format. Please email the files to [info@productivity.govt.nz](mailto:info@productivity.govt.nz).

## What the Commission will do with the submissions

The Commission seeks to have as much information as possible on the public record. Submissions will become publicly available documents on the Commission's website. This will occur shortly after receipt, unless your submission is marked "in confidence" or you wish to delay its release for a short time. Please contact the Commission before submitting "in confidence" material, as it can only accept such material under special circumstances.

## Other ways you can participate

The Commission welcomes feedback about its inquiry. Please email your feedback to [info@productivity.govt.nz](mailto:info@productivity.govt.nz) or contact the Commission to arrange a meeting with inquiry staff.



# Contents

<b>Terms of reference .....</b>	<b>iii</b>
<b>About the draft report.....</b>	<b>v</b>
Key inquiry dates.....	v
Why you should register your interest.....	v
Why you should make a submission.....	v
How to make a submission.....	v
What the Commission will do with the submissions.....	vi
Other ways you can participate.....	vi
<b>Overview .....</b>	<b>1</b>
What this inquiry is about.....	1
What is state sector productivity?.....	1
Why state sector productivity matters.....	1
Measuring productivity in public services.....	2
Moving forward with measurement.....	3
Encouraging productivity improvements.....	4
<b>1 State sector productivity – what it is and why it matters.....</b>	<b>5</b>
1.1 The Commission’s task.....	5
1.2 The New Zealand state sector.....	6
1.3 What is productivity?.....	6
1.4 Why is state sector productivity important?.....	7
1.5 What are the barriers to measuring and improving state sector productivity?.....	8
1.6 Recent experience.....	12
1.7 What is needed to achieve future gains?.....	16
1.8 The structure of this report.....	17
<b>2 Measuring state sector productivity .....</b>	<b>19</b>
2.1 International practice in measuring state sector productivity.....	20
2.2 Measures of state sector productivity in New Zealand.....	25
2.3 Key steps in developing a productivity measure.....	27
2.4 What kind and level of productivity measure?.....	28
2.5 Selecting a measurement technique.....	32
<b>3 Measuring outputs.....</b>	<b>35</b>
3.1 Introduction.....	35
3.2 Defining outputs.....	35
3.3 Selecting which outputs to measure.....	37
3.4 Challenges in measuring outputs.....	39
3.5 Calculating a total output metric.....	43
3.6 Conclusion.....	46
<b>4 Selecting and costing inputs.....</b>	<b>47</b>
4.1 Introduction.....	47
4.2 Types of inputs and attribution.....	47
4.3 How to calculate inputs?.....	49
4.4 Systems to capture and allocate inputs.....	51
4.5 Co-financing, missing inputs and costs.....	53
4.6 Summary.....	55
<b>5 Adjusting for operating environments, quality and price .....</b>	<b>56</b>
5.1 Introduction.....	56
5.2 Why adjustments for operating environments, quality and price matter.....	56
5.3 Differences in operating environments.....	57
5.4 Quality adjustment.....	59

5.5	Accounting for price changes .....	64
5.6	Conclusion.....	68
<b>6</b>	<b>Putting productivity measures in place.....</b>	<b>69</b>
6.1	Introduction.....	69
6.2	Productivity is a companion to effectiveness.....	70
6.3	Much to build off, but also significant gaps.....	70
6.4	Why so little interest in or analysis of productivity?.....	72
6.5	How to make progress .....	77
6.6	Conclusion.....	84
<b>7</b>	<b>Capability and systems needed to measure productivity .....</b>	<b>85</b>
7.1	Introduction.....	85
7.2	Productivity measures do not have to be sophisticated or complete to be useful .....	86
7.3	Understanding an agency’s operational business systems and practices .....	87
7.4	Data analytic and interpretive skills are needed.....	89
7.5	Leaders need to drive productivity measurement .....	92
7.6	Sound financial management systems.....	94
7.7	Robust and integrated data systems and tools .....	96
7.8	Data use rules, protocols and protections.....	99
7.9	Using the Performance Improvement Framework .....	101
7.10	Conclusion.....	102
<b>8</b>	<b>Encouraging state sector productivity improvement.....</b>	<b>103</b>
8.1	Sources of productivity improvement in the public sector .....	104
8.2	Internal drivers .....	106
8.3	External drivers .....	116
8.4	Organisational choices .....	125
8.5	Data and ICT .....	129
8.6	Conclusion.....	137
	<b>Summary of questions .....</b>	<b>138</b>
	<b>Findings and recommendations .....</b>	<b>140</b>
	<b>Appendix A Public consultation .....</b>	<b>144</b>
	Submissions.....	144
	Engagement meetings .....	144
	Roundtable .....	145
	Conferences and seminars .....	145
	<b>References.....</b>	<b>146</b>

## Tables

Table 1.1	Challenges measuring service productivity.....	8
Table 1.2	Productivity case studies .....	18
Table 2.1	Overseas studies of productivity in specific public services – selected examples.....	24
Table 2.2	Purposes of performance measurement .....	29
Table 2.3	Levels at which public sector productivity analysis can be conducted.....	29
Table 2.4	Different decision-makers and types of productivity measures.....	30
Table 2.5	Identification of possible measurement techniques.....	33
Table 3.1	Outputs and outcome in the state sector .....	36
Table 4.1	Outputs and possible inputs in the state sector .....	48
Table 4.2	Options for calculating labour input.....	49
Table 4.3	iCAM Service Delivery components .....	52
Table 5.1	Age of Impatient Events, 2002-2014 .....	57
Table 5.2	Possible approaches to quality adjusting data on schools .....	60
Table 5.3	Possible approaches to adjusting health sector data.....	62
Table 5.4	Possible approaches to adjusting data in other sectors.....	63
Table 5.5	Quality Criteria for Deflators for Government Services .....	65

Table 6.1 Budget 2017 outputs with efficiency-related performance measures ..... 74  
 Table 7.1 Number of domestic students completing Bachelor's degrees by selected field of study ..... 91  
 Table 7.2 Departments' PIF organisational management ratings (selected measures)..... 101  
 Table 8.1 Baseline and output price reviews, 1996-2004 ..... 121  
 Table 8.2 Potential mechanisms to motivate performance ..... 122

**Figures**

Figure 0.1 Forecast government expenditure, as a share of GDP ..... 2  
 Figure 1.1 The relationship between inputs, outputs and outcomes ..... 7  
 Figure 2.1 Countries measuring public sector productivity (from a sample of 32) ..... 22  
 Figure 2.2 Labour and multi-factor productivity indexes, 1996–2015..... 25  
 Figure 2.3 Education productivity estimates in New Zealand and the United Kingdom, 1997–2014 ..... 27  
 Figure 2.4 Main steps in developing a productivity measure ..... 28  
 Figure 3.1 Matching outputs with the level of analysis..... 38  
 Figure 3.2 Typology of state sector outputs ..... 40  
 Figure 4.1 NZ Police “Know Your Business” model ..... 53  
 Figure 6.1 Mean scores for selected learning culture questions by sector ..... 80  
 Figure 6.2 Mean scores against selected innovation questions by sector ..... 81  
 Figure 7.1 Core capability and systems required to measure productivity..... 86  
 Figure 7.2 A shift in thinking from transactional to strategic financial management ..... 96  
 Figure 8.1 Reported sources of productivity improvement in public services, selected literature ..... 106  
 Figure 8.2 Barriers to innovation in the public sector..... 107

**KEY**



Questions



Findings



# Overview

## What this inquiry is about

The terms of reference ask the New Zealand Productivity Commission to undertake an inquiry into measuring and improving productivity in public services, and to provide guidance and recommendations on:

- how to measure productivity in “core” public services (health, education, justice, social support) at sector and service level;
- what role productivity measures should play in public sector performance frameworks; and
- how to develop the culture, capability and systems needed within government agencies to measure, understand and improve productivity.

The Commission was asked to not provide detailed recommendations on service provision or access in the core sectors, and to not duplicate work on where to invest in public services.

## What is state sector productivity?

Productivity is a measure of how well an organisation uses its resources (eg, money, labour, capital) to produce goods and services. Properly measured, productivity considers changes in both the volume and the quality of public services provided.

Productivity leads to greater efficiency. Efficiency can be defined in several ways:

- technical efficiency – the optimal method of producing services (“doing things right”);
- allocative efficiency – the optimal distribution of resources to produce the right set of services (“doing the right things”); and
- dynamic efficiency – the structures, behaviour and incentives that create improvements in allocative and technical efficiency over time (“doing things right and the right things over time”).

The Commission focused primarily on technical efficiency, but these concepts do overlap. A system focused on delivering public value needs to be concerned with all three.

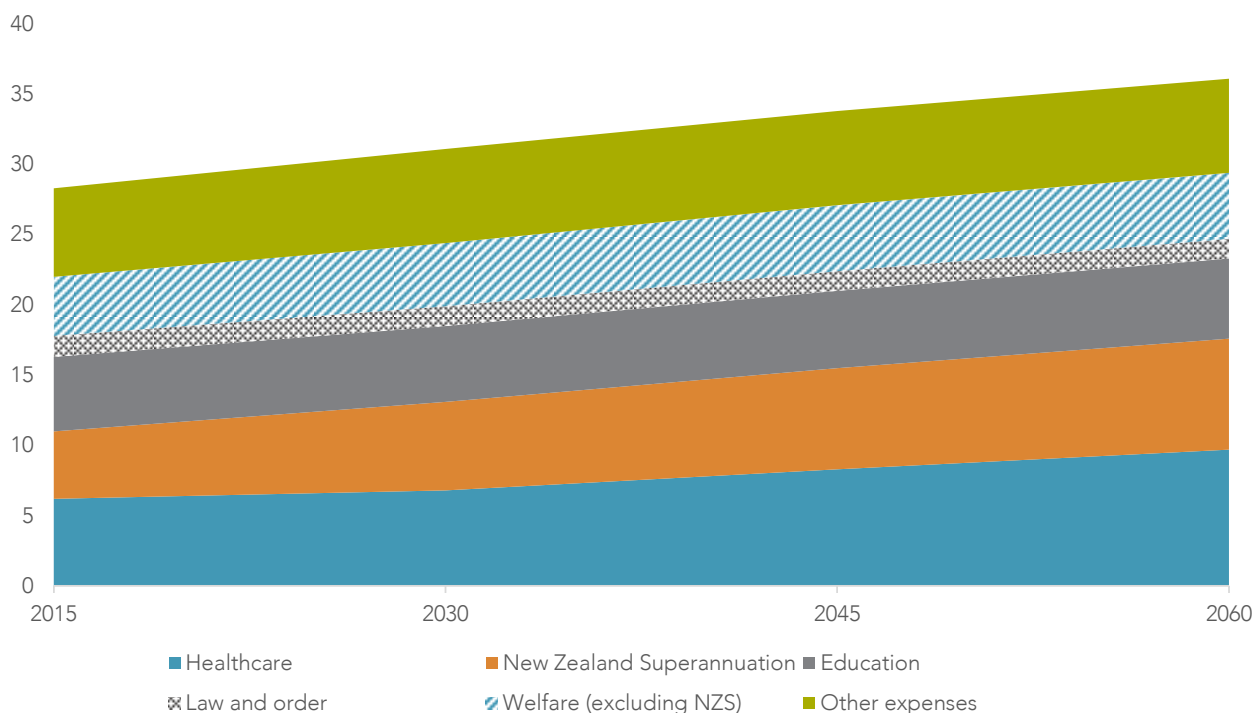
## Why state sector productivity matters

### Government makes up a large part of the economy

The state sector spends more than \$40 billion every year on public services (New Zealand Treasury, 2017).<sup>2</sup> Further, the New Zealand Treasury forecasts total government spending (as a proportion of GDP) to grow (Figure 0.1). Higher productivity in the state sector helps ensure the nation’s resources are being well-used and are making the greatest contribution to wellbeing.

---

<sup>2</sup> Government expects to spend \$41 billion on services delivered by the Justice, Health, Social Development and Education sectors in 2017/18 (excluding transfer payments such as New Zealand superannuation).

**Figure 0.1 Forecast government expenditure, as a share of GDP**

Source: New Zealand Treasury, 2016.

### Higher productivity provides more choices and greater sustainability

The more productive the state sector is, the more public services New Zealanders will receive for a given level of government spending. Higher public sector productivity opens up choices about how the revenue freed up by productivity is used – for example, for additional or new services, or returned to New Zealanders through transfers or tax cuts.

A more productive state sector will be better-placed to respond to the pressures that will arise as New Zealand's population ages. Higher productivity is a key way of ensuring the community can afford the services it values into the future.

### Greater public service productivity promotes better economic performance

In addition to enhancing individual wellbeing, improved public sector productivity also matters for the private sector. For example, publicly-funded education and training increases the supply of skilled workers and entrepreneurs. Police and court services prevent economic loss from crime and support the fair and timely resolution of business and contractual disputes. Therefore, productivity improvements in public services are likely to contribute to better economic performance.

### Productivity is an essential component of state sector performance

A high-performing state sector is both effective and efficient; governments should pursue both ends. An unbalanced focus on effectiveness and outcomes risks wasting valuable resources. The pursuit of greater productivity is consistent with the achievement of better social outcomes for New Zealanders, and with the protection of fundamental values, such as justice.

## Measuring productivity in public services

### Better measurement is feasible and desirable

Although a few areas of the state sector creates challenges, for most public services, productivity measurement is possible and reasonably straightforward. The key steps in measuring public service productivity are to:

- identify the output (ie, the final good or service delivered by a government agency to an individual or organisation);
- calculate the input costs associated with the production of the output (eg, staff, capital, money);
- account (where necessary) for any changes in the quality of outputs (eg, are fewer people who receive hip operations having to return to hospital because of complications?); and
- develop an index to measure how the ratio of outputs to inputs has changed.

More sophisticated measurement techniques (eg, frontier analysis) can be a useful supplement to productivity indices, but significant gains in understanding can be made through simple assessments of productivity trends. Measures do not need to be complete or comprehensive to provide valuable insights into how well an agency or service is performing. Furthermore, simple measures can be developed into more sophisticated indicators over time.

## The poor state of productivity measurement and understanding

The New Zealand government has been organised around outputs since the late 1980s; the law requires these outputs to be fully costed; this legislative framework provides a sound foundation for productivity measurement. Despite this, the Commission found that for large parts of the state sector, there is little measurement or understanding of productivity. In some sectors covered by the inquiry, agencies were either unwilling to conduct productivity assessments, or were unable to bring together the data needed for such studies. Elsewhere, data was not available at the level of detail or quality needed to conduct proper measurement. Independent reviews of public sector organisations have consistently pointed to poor agency performance in terms of improving their efficiency.

The culprits for this lack of measurement and understanding are cultural and institutional. The public sector financial management system provides weak incentives for agencies to seek productivity gains, with most attention focused on gaining new revenue rather than optimising existing spending. Concepts such as “productivity” or “efficiency” also have negative connotations for some public service organisations or sectors. There appears to be little demand for productivity measures from Ministers, and insufficient demand from senior public sector leaders for the necessary information and analysis.

## Moving forward with measurement

Productivity measures have an important role to play in public sector performance frameworks and decisions. To expand and integrate productivity measurement into state sector practice, agencies should:

- design productivity measures to complement outcomes;
- design measures with the involvement of staff who deliver services;
- collect productivity data as part of “business-as-usual” activity;
- use productivity information primarily as the basis for conversations and learning about service improvement;
- ensure agency leaders actively support the use of productivity measures;
- develop measures that enable comparisons between similar organisations and business units or outputs, to identify and promote the sharing of good practice; and
- treat productivity measures as one input into performance decisions, rather than the sole factor with high stakes impacts.

Agencies may be deterred from measuring productivity by methodological questions, or limited access to data, but they should not be. Regular measurement is the first step to improve data quality and overcoming other challenges. It can also contribute to a better understanding of what drives public service delivery and quality, and identification of areas where greater value can be added.

## Encouraging productivity improvements

Better measurement is a necessary but not sufficient condition for improving state sector productivity. Productivity improvements emerge from innovation – that is, changes in how services are designed and delivered. Although there are many examples of innovation throughout New Zealand’s public sector, the “pockets of excellence” do not appear to spread far or fast. This lack of diffusion is a system failure.

There is a lack of evidence about the effectiveness of strategies available to agencies to boost innovation and productivity. The Productivity Commission is interested in learning more from stakeholders about the potential to lift state sector productivity through:

- changes to budget rules and processes to place more scrutiny on existing programmes;
- human resource practices that better engage staff;
- policies that shift resources from less productive to more productive service providers;
- measures that encourage the diffusion of successful innovations across sectors and public services; and
- smarter use of information and communications technology to re-engineer agency business processes.



# 1 State sector productivity – what it is and why it matters

## Key points

- “Productivity” is a measure of how well an organisation uses its resources to produce goods and services. In other words, it is a measure of efficiency. In the context of public services, greater productivity means using less of what the community values to obtain more of what it wants, over time. Properly measured, productivity captures changes in both the volume and quality of services provided.
- In line with the inquiry’s terms of reference, the primary focus of this report will be on technical efficiency – the optimal method of producing outputs. However, many productivity gains come from changes in the mix of services provided (“allocative efficiency”). The inquiry will also look at processes and systems that promote ongoing improvements in efficiency (“dynamic efficiency”).
- Higher state sector productivity provides more choices and better services, more sustainable public services, and flow-on benefits to the wider economy.
- Measuring and improving state sector productivity can be challenging. These challenges include:
  - difficulties defining the output of services and adjusting measures to reflect changes in the quality of those outputs;
  - regulations and practices within government, which limit opportunities for greater productivity and innovation within the state sector;
  - the fact that the state sector often has multiple stakeholders, who can have differing perspectives on what constitutes quality and value in public services, and may place their own needs ahead of a general concern for efficiency; and
  - the multiple tasks required of some public agencies, often with conflicting objectives. There may be unclear guidance about how trade-offs between objectives should be managed.
- Despite these challenges, measuring productivity is both feasible and desirable. New Zealand has made improvements to state sector productivity in the past, based on reforms to the state sector management framework in the 1980s, changes to business models and the introduction of new institutions. However, in recent times, the state sector management system has moved away from a focus on outputs. This may partly explain the reduced attention to state sector productivity.
- Achieving further gains in state sector productivity requires an openness to change and new ideas, and the flexibility and authority to revise and alter policies and processes. The guidance and recommendations in this report are designed to help New Zealand’s public services make these shifts.

## 1.1 The Commission’s task

The Productivity Commission has been asked to undertake an inquiry into how the “New Zealand State sector can effectively measure and improve productivity in core public services, with a particular focus on health, education, justice, and social support.” In doing so, the Commission is required to consider and provide advice on:

- How to measure efficiency / productivity in each of the identified core public service sectors: health, education, justice, social support. This should focus on meso (sector) and micro (function or

service) level measures. Guidance should consider key measurement and accuracy issues, and how imperfect measures are most appropriately and usefully employed.

- The appropriate role of identified efficiency/productivity measures in public sector performance frameworks, with the goal of improving assurance to Ministers and incentives on agencies for improvement. This should draw on theory and evidence of incentive and disincentive effects of measurement and other performance approaches on different workforces.
- Developing the capability, culture and systems that can support agencies to better measure, understand and improve productivity.

In carrying out this work, the Commission should not:

- carry out in depth analysis or provide detailed recommendations on specific policies relating to service access or provision in sectors; or
- duplicate work on issues, such as where to invest, or service effectiveness being developed as part of the social investment approach.

## 1.2 The New Zealand state sector

The state sector covers all organisations reporting to the Crown, including public service departments, Crown entities (eg, district health boards, schools) and Offices of Parliament, but excludes local government. The terms of reference direct the Commission to pay particular attention to specific sectors and service types, namely the core services in the health, education, justice, and social development sector: “teaching, hospitals and primary health care, policing, courts, corrections and work and income services.” While this report focuses on these sectors, the advice is intended to also be of use to the wider state sector.

## 1.3 What is productivity?

Delivering services – such as teaching, medical and justice services – is the main way in which the government works to improve the wellbeing of New Zealanders.

“Productivity” is a measure of how well an organisation uses its resources (inputs, such as labour and capital) to produce goods and services (outputs) and is typically expressed as a ratio of outputs to inputs. As such, productivity is a measure of efficiency.

There are three dimensions of efficiency.

- **Technical efficiency** is concerned with the optimal method of producing outputs.
- **Allocative efficiency** is concerned with the optimal distribution of resources to produce the right set of outputs.
- **Dynamic efficiency** is concerned with the structures, behaviours and incentives which create improvements in allocative and productive efficiency over time.

Simply stated, state sector productivity means using less of what the community values in order to obtain more of what it wants, over time.

Measuring and improving state sector productivity helps ensure that public resources are being used to their greatest effect.

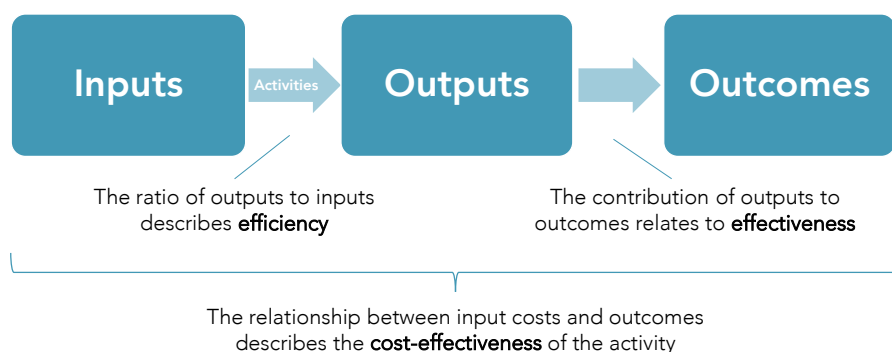
### What about effectiveness and cost-effectiveness in achieving outcomes?

To ensure that the community receives the most value from public services, officials need to be concerned with the effectiveness and cost-effectiveness of services, as well as how efficiently they are delivered (Figure 1.1).

In recent years, the call has been for commissioning agencies and providers to establish a clear intervention logic between the services delivered and the outcomes government is seeking (NZPC, 2015a).

For some public services, sophisticated methodologies for measuring the cost-effectiveness of interventions have been refined and embedded into prioritisation decisions. For example, the methodology for measuring the cost-effectiveness of pharmaceuticals is well established and embedded into the purchasing framework for drugs (Box 1.2).

**Figure 1.1 The relationship between inputs, outputs and outcomes**



The terms of reference for this inquiry ask the Commission to focus on a third and equally important aspect of ensuring New Zealanders receive value from public services – the technical efficiency with which public services are delivered.

### Technical efficiency... and allocative and dynamic efficiency

Technical efficiency is often thought of as “more for less”, or simply about reducing inputs. While this is one description of technical efficiency, it is not the only definition. Efficiency gains include changes that:

- reduce the quantity of inputs while maintaining the same level of service provision;
- reduce the quantity of outputs while reducing the quantity of inputs even further;
- lower prices for the inputs needed to provide public services;
- lead to additional or better outputs – such as enhanced quantity or quality<sup>3</sup> of service – for the same level of inputs; or
- result in additional or better outputs for an increase in inputs that is less than the value of the increased outputs.

The terms of reference for the inquiry steer the inquiry away from the “prioritisation of interventions” – however, many productivity gains come from allocative decisions that change the nature and mix of services provided. While the Commission does not consider or make recommendations about the allocation of resources to specific programmes or policies, it does consider the efficiency of where and how services are delivered. The inquiry also looks at mechanisms to provide ongoing improvements in efficiency over time.

## 1.4 Why is state sector productivity important?

### More choices, better services

In aggregate, productivity improvements release resources that can be used for more or different services, and result in better services for citizens. Greater productivity in public services also frees up resources for other uses, such as investment in new infrastructure, transfers back to citizens, or tax cuts. Measuring productivity can make these decisions more transparent and the trade-offs clearer.

### Sustainability

Higher state sector productivity allows the community to continue to afford the services it values. Although New Zealand currently has low levels of public debt and an operating balance surplus, there are likely to be

<sup>3</sup> Properly measured, productivity captures changes in both the quantity and quality of services provided.

increasing pressures on government spending and household budgets as the population ages and service expectations continue to rise. One important way of managing these rising fiscal pressures is to raise productivity. In the 2009 Long-Term Fiscal Statement, the Treasury looked at the impact of different public sector productivity growth rates on the delivery of public services over time and found that a:

lift in public sector productivity would have a positive impact on the notional basket of services that could be delivered to the average New Zealander for a given level of spending. A 0.5 percentage point increase in our baseline assumption for annual public sector productivity growth, if sustained, would result in around 20% more public services per person after 40 years. (New Zealand Treasury, 2009, p. 33)

An ongoing focus on productivity is important for ensuring the public receives the greatest value from its investment in government services.

## Flow-on benefits to the wider economy

The state sector makes up a large share of national income and employment; therefore, improvements in its productivity – both in terms of reduced costs and better outputs – help raise the performance of the wider economy, and its ability to provide higher living standards. While the government can only indirectly affect the productivity of private sector (eg, through regulation, policy and incentives), it has direct control and responsibility for the public sector.

The state sector also directly contributes to economic performance through the delivery of services that benefit the private sector – for example, publicly-funded education and training increases the supply of skilled workers and entrepreneurs. Police and court services help prevent economic loss through crime and support the fair and timely resolution of business and contractual disputes. As the New Zealand Nurses Organisation noted in its submission, health makes important contributions to “economic progress, as healthy populations live longer, are more productive, and save more” (World Health Organisation, cited in sub.14, p.3).

## 1.5 What are the barriers to measuring and improving state sector productivity?

While greater state sector productivity is clearly beneficial, achieving these gains can sometimes be challenging. Barriers to improvement include general difficulties measuring or attributing the outputs of services. There are measurement challenges that are particularly-relevant to public services, such as the political and institutional environment in which state sector organisations operate.

### Productivity measurement of services can be challenging in general...

Problems with measuring service productivity are not specific to the state sector, and a number apply equally to the private sector (Table 1.1).

**Table 1.1 Challenges measuring service productivity**

Issue	Implications
Output is “fuzzy”: the process of producing a service does not result in a tangible good, but in a “change of state” – eg, it can be harder to define the output of a bank, consultancy, or hospital than a manufacturer.	<ul style="list-style-type: none"> <li>• It can be hard to clearly identify the output of a service.</li> <li>• It might be difficult to separate the output of services from the factors of its production (ie, distinguishing the output from the process).</li> <li>• Possible challenges with the identification of innovation or quality improvements.</li> </ul>
Output in the service sector is often co-produced with customers – eg, customers often determine what kind of service they want (such as the type of haircut, or fitness programme), rather	<ul style="list-style-type: none"> <li>• Problems defining and identifying a standardised unit of output, as the customer’s involvement in production means each output is different and adapted to specific needs.</li> <li>• Difficulties identifying the value added by the provider, as opposed to the customer.</li> </ul>

Issue	Implications
than simply receiving standardised outputs.	

Source: Djellal & Gallouj, 2008

There are also technical challenges regarding whether an item counts as a service or product (a “good”). The Commission outlined this definitional issue in its report on *Boosting productivity in the services sector*, using the example of takeaway coffee:

Making the distinction between a good and a service can be complex – even for an apparently simple product such as takeaway coffee. Is the customer paying for the coffee (a good) or the application of the barista’s skills (a service)? The two perspectives ... reflect these differing views:

- The *production perspective* classifies takeaway coffee as a service, as it is an output of the accommodation and food services industry.
- The *transaction perspective* classifies the sale of a takeaway coffee as a goods transaction, as the cup and its contents change ownership at the point of sale. (NZPC, 2014a, p. 21)

## The state sector has distinct characteristics...

Some features of public services further complicate measurement. First, the outputs of some public services are intangible and difficult to attribute (eg, threatened species conservation).<sup>4</sup> Problems observing and verifying outputs also make it difficult to monitor and reward performance (or punish shirking).

Second, a number of public services are non-excludable and non-rival (eg, national defence, environmental protection, international diplomacy). The inability to exclude anyone from “consuming” the service (or identify specific “consumers”) can make it difficult to value the output.

Third, public services can be thought of as having multiple consumers – those directly receiving the service (eg, patients, school children) and the wider citizenry. These two groups may have different perspectives on a particular public service. Djellal & Gallouj (2008) distinguish between these direct and indirect consumers of public services and say

[a]lthough they are ultimately the same group of individuals, their perceptions of the output and performance can be said to be different. After all, direct consumers emphasize the nature and quality of the service provided and its mode of delivery, whereas indirect consumers are more concerned by the economic and financial aspects (as taxpayers) and the long-term socio-economic effects (as citizens). (p.67)

This raises the question of whose perspectives should be prioritised in the definition of an output and any quality adjustments.<sup>5</sup>

Finally, for many public services, there is either no price information (as services are provided without charge) or only limited price data (as they are not set in a competitive market, and may be partially subsidised). Prices play a key role in measuring productivity in the market sector, by:

- providing information about the relative value of different goods and services and changes in their quality;
- serving as weights when aggregating up goods and services (eg, into industry or national measures); and
- permitting analysis of output growth by removing the effect of changes in inflation.

However, the absence of price information is not unique to the state sector. For example, similar issues arise within large private sector firms that seek to measure the performance of their business units. Therefore,

<sup>4</sup> Chapter 3 discusses these issues in more detail.

<sup>5</sup> Quality adjustments are modifications made to output measures, to reflect changes in their quality. An example would be adjusting school teaching hours by exam achievement or measures of learning. Quality adjustment is discussed in Chapter 5.

having clear and accurate information about costs and outputs can be critical for productivity measurement in both state and private sectors.

### **...and it operates in a complex environment...**

A significant portion of productivity growth in the market sector is the result of influences that are external to individual organisations (eg, competition, and knowledge spillovers) (NZPC, 2014a). Some of these external influences do not apply as strongly to the state sector. As Dunleavy (2015) notes:

[e]arly private sector estimates said that 60 to 70 per cent of productivity gain in an industry was produced by displacement – that is, by customers leaving a firm that was less efficient or had an inferior product and moving to a firm that was more efficient and had a better product. That ratio has come down over time, and nowadays home-grown innovation is far more important in private sector productivity gain. But 50 per cent of all productivity gain still comes from people moving from the worst firms to better firms, while the bad firms go out of business. This is an effect that cannot be reproduced in the public sector (pp.39-40).

In practice, benefits from reallocation are achievable in public services, where there is contestability of service. However, contestability may not always be feasible, or may be difficult to introduce.

Public service actors also face a different set of incentives from those in the market sector, reflecting the distinct operating environment of the state sector. Important factors for productivity improvement in public services are the political environment, institutional settings, multiple stakeholders, multiple or misaligned goals, and agency motivations.

#### **Political environment**

The shortness of electoral cycles and of ministerial tenure may discourage investments in systems, infrastructure or capability with longer-term or uncertain gains, because elected representatives can worry about having to justify short-term costs in the absence of offsetting benefits, and officials may be wary of investing in specific assets that may have zero or negative value if policies change.

These risks can be managed to an extent by making credible commitments to make decisions or investments with timeframes outside of political cycles. Examples in New Zealand include the Public Finance Act 1989 (which requires Governments to set out medium-term fiscal goals) and the State Sector Act 1988 (which gives state sector chief executives responsibilities to maintain the long-term capability and performance of their departments) (NZPC, 2015a).

The adversarial nature of politics – in particular, the incentives for opposition parties and interest groups to highlight, or create the perception of flaws in Government policies and programmes<sup>6</sup> – can also encourage risk-averse behaviour in the state sector (Moszoro & Spiller, 2012). The Commission has observed such behaviour in its earlier inquiries into aspects of the state sector:

The culture of some New Zealand regulatory bodies appear to place significant weight on managing risks to the organisation, at the expense of the efficient management of social harm. Such cultures can resist innovation in regulatory practices. (NZPC, 2014b, p. 98)

The need for accountability and political risk management favours the use of prescriptive contracts, short contract periods and onerous reporting requirements. These factors work against the development and spread of innovation, and discourage productive and trusting relationships between government agencies and non-government providers, (NZPC, 2015b, p. 7)

...the Government and its agencies sometimes have strong incentives to suppress results that show a programme is performing poorly. New initiatives are often associated with an agency or political brand, and a perception of poor performance puts that brand at risk. One senior official, in meeting the Commission, described a government agency's internal evaluation unit as a "bomb factory" because evaluations were late and found faults. (NZPC, 2015b, p. 198)

<sup>6</sup> In a New Zealand context, these incentives have been described as "Ladley's Law", which states that oppositions "do not criticise government policy to improve it, they attack the policy in order to overthrow the government" (cited in Gill, 2011, p.186)

## Institutional settings

Rules and conventions within the government may act to constrain productivity gains.

- Budgetary systems which focus attention primarily on new, additional funding may weaken incentives to seek improvements within existing programmes.
- State sector employment conditions which set pay rates out of line with comparable market levels may limit the ability to attract and retain suitably-skilled staff, and overly-restrictive staffing rules or requirements (eg, through conditions attached to funding contracts) can constrain agencies from making internal arrangements that provide greater productivity.
- The OECD (2017) has also pointed to the impacts of “regulation inside government” on productivity, noting that public service delivery “eg, in education, health, and transport, can be improved by addressing constraints and burdens imposed by superfluous, outdated, or badly targeted regulation, and burdensome information obligations” (p. 29).

Submitters identified state sector rules and practices that limited opportunities for innovation or productivity. For example, the Methodist Mission identified:

Extensions to contracts for highly specific add-ons (e.g. two one-hour budget advice sessions) that are not necessarily wanted or needed by the clients but which a mandatory number must undertake for the contract to be fulfilled (sub.1, p.2)

The New Zealand Nurses Organisation highlighted the “continued failure to update obsolete medicines legislation” as another barrier to improved performance (sub.14, p.4).

## Multiple stakeholders

Service providers often have to respond to multiple groups - not just elected officials, but also their clients and the wider citizenry (Burgess & Ratto, 2003). This can constrain the ability of agencies to make changes to their input mixes or the types and amounts of outputs they provide. Some stakeholders may place their own needs ahead of a general concern for productivity. As Dixit (2002) comments:

[p]eople care not only about the outcomes of government agencies but also about some of the input and the methods of production, most notably about employment, and incomes in particular regions or communities. Thus there is no clear separation between ends and means. Labor unions are an important stakeholder or principal in many government agencies, and they care about various aspects of the inputs or means such as working conditions, and even about the incentive schemes themselves as they affect the incomes of their media members. And many citizens feel strongly about the methods used by law-enforcement agencies in combating crime.

More generally, some of the agents whose behaviour the policy or the bureaucracy is trying to influence are themselves principals in the prior political game that set the rules of the subsequent policy implementation or regulation game. They are selectively interested in specific dimension of the outcomes and inputs. (pp. 711-712)

There is often a community expectation that some public services will be provided in accordance with principles such as equality of treatment, or universal access regardless of location. Such expectations can be “the basis for very precise rules and a high degree of centralization, which may serve to make procedures more cumbersome when the drive for productivity gains demands rapid and flexible responses.” (Djellal & Gallouj, 2008, p.67). They may also require some sacrifice of efficiency (Dormer, 2011), although this may be offset by gains in quality or outcomes.

## Multiple or misaligned tasks

Agencies often carry out multiple tasks, with conflicting objectives and unclear guidance about how trade-offs between these objectives should be managed (Dixit, 2002; Burgess & Ratto, 2003). This lack of clarity about goals and trade-offs, combined with service outputs that can be hard to observe or measure, can make it difficult to monitor and encourage performance.

## Agency motivations

Finally, public service workers are sometimes motivated agents – that is, they “pursue goals because they perceive intrinsic benefits from doing so” (Besley & Ghatak, 2005, p.616). Intrinsic motivation can be positive for productivity, especially where the motivation of staff aligns with the objectives of their organisation:

Given the nature of the output in the public sector...managers and workers may care about the amount produced, and therefore be less inclined to shirk. For example, they may care about the sick, the old, or the unemployed. If the welfare of the clients is the sole goal of the organization itself, workers will internalize its objectives. (Burgess & Ratto, 2003, p.290)

However, as noted above, public service agencies often have multiple tasks and objectives, and differences between staff motivations and organisational goals, which may impede productivity improvements. Changes in an organisation’s goals and values can be particularly disruptive to morale or effort; this can lead organisations that place a high weight on their mission to be conservative and resist innovation (Besley & Ghatak, 2005). Other scholars (Tullock, 1976; Niskanen, 1973) have argued that public sector workers have strong incentives to raise their status by seeking increases to their budgets, thereby reducing productivity.

## ...but the barriers to measurement and improvement can be overcome

While there are clearly challenges in measuring and improving state sector productivity, they are not insuperable. Academic and official literature has clarified techniques for measuring productivity at the agency and service level (eg, Atkinson, 2005; Dunleavy & Carrera, 2013) and changes to government institutions across a range of countries have improved the information bases and incentives for efficiency gains. Indeed, as the following section describes, New Zealand in recent times was one of the leaders in reforming state sector processes to increase productivity.

## 1.6 Recent experience

### New Zealand has done this before...

The goal of achieving higher productivity from the state sector is not new. The two main public sector management statutes (the State Sector Act and Public Finance Act) were introduced in response to dissatisfaction by the then-Government with the “relative lack of responsiveness of large bureaucracies... [and] the very considerable difficulties in ensuring those bureaucracies used resources in a way that was efficient, responsive and flexible” (Palmer, 1988, p.2). The new management frameworks sought to achieve a number of objectives, including to “encourage a responsive and efficient public sector” (New Zealand Treasury, 1996, p.7).

The accompanying reforms to the state sector in the late 1980s and early 1990s removed detailed internal controls on departments and agencies, delegated decision rights to departmental chief executives (with reporting and accountability obligations), and significantly improved reporting to ministers, Parliament and the public on Government’s spending and fiscal position.

The reforms also organised Government’s budgets onto an output basis; instead of paying for inputs (such as staff and rents), ministers were explicitly buying services. Key features of these reforms were that:

- by law, output expenses must be fully costed, including “the full cost of producing and supplying outputs measured in accrual accounting terms; and...the full allocation of overhead costs” (Public Finance Act 1989, s 2);
- agreed and costed outputs form the basis of the Government’s budget Estimates; and
- in the early years of the reforms, each output class (groupings of outputs of a common type, such as departmental output expenses) was a distinct and legally binding appropriation, leaving departmental chief executives with “very little flexibility over switches between output classes” (Scott, 2001, p.178). This constraint was designed to force conscious choices about which, and how many, outputs to buy.

This combination of greater clarity about spending intentions, and more managerial freedom for agencies led to a number of productivity gains in the 1990s and 2000s. Many of these were achieved through the redesign of business processes (Box 1.1).



**Box 1.1 Productivity gains from public sector reforms**

Restructuring of the Income Support service in the Department of Social Welfare saw dramatic improvements in the time taken to process benefit applications:

From June 1992 to June 1993 the average time taken to grant a benefit had been cut from 13 days to less than 6 days. (The standard required in the department's 1992/93 purchase agreement with the Minister was 20-35 days, depending on the benefit type.) By the end of May 1994 IS [Income Support] had achieved its goal of a one day benefit turn around time. (Petrie, 1998, p.27)

These improvements were achieved with reduced staff numbers and flat operational funding (Smith & Norman, 1998, p.9). A 1996 Treasury study of the effects of public sector financial management reform said of Income Support

...over 1991/92 to 1993/94 a 20% increase in primary applications and a 60% plus increase in total (primary plus supplementary) applications have been accommodated with barely an increase in operating expenses. (Brumby et al, 1996, p.20)

Petrie (1998) attributes these productivity improvements in part to the public sector management reforms, commenting that the initial gains in Income Support reflected the general emphasis in the state sector on "the efficiency of output delivery" (p.62).

Brumby et al (1996) also found the average unit cost for property titling services in Valuation New Zealand fell in "the order of 10-20 percent in nominal terms between 1989/90 and 1994/95." (p.19) Similar improvements were found in immigration application processing.

Whitehead (2006) cites Inland Revenue and the Companies Office as examples of productivity improvements. For Inland Revenue, between 2002 and 2005, the "cost of collecting \$100 of tax... reduced from \$0.89 to \$0.77 – a 13% decrease" and the "number of customers per staff member...increased" (p. 6).

The Companies Office, using "greater managerial freedom to apply new technology judiciously combined with clear performance expectations...was able to reduce the cost of registering a company from \$200 in 1995 to \$70 in 1999. Its average turnover time for registering a company fell from two weeks to just thirty minutes" (2006, p. 6).

Other gains were achieved through the design of new public institutions, such as PHARMAC (Box 1.2).

**Box 1.2 PHARMAC – controlling costs, increasing access and effectiveness**

The Pharmaceutical Management Agency (PHARMAC) was established in 1993 with the aim of securing the best health outcomes from pharmaceutical treatment within the available funding. Its establishment was a response to rapidly-rising community drug treatment costs, which grew at around 15% a year over the 1980s.

PHARMAC's key roles are deciding whether a medicine will be subsidised, the price the government is willing to pay, and the conditions of access to the drug. Medicines are assessed against a number of criteria, including population health needs, clinical benefits and risks, cost effectiveness, and contribution to Government priorities. PHARMAC also uses a range of strategies to deliver value for New Zealanders, including tendering, negotiations with supplier, and reference pricing.

PHARMAC's role has expanded over time. In its earlier days, the Agency focused on managing community medicines. In 2001, it began making funding decisions on cancer medicines used in public hospitals and other functions were added in 2012 (management of the national immunisation schedule) and 2013 (management of hospital medicines and negotiation of national contracts for hospital medical devices).

The benefits of PHARMAC's activities are well-documented. PHARMAC estimates that between 2005 and 2016, it avoided nearly \$6 billion in extra costs for district health boards (PHARMAC, 2017). At the same time, the number of subsidised medicines increased, with 188 items added between June 1993

and 2007 (Cummins et al, 2010). Grattan Institute research found that New Zealand pays substantially-lower prices for a range of medicines than Australia:

If Australia adopted New Zealand's prices for 62 identical doses available in both countries, it would save \$1.1 billion a year...for the 73 doses we compared, Australian wholesale prices are eight times higher than New Zealand's. For identical drugs – a more conservative comparison – our prices are six times higher (Duckett, 2013, pp.7-8).

Submissions show continued efforts within public agencies and wider social services to raise productivity. For example, the Inland Revenue Department highlighted the benefits of their investments in Information Communication Technology (ICT) in terms of lower costs and better services for citizens.

- Online services for customers [enable them to] manage their tax and social entitlements. More than 2.5 million customers have an active myIR account and our GST filing service was improved in February 2017 as part of Stage 1 of our transformation. A total of 79% GST returns are filed through this service within myIR (up 6% on 2016) which saves our customers time and money.
- ...We set up a Reducing Paper Output project in 2015 to give customers access to statements, notices and letters electronically. eAlerts are used to let customers know that a document is available in their myIR account. The documents are stored as PDFs and can be viewed, saved or printed. Reducing paper output has been a success with nearly 32,000 fewer reams of paper being used each year...
- Voice Biometrics – allowing customers to register for VoiceID has removed the need to manually identify customers each time they call. The nearly two million customers using this service are now identified by our automated (IVR) phone system, and their calls are 20-30 seconds shorter than calls requiring manual identification. (Inland Revenue Department, sub.7, p.6)

The New Zealand Nurses Organisation referred to co-operative health care demand management programmes being implemented between unions and district health boards:

CCDM [Care Capacity Demand Management] is a management system focused on improving the quality of care for patients, the work environment for staff and organisational efficiency. The programme is an organisational approach to ensuring the demand for patient care is matched accurately and effectively with the resources required, in 'real time'. It was developed as part of a safe staffing/health workforces initiatives, and is being implemented and resourced via an agreement between the 20 DHBs (shared services), the New Zealand Nurses Organisation (NZNO), the Public Service Association (PSA), and the Service and Food Workers Union (SFWU). (sub.14, pp. 6-7)

The Methodist Mission discussed the "significant improvements in both efficiency and effectiveness" they have achieved in service delivery as a result of their use of "real-time monitoring across KPIs focussed on the appropriate use of the tools that the evidence shows will deliver the outcomes we seek." (sub.1, p.1)

### **...but the focus has come off outputs**

However, the existence of this inquiry and evidence from other quarters suggests that the earlier focus on the productivity of public services has waned somewhat. In part, this may reflect the attention of the public management system shifting away from outputs.

#### **The rise of outcomes...**

New Zealand's public finance management system was designed with outcomes in mind (Hitchener & Gill, 2011a; Scott, 2001). However, shortly after its introduction, concerns were expressed by ministers, public servants, and external commentators that the system's focus on outputs could be crowding out outcomes. SSC (1999a) note that by 1992:

there was a growing sense of disquiet that New Zealand public managers had become so focused on outputs, that is, what they were doing and how they were doing it, that they were starting to lose sight of the reasons why they were doing things. (p.11)

Minister of State Services Paul East commented to a 1997 public sector senior management conference that the preoccupation with outputs had tended to make departments “more defensive, less objective than they should ideally be”, impairing coordinated efforts across departments (1997, p. 7).

A number of steps were taken, starting in the mid-1990s, to put greater emphasis on outcomes in public sector strategic and accountability documents. In their earlier manifestations, these efforts explicitly linked outputs to outcomes. For example, under the Strategic and Key Result Areas (SRAs and KRAs) initiative in the mid-1990s, ministers stated their desired outcomes (in the SRAs) and departmental chief executives identified the outputs that they would deliver that contributed to each outcome (in the KRA).

Over time, however, the link between outputs and outcomes has become less clear. Current accountability mechanisms for departmental chief executives under the Better Public Services seek to make agency heads directly accountable for outcomes. Scott & Boyd (2017) report that under new chief executive performance assessments,

Chief executives were either rewarded or sanctioned based on the collective performance of the interagency group in addressing the nominated problem. Performance bonuses for chief executives were awarded on the basis of collective (group) achievement. (p.38)

Changes to public finance legislation aimed at increasing agency flexibility have also weakened incentives on departments to focus on outputs. Lonti and Gregory (2007) comment that while these amendments were designed to give managers

more flexibility in defining and measuring outputs, and be more creative in shifting resources across output classes, all with an eye to the strategic achievement of policy outcomes, they will now be held to account less for their ability to produce specified outputs. (p.471)

Hitchener and Gill (2011 a) similarly conclude that in practice

New Zealand has moved from the goal of budgets based on costed outputs, to baselines for each output class (output expense appropriation) and then to pseudo baselines operating at the departmental level. Performance expectations presented broadly as a result, guide chief executives in selecting the mix of interventions within these ‘departmental’ baselines. (p.91)

Hitchener and Gill also observe a deterioration in output performance information, as a result of the growing focus on outcomes (pp.89-90).

### **...and the return of input controls**

Developments in state sector performance management since 2008 have seen a renewed interest in revealing and controlling input costs. Examples include the Benchmarking Administrative and Support Services (BASS) surveys, the establishment of state sector wide “functional leads” for key state sector inputs, and the introduction of a “cap” on core government administrative staff in 2009.

#### **The BASS surveys**

Beginning in 2011, the Treasury conducted annual surveys of administrative and support services across government, covering a range of departments and Crown entities. The surveyed services include human resources, finance, information and communications technology, procurement, property management, and corporate and executive services. The expressed aim of the survey was to respond to “Government demands for better, smarter public services for less” and to provide “an evidence base for assessing current performance, setting targets, identifying and quantifying opportunities for improvement, and tracking changes in performance over time” (New Zealand Treasury, 2011, p. 8).

#### **Functional leadership**

In 2012, Cabinet established functional leadership aimed at “securing economies or efficiencies across departments, improving services or service delivery, developing expertise and capability across departments, and ensuring business continuity” (Office of the Deputy Prime Minister and Minister of State Services, 2012, p.5). The original concept for functional leadership came out of the Better Public Services Advisory Group report, which argued that reduced duplication and inconsistency in activities common to many agencies would create fiscal savings and service improvements. (Better Public Services Advisory

Group, 2011). Departmental chief executives and a number of Crown entity Boards were required to seek agreement from functional leads before taking specified decisions or actions, and to adopt common standards or contracts set down by the leads. Although the policy was not explicitly targeted at inputs, the functional leaders appointed to date have been for agency inputs rather than outputs – namely, property, procurement and ICT.

### Core government administration cap

In 2009, the Government set a cap on the number of “core government” administrative staff, reflecting a pre-election commitment. The aim of the cap was to give “priority to delivering frontline services that directly benefit New Zealanders” and help manage public expenditure (Office of the Minister of State Services, 2009, p.1). The cap was originally based on “the number of FTE staff in Public Service departments and selected Crown entities<sup>7</sup> plus unfilled vacancies as at 31 December 2008” (Hon Tony Ryall, 2009).<sup>8</sup> “Frontline services” (including Work & Income and Child Youth & Family staff in the Ministry of Social Development; sections of the Department of Corrections; and the Police and New Zealand Defence Force) were excluded.

### Implications

This reduced attention to outputs has several potential implications for state sector productivity:

- reductions in the quality of output information may make measurement and analysis more challenging;
- the effective move towards “departmental baselines” and looser constraints around outputs reduces incentives to cost outputs accurately;
- input controls, such as the staffing cap, may limit the ability of chief executives to allocate inputs for maximum effect; and
- effectiveness is prioritised over efficiency, rather than pursuing both goals.

## 1.7 What is needed to achieve future gains?

The aim of productivity measurement is not to simply count inputs and outputs. The point is to understand what is happening, and create an environment in which agencies are able and encouraged to seek improvements in productivity. As noted above, productivity gains can come from improvements to business processes and the design of state sector institutions. Another source of productivity gains is technological advances (Box 1.3).

### Box 1.3 Productivity improvements in cataract treatment – better care, lower costs

Cataract surgery is one of the most common elective operations conducted in New Zealand, with around 16 000 publicly-funded procedures carried out in 2016 (Broatch, 2017). They also provide good value for money, with similar QALY (quality-adjusted life year) results to hip replacements, but at a fraction of the cost.

In the first half of the twentieth century, cataract surgery was a time-consuming and invasive process for patients, often requiring a week-long stay in hospital. Over time, surgical techniques improved, resulting in faster and better healing, fewer complications and more certain results. Cataract surgery is now delivered through outpatient clinics and generally takes less than an hour, allowing patients to recover at home and return to work faster.

Improvements in the treatment of cataracts have not only improved the experience for patients, but have also reduced costs. Shapiro, Shapiro and Wilcox (2001) constructed a “prototypical CPI” for cataract surgery in the United States over the period 1969-1994, and found that “*even without taking*

<sup>7</sup> Housing New Zealand Corporation, New Zealand Qualifications Authority, New Zealand Transport Agency, New Zealand Trade & Enterprise, and the Tertiary Education Commission

<sup>8</sup> The cap was reset in 2012, based on FTE numbers and unfilled vacancies as at 30 June 2011.

*quality improvement into account, the prototypical index shows that the cost of cataracts fell hugely relative to the published general price level.” (p. 431)*

All these sources of productivity require an openness to change and new ideas, and the flexibility and authority to revise and alter policies and processes. Perhaps more than anything else, this involves creating an environment that enables innovation in the delivery of public services. Dunleavy (2015) argues that in the absence of the pressure for improvement caused by competition, “home-growth innovation produced by managers – particularly medium managers, lower managers and staff themselves – is the lifeblood of change in the improvement of [public sector] productivity” (p.40).

Allowing and encouraging this innovation to take place requires shifts on multiple levels; this includes:

- creating or enhancing agency cultures that are supportive of experimentation and evaluation;
- building the datasets and analytical capabilities to understand changes in performance and productivity;
- removing institutional or systemic barriers to innovation and productivity; and
- establishing authorising environments that give ministers, Parliament, and the public confidence, thereby permitting experimentation and innovation within public services.

The guidance and recommendations in this report are designed to help the New Zealand state sector make these shifts.

## 1.8 The structure of this report

This report is designed with the needs of those who would develop or use productivity measures in mind, and is in two parts.

Part 1 is aimed at providing guidance to officials who may be asked to develop productivity measures for their agencies, or for particular service lines.

- Chapter 2 introduces the key steps in measuring state sector productivity, discusses the importance of being clear about what is being measured and who the primary client is, and describes key measurement techniques.
- Chapter 3 provides advice on how to define outputs in different circumstances, how to weight them in productivity measures, and how to deal with output measurement challenges.
- Chapter 4 focuses on issues related to measuring inputs, including how to attribute them to outputs, how to deal with co-financing, and how to assign capital costs.
- Chapter 5 discusses quality and inflation adjustment of productivity measures, including when these adjustments are important, and how to do them.

Part 2 is designed for senior officials, ministers and parliamentarians looking to make changes to measure and improve productivity.

- Chapter 6 discusses the place for productivity measures in the performance management system, and lays out some principles for integrating such measures into state sector practice.
- Chapter 7 discusses the systems and capabilities needed to measure productivity across the state sector.
- Chapter 8 explores potential sources or drivers of productivity in state sector services.

Put simply, Part 1 describes *how* productivity can be measured, and Part 2 outlines *why* it should be measured.

Alongside this work, the Commission is also developing a series of case studies, which take a detailed look at issues relating to the measurement of productivity in the core public services identified in the terms of reference. They are designed to provide practical guidance to officials grappling with definitional and measurement challenges. The case studies cover the following sectors and topics (Table 1.2).

**Table 1.2 Productivity case studies**

Project title and brief description
<p><b>Productivity measurement case study: Early childhood education</b> – uses publicly-available information to construct raw measures of productivity in the ECE sector, and discusses options for quality-adjusting ECE outputs to provide a more accurate and complete picture of changes in early childhood productivity.</p>
<p><b>Quality-adjusting sector level data on New Zealand schools</b> – estimates a range of quality adjusted productivity measures for New Zealand schools and discusses the benefits and risks of different approaches.</p>
<p><b>Productivity measurement case study: Ministry of Social Development</b> – identifies potential ways that MSD’s individualised Cost Allocation Model (iCAM) can be used to develop productivity measures for MSD’s Income support benefit payment services. The iCAM model uses existing administrative data taken from the client management IT systems, to calculate how much staff time is used on a range of different activities, and then a range of other costs such as corporate overheads are added to the data set. This case study highlights how routinely-collected administrative data can be used to measure changes in productivity over time.</p>
<p><b>Understanding health sector productivity</b> – illustrates both the changing economic complexity of hospital output and the status of key sources of data in the health sector. These two topics fit neatly together because, for example, while discharges are aging it is unclear whether this is evidence of the “healthy aging” hypothesis or evidence about the interaction between hospital care and other parts of the health system (such as primary and residential care). To fully answer these questions more system-wide data are required that could potentially be used as the basis for productivity measures.</p>
<p><b>Quality-adjusted productivity in tertiary education</b> – constructs quality adjusted productivity indices for the tertiary education sector. It proposes a number of methods for making quality adjustments to “basic” measures of labour and multi-factor productivity for teaching across the New Zealand tertiary sector as well as measures of research productivity for universities.</p>
<p><b>Productivity measurement case study: Police</b> – develops a productivity measure for police responses to mental health incidents – an area of particular interest for the New Zealand Police.</p>
<p><b>Productivity measurement case study: Courts</b> – develops productivity measures for New Zealand’s court system using publicly-available data. The available data is significantly limited, and accordingly, a particular focus of this case study is on how these data limitations affect the utility of productivity measures.</p>

## 2 Measuring state sector productivity

### Key points

- While measuring and understanding the productivity of the public sector matters for all countries, progress toward developing robust productivity measures has been slow.
- Many countries have valued public sector outputs at their cost of production. The convention that “inputs=outputs” meant public sector productivity was assumed to be constant through time. More recently, the inputs=outputs convention has been rejected, with several influential studies recommending the direct measurement of outputs.
- The broad steps involved in developing a productivity measure for a state sector agency are:
  - identify the core outputs that the relevant agency produces, identify the unit costs associated with each output, and then develop a cost-weighted total output metric;
  - calculate the input costs associated with the production of outputs – this will include costs such as staff salaries, the cost of intermediate inputs such as services provided by contractors, and the capital costs for buildings and major pieces of equipment; and
  - determine whether it is necessary to account for changes in the quality of outputs, and if so, how to achieve this.

Advice regarding each of these steps is set out in other chapters of this report.

- Before beginning any technical aspects of developing a productivity measure, a first step should be to clearly establish the underlying purpose for measuring productivity. Performance measures can serve several purposes:
  - to learn (what’s working and what’s not? why? how can improvements be made?);
  - to steer and control (are policies and programmes on track?); and
  - to give account (can performance be justified?).
- Productivity measurement can occur at a number of levels (eg, sector, agency, or service level). Measures are likely to be most useful where they align with the scope of control of decision makers.
- Different types of output and input measures will answer different questions. For example, *gross* output measures are useful for understanding the total output of a sector or agency, while *value added* measures are useful for assessing the marginal additional value added (or removed) by the delivery of a particular service.
- A number of techniques could be used to analyse productivity, ranging from relatively straightforward ratio analyses, to more complex approaches based on frontier techniques.

This chapter introduces the high-level steps involved in developing productivity measures for the state sector. It also discusses the different kinds and levels of productivity measurement, and the importance of designing measures to meet the needs of different clients and answer the right questions.

Before delving into methodological issues, the chapter begins by setting out the current state of play regarding state sector productivity measurement internationally (section 2.1) and in New Zealand (section 2.2). Section 2.3 summarises the key steps involved in developing a productivity measure. The chapter then discusses different levels of productivity measurement, and the importance of designing measures to meet

the needs of different clients, and answer the intended questions (section 2.4). The chapter concludes (section 2.5) by outlining some of the main techniques that can be used to analyse productivity.

## 2.1 International practice in measuring state sector productivity

Improving the productivity performance of the public sector is important for all countries. The reasons public sector productivity matters for New Zealand such as demographic shifts and the flow-on benefits from a well-performing public sector to the rest of the economy are common to most developed countries. But despite this, efforts to accurately measure and record public sector productivity are limited:

To date, the lack of measures to appropriately capture public sector productivity building on, and going beyond, the System of National Accounts, has meant that major policy decisions are being taken without adequate understanding of their implications for the economy as a whole. (Lau, Lonti & Schultz, 2017, p. 182)

As discussed in Chapter 1, measuring the productivity of services can be challenging, particularly when those services are provided in the public sector. Murray (2010) notes that these complexities, along with differences in the way public services are viewed, have resulted in disappointingly slow progress in understanding productivity of public services.

In the market sector there is a long tradition of studying production functions, demand for inputs, average and marginal cost functions, elasticities of supply, productivity and technical progress. The non-market sector has gone largely unnoticed. In part this can be explained by general difficulties in measuring the output of services, whether public or private. But in part it must be explained by a completely different perspective on public services apart from private services. Resource use for the production of public services has not been regarded as inputs into a production process, but as an end in itself, in the form of public consumption. Consequently, the production activity in the government sector has not been recognized. (p. 414)

### Evolution of public sector productivity measurement

In the absence of comprehensive public sector productivity measurement, many countries have fallen back on valuing public sector outputs at their cost of production. The convention that “inputs=outputs” means that public sector productivity is assumed to be constant through time (ie, no productivity growth). Additionally, it implies that all spending was equally worthwhile; each dollar spent anywhere in government was assumed to create a dollar of value, regardless of what was delivered.

But clearly, productivity improvements in the public sector are possible - for example, by the public sector delivering more outputs for a given level of spending.

Dunleavy (2016) notes that advances in data collection and analytics, along with improved understanding of productivity measurement, provide an opportunity to improve the understanding of government productivity:

It should now be feasible to avoid many past problems of measurement, and to exploit the current potential of improved government information systems, so as to move decisively away from the conventional assumption of static government productivity. (p. 4)

One way to avoid having to assume that inputs=outputs is to attempt to directly count the outputs in a given area of the state services – for example, to count the number of court trials, and divide it by the total cost of administering trials.

This approach does not provide a measure on the value of these areas of government activity, but it does provide a measure of changes in the volume of outputs over time.

In the early 2000s the statistical office of the European Union, Eurostat, published their *Handbook on Price and Volume Measures in National Accounts*, which required member states to develop direct output measurements (Eurostat, 2001). The handbook rejected the inputs=outputs convention and provided guidance on three different classifications of output measurement techniques: “preferred methods”; methods to be used when it is not possible to apply the preferred methods; and methods that should not be used.



To classify as a preferred method, output measures need to be comprehensive (covering all services provided); weighted by the cost of each type of output in the base year; as detailed as possible; and be adjusted for changes in quality.

In 2005, the United Kingdom Office of National Statistics commissioned the *Atkinson Report* on the measurement of government output and productivity. This report also found that the direct measurement of outputs was best practice:

There is an intrinsic case based on public accountability for seeking to measure what is achieved by spending on public services. We cannot simply assume that outputs equal inputs in such a major part of the economy. To fail to measure the output would be to miss the essential complementarity between public services and private economic growth. (Atkinson, 2005, p. 182)

The report set out principles covering the measurement of outputs, inputs, and productivity including that:

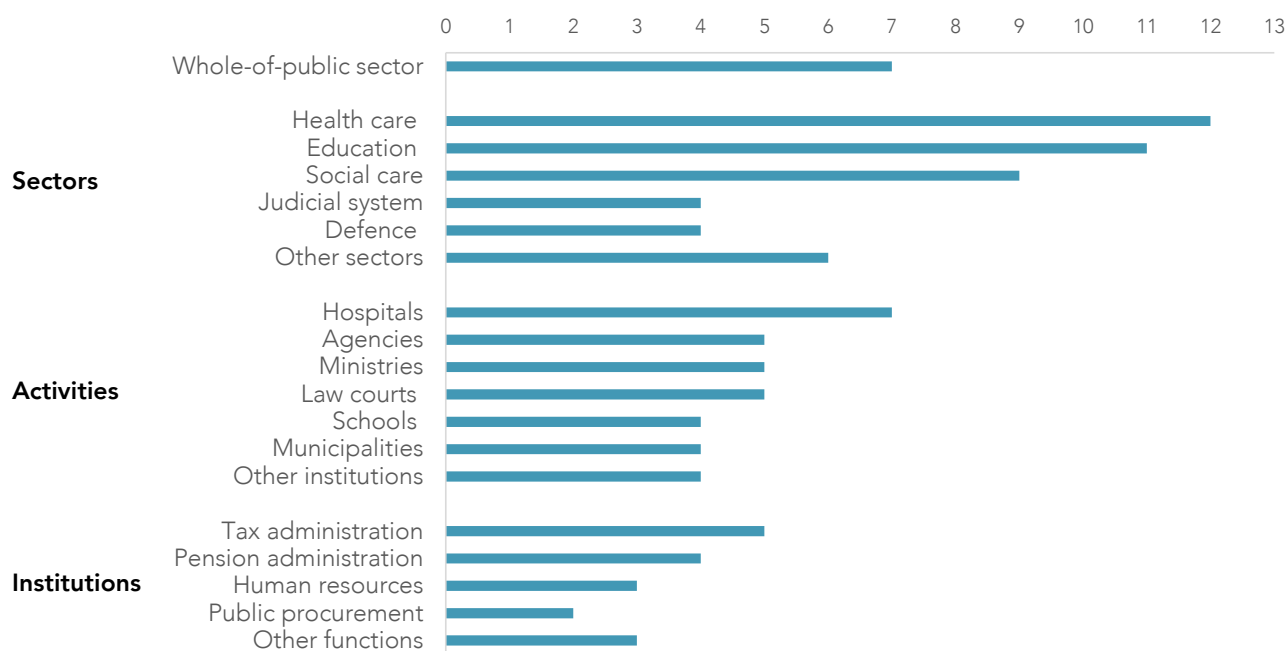
- outputs should be adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome;
- output indicators should cover the full range of services for that functional area;
- the measurement of inputs should be as comprehensive as possible, and in particular should include capital services; and
- independent corroborative evidence should be sought on government productivity, as part of a “triangulation” process, recognising the limitations in reducing productivity to a single number.

The OECD has also offered guidance regarding the measurement of public sector productivity – in particular the measurement of output for health and education services (OECD, 2010). More recently, the OECD has reviewed current practices and challenges to measuring public sector productivity and set out five areas to enhance measurement efforts:

- improvements to input measurement and cost accounting;
- standardisation of measurement practices to enhance comparability;
- output measurement beyond the education and health sectors;
- building a better understanding of countries’ existing micro level productivity measurements; and
- intra-governmental coordination on productivity measurement (Lau, Lonti & Shultz, 2017).

## **Public sector productivity measurement in other jurisdictions**

International practice in measuring government productivity is variable, although overall it is not well developed. The OECD recently conducted a survey (2016 and 2017) on countries’ existing efforts to measure public sector productivity. Thirty member countries and two partner countries responded to the survey. It revealed that public sector productivity measurement is not widespread, with health care, education, and social care sectors being the most common areas measured (Figure 2.1).

**Figure 2.1 Countries measuring public sector productivity (from a sample of 32)**

Source: Lau, Lonti & Schultz, 2017

**Notes:**

The results don't indicate whether there are a core of countries that measure productivity for most sectors, or whether there are many countries, each measuring something different.

Box 2.1 provides examples of how countries approach the measurement of public sector productivity.

### Box 2.1 Measurement of public sector productivity – selected examples

#### United Kingdom

In the 1950s and 1960s, the UK statistics office attempted to directly count outputs such as the number of court trials or the number of hospital patients. Its view was “even a crude measure of output is assumed to be preferable to an index based on total cost” (Central Statistical Office, 1956, p. 42, cited in Atkinson, 2005, p. 17). But the indicators developed at that time were criticised and later abandoned. A later review wrote

From this earlier experience, we draw two main conclusions. The first is that the design of direct output measures needs considerable care. It is not necessarily the case that ‘even a crude measure of [government] output is ... preferable to an index based on total cost’. The fact that it is not easy to obtain direct indicators means that better measures are likely to require significant investment of resources. Direct measures of output need to be continuously monitored to ensure that they are capturing changes in quality. The second conclusion is that ONS [the Office of National Statistics] has to steer a careful course with regard to changes in government policy, guaranteeing the independence of the approach to measuring government output while ensuring that its implementation reflects the realities of public spending and circumstances. (Atkinson, 2005, p. 17)

In the 1990s, the UK recommenced the direct measurement of government outputs and is now considered “one of the pioneers in measuring productivity and efficiency of the public sector” (Lau, Lonti & Shultz, 2017).

- In 1998, the UK Office of National Statistics (ONS) attempted to directly count government output, first in the health, education and social security sectors, and later in other areas of government.
- The first public service productivity statistics were released in 2004.
- Measures of total public service productivity were published for the first time in 2009.

In addition to regular publications of public sector productivity statistics, the UK government has carried out a comprehensive review of public sector efficiency in 2014–15. The review gathered evidence on efficiency trends and drivers, including ONS productivity estimates for the whole public sector, and identified areas for achieving further efficiency improvements.

### **Finland**

Finland began a project to measure public sector productivity in 1995. The project was headed by Finland's national statistical institution, and expanded in 1997 to include productivity measurement for local government services. The agencies for which input/output data was collected accounted for roughly 70–80% of central government employee remuneration.

The way that Finland's public sector productivity data has been used has changed over time. In the early 2000s the data was used to improve processes in preparation for the growing number of public servants that were retiring. From 2005, there was a deliberate effort to raise productivity to enable fewer personnel, with any cost savings divided between the government and the ministry that achieved them (Pekonen, 2011).

In 2014, Statistics Finland announced that, owing to a reduction in their resources, it would discontinue its central and local government productivity statistics (Statistics Finland, 2014).

### **Denmark**

In 2014, Denmark moved from the use of an input-based method to an output-based method for measuring the volume of government production. The Danish National Accounts measure the volume of government services in the areas of health care, social protection, education, and recreation and culture and applies a weighting based on the service's unit cost. For collective services (eg, defence, public order and safety, environmental protection) the input-based method is still used (Lau, Lonti & Shultz, 2017).

### **Australia**

The Report on Government Services (RoGS) is an important source of performance data in Australia. This annual publication compares the efficiency and effectiveness of Commonwealth and State/Territory government services such as education, health, justice, emergency management, community services and housing.

The court system is one example of a government service that is analysed in the RoGS. A selection of efficiency and effectiveness measures are reported for most State and Territory courts and Federal Courts (APC, 2017a).

There is evidence that the information in RoGS has played a significant role in informing policy development across a broad range of services. In the justice sector for example, RoGS illustrated the significant efficiency gains associated with Victoria's use of electronic courts for minor traffic infringements, which has been adopted by other jurisdictions (APC & NZPC, 2012).

## **Productivity measures for specific services**

In addition to public sector productivity measures collected and reported by countries' statistical agencies, there are numerous studies examining productivity at a more micro level. Table 2.1 provides a summary of some selected examples.

**Table 2.1 Overseas studies of productivity in specific public services – selected examples**

Title and authors	Summary
<i>Measuring the performance of police forces in Taiwan using data envelopment analysis</i> (Wu, Chen & Yeh, 2010).	<p>This study uses data envelopment analysis to construct an efficiency measure for Taiwan's 22 police precincts. The outputs measured in the analysis were the number of general and special services provided, and the number of burglaries, violent crimes, and other crimes solved. The inputs were labour cost, general operating costs, and equipment purchasing costs.</p> <p>In addition to examining overall efficiency, the report also considers how differing levels of urbanization and other external environmental factors (including the number of public housing units, the unemployment rate, and average household yearly revenue) affect efficiency. None of these variables were shown to have a statistically significant effect on efficiency.</p>
<i>Efficiency in Northern Ireland Hospitals: A non-parametric analysis</i> (McKillop et al, 1999).	<p>This study examines the efficiency of 23 hospitals in Northern Ireland. Non-parametric analysis is used to measure the efficiency of larger and smaller hospitals relative to best practice.</p> <p>The results show that the smaller hospitals not only commenced with a lower average level of overall efficiency but also experienced a much greater level of decline over the timeframe of the study (1986–1992).</p>
<i>Return on Educational Investment: 2014. A District-by-District Evaluation of U.S. Educational Productivity</i> (Boser, 2014).	<p>This study measures the productivity of nearly all the school districts in the United States. Reflecting the complexity involved in measuring school productivity, the study presents three different measures of productivity rather than relying on a single measure.</p> <ul style="list-style-type: none"> <li>• A basic return on investment index which measures how much academic achievement is obtained for each dollar spent.</li> <li>• An adjusted return on investment index, which accounts for certain factors outside a district's control such as the added costs of educating low-income, non-English speaking, and special education students.</li> <li>• A predicted efficiency index, which measures whether a district's achievement is higher or lower than would be predicted after accounting for its per-pupil spending and concentrations of students with additional needs.</li> </ul> <p>The report finds that more than 275 school districts (from a total of 7 000) were highly inefficient across all three measures. Schools in these districts served about 3% (1.23 million) of the students covered in the study.</p>
<i>Rapid productivity growth – customs regulation</i> (Dunleavy and Carrera, 2013).	<p>This study examines the productivity of customs regulation in the United Kingdom over a ten-year period from 1997/98. The outputs measured are the total number of import and export declarations processed per year. These were weighted by the relative unit costs in each year to create a total outputs data series. The inputs used were staff costs, cost of direct material and other costs, procurement, outsourcing of services provision, and capital investment.</p> <p>The results show a dramatic improvement in productivity over the course of a decade. The authors suggest this reflects a shift away from staff-intensive processes toward more risk-based deployment, and an early investment in proven technologies.</p>

## 2.2 Measures of state sector productivity in New Zealand

### Official productivity statistics

In 2013, Statistics New Zealand presented the first official estimates of productivity for the education and training, and health care and social assistance industries. The initial series covered the period 1996–2011, and has subsequently been updated each year.

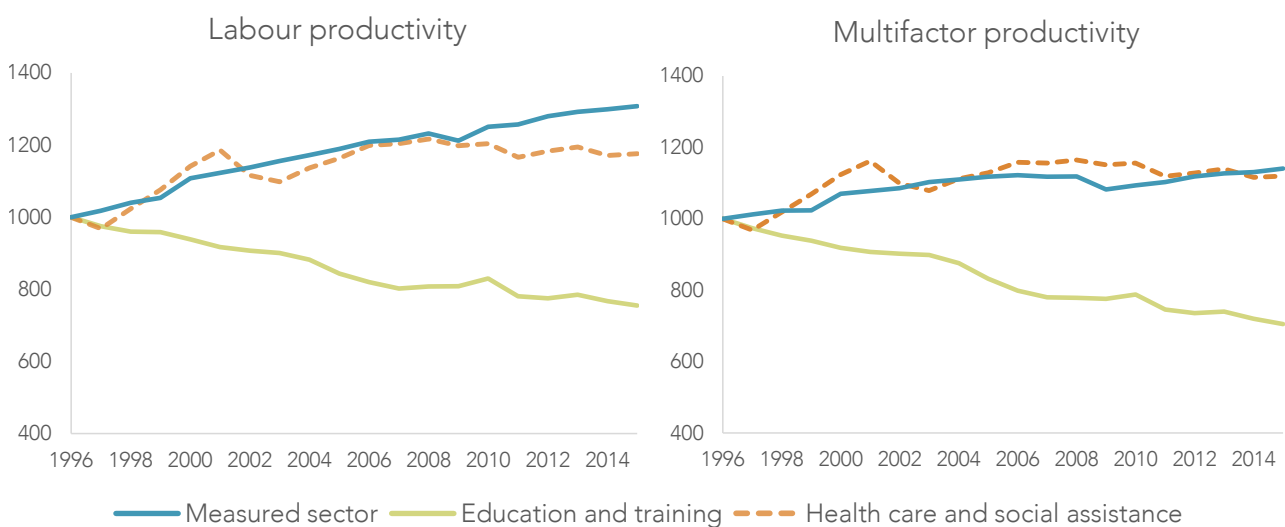
The productivity measures reflect output growth relative to input growth, they do not reflect other performance indicators such as effectiveness, and they are not quality adjusted. Also, the measures reflect the productivity of the industries as a whole, and hence cannot be used to attribute productivity change to a particular part of the industry. There could be considerable differences in productivity performance across the producers in a given industry (Statistics New Zealand, 2013).

Figure 2.2 shows the labour productivity and multi-factor productivity indexes for education and training, health care and social assistance, and for the measured sector.

- The education and training industry includes preschool, school, tertiary, and adult, community and other education. Both market and non-market activities are included in the measure, with non-market activity accounting for 87% of industry GDP in 2010. School education is the largest sub-industry (accounting for 50% of industry GDP), followed by tertiary education (33%), preschool education, and then adult community and other education.
- The health care and social assistance industry includes hospitals, medical and other health care services, and residential care services and social assistance. Non-market activity accounted for 57% of industry GDP in 2010. Hospitals are the largest component, accounting for 45% of industry GDP, followed by medical and other health care services (eg, general practitioners and dentists, which accounts for a further 34%), and residential care services and social assistance (21%) (Statistics New Zealand, 2013).
- The measured sector includes industries comprised of enterprises that sell their products for economically significant prices (for example, the manufacturing industry and the retail trade industry).

Labour productivity reflects the amount of output produced from each unit of labour employed (such as the number of hours worked). Multi-factor productivity reflects the efficiency with which a combination of productive inputs is used to produce output.

**Figure 2.2 Labour and multi-factor productivity indexes, 1996–2015**



Source: Statistics New Zealand, 2017a.

For both industries, multi-factor productivity tracks closely with labour productivity, reflecting the strong labour intensity of both industries (Statistics New Zealand, 2013). In the health industry, outputs grew strongly in the late 1990s, while input growth remained steady, leading to strong growth in productivity; this output growth was due to strong growth in market health activity (Statistics New Zealand, 2013). Since 2000,

input and output growth in the health sector have been broadly similar, resulting in flat productivity growth. Except for a period between 2008 and 2010, labour and capital inputs in the education industry have generally exceeded output growth, resulting in steadily declining productivity most years.

## Productivity measures for specific government services

In addition to Statistics New Zealand's official productivity statistics, government agencies and other organisations have attempted to measure productivity for specific government services.

### Productivity measures for hospitals

One example of service-level productivity measurement is the New Zealand Treasury's assessment of district health boards (DHB) using trends in case weighted discharges (CWDs), and average length of inpatient hospital stay (ALOS).

CWDs provide a standardised measure of the volume of hospital inpatient activity, excluding mental health and disability support. It measures total output (assigning greater weight to more complex procedures) per total cost of production (expenditure on medical and nursing personnel, clinical supplies, interest, depreciation, and capital charge).

The Treasury also uses average length of hospital stays (ALOS) as a measure of hospital efficiency.

ALOS can be reduced by measures such as advances in treatment technologies, more effective drugs, improved community and follow-up care, and more effective hospital administration. (New Zealand Treasury, 2017a, p. 41)

Shorter hospital stays are generally considered to be positive, as longer stays tend to reduce patient wellbeing and increase costs. But Treasury acknowledges that CWDs and ALOS are incomplete as indicators of DHB productivity:

There is a mismatch between inputs and outputs because inpatient CWDs are a subset of hospital activity and we are not able to exclude provider-arm inputs that relate to other (non-CWD) activity... [The measure] tells us nothing about service quality and does not recognise the benefits of DHB programs such as "releasing time to care" which aim to improve hospital processes allowing staff to spend more time with patients. Our ALOS analysis could be strengthened by assessment of hospital readmission rates (readmission rates would be expected to remain the same or fall as length of stay reduces if the system is working well). (2017, p. 42)

More broadly, Treasury (2017a) suggest that the measures should not be used systematically to manage performance, and that a more comprehensive monitoring framework covering the entire health system is required.

In another study looking at the efficiency of DHBs, Sandiford, Consuelo & Rouse (2017) use data envelopment analysis to measure the efficiency of DHBs at achieving gains in Māori and European life expectancy. The study finds reasonably consistent levels of efficiency in New Zealand DHBs and notes that this:

... may reflect a relatively uniform standard of health sector management, but it could also be because policy directions and service guidelines provided by the Ministry of Health allow little scope for any particular DHB to shine over the rest, or because improvements in health outcomes are largely determined by national historical trends (such as the decline in smoking and the obesity epidemic), that affect all DHBs similarly, even if their starting levels of LE [life expectancy] are quite different. (p. 128)

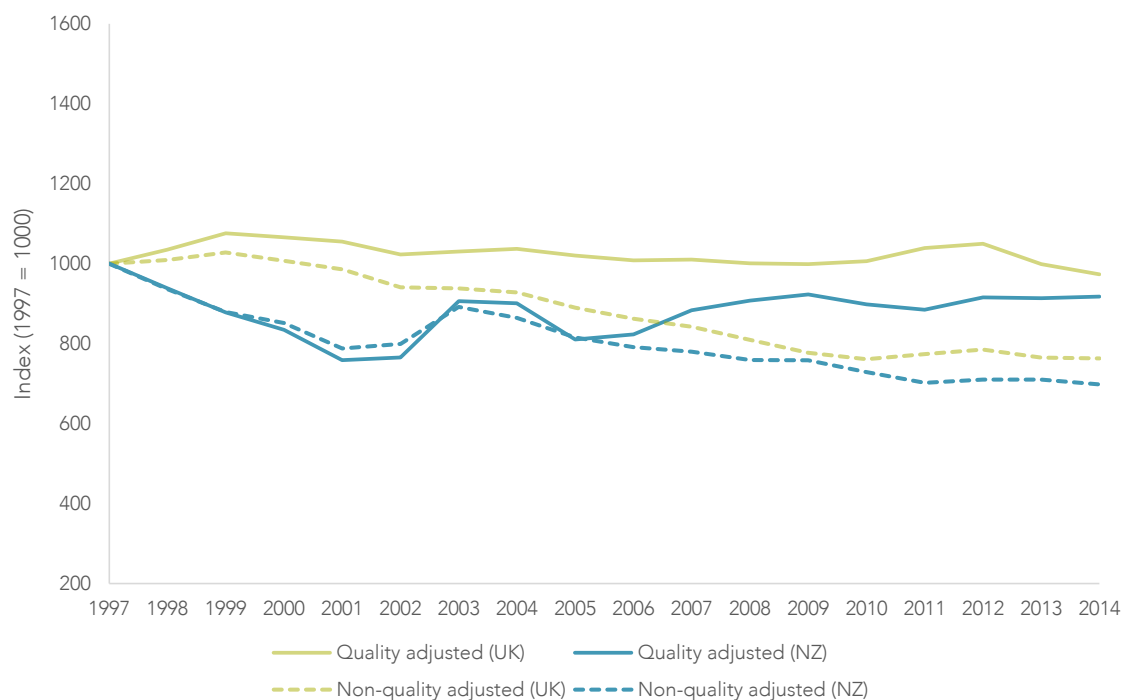
The study also finds that efficiency was significantly correlated with DHB financial performance, and that changes in life expectancy did not depend on the social characteristics of the DHB.

### Productivity measures for schools

Gemmell, Nolan and Scobie (2017) developed productivity measures for New Zealand schools using full-time equivalent teachers as the labour input, revenue as an indicator of total inputs, and student places provided as the output. The authors then applied adjustments to account for changes in the composition of the labour input, changes in student attainment, and changes in student outcomes (measured by their earnings).

One of the quality adjustments used in this study (student achievement of NCEA Level 2) is broadly comparable with a productivity measurement developed by the Office of National Statistics (ONS) in the United Kingdom (Figure 2.3). In both countries the unadjusted series imply decreasing productivity reflecting policy choices regarding smaller class sizes and more support staff. However, when quality adjustment based on student attainment is applied, average labour productivity growth is around zero in both countries between 1997 and 2014. Gemmell, Nolan and Scobie’s study illustrates both the importance and the difficulty of quality adjusting sector-level productivity data.

**Figure 2.3 Education productivity estimates in New Zealand and the United Kingdom, 1997–2014**



Source: Gemmell, Nolan and Scobie, 2017.

## Summary

As set out above, statistical agencies in several countries, including New Zealand, have made some progress in the development of public sector productivity measures. These efforts built on major methodological studies in the early 2000s. However, since the mid-2000s, international progress in the development of productivity measures has stalled somewhat. Productivity measurement in most countries is limited to health and education sectors, and little progress has been made in the area of quality adjustment.

In addition to official productivity statistics, there is a significant volume of academic research applying different techniques to the measurement of public sector productivity, often in specific areas of service (eg, hospitals). These studies provide useful insights – however, given that they are typically one-off pieces of research, they are limited in terms of providing a useful evidence base regarding public sector productivity performance over time.

### F2.1

There is a significant volume of research that applies different approaches to the measurement of state sector productivity, from which a lot can be learned. However, in recent years, progress in the development of robust and regularly-reported public sector productivity measures has slowed, both in New Zealand and internationally.

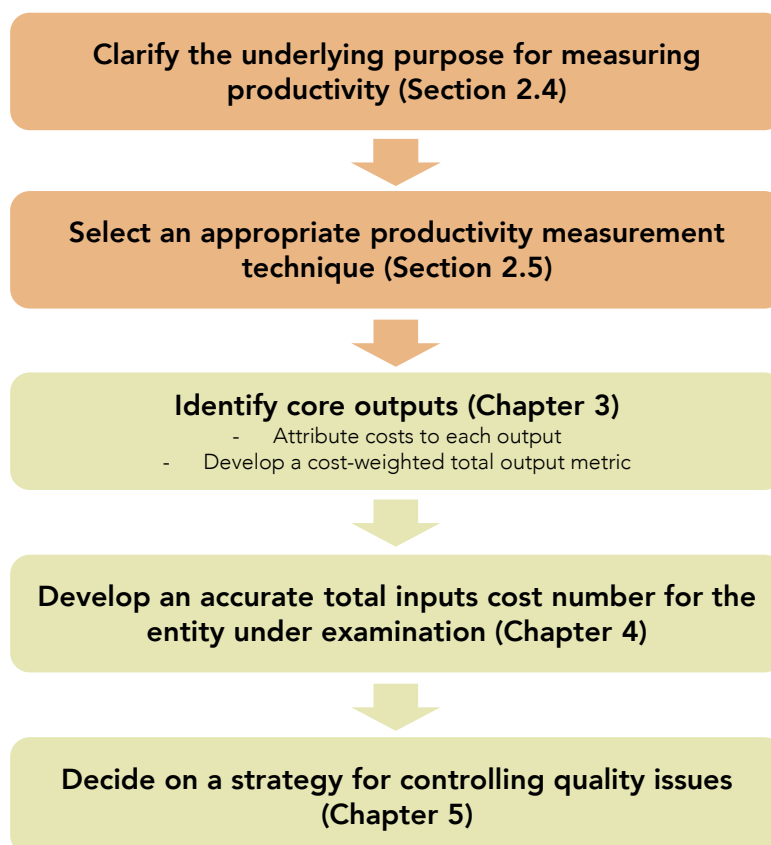
## 2.3 Key steps in developing a productivity measure

Figure 2.4 sets out the main steps involved in developing a productivity measure, drawing on the guidance set out by Dunleavy (2016).

Before beginning any technical aspects of developing a productivity measure, it is important to clearly establish the underlying purpose for measuring productivity (section 2.4). In turn, this assessment should then inform the selection of the measurement technique (section 2.5).

The first step in developing a productivity measure is to identify the core outputs that the relevant agency produces, identify the unit cost associated with each output, then develop a cost-weighted total output metric (Chapter 3). Following this, the total cost of inputs for the agency needs to be calculated (Chapter 4) – this will include costs such as staff salaries; the cost of intermediate inputs such as services provided by contractors; the agency’s annual running costs for property, equipment maintenance, and materials; and the capital costs for buildings and major pieces of equipment. A final step in the process is to decide whether it is necessary to account for changes in the quality of outputs, and if so, how to achieve this (Chapter 5).

**Figure 2.4 Main steps in developing a productivity measure**



Source: Adapted from Dunleavy, 2016

## 2.4 What kind and level of productivity measure?

Establishing a performance measure without consideration of the context within which it will be used, or attempting to use a single measure for too many purposes, can lead to unintended consequences. As such, a shared understanding of why productivity is being measured, and for whom, is an important prerequisite to developing a useful measure.

### Measurement for what purpose?

Performance measures can serve a number of distinct purposes. Various authors have described these purposes in a range of ways (eg, Behn 2003; Franco-Santos et al, 2007; Henri, 2006; Hood, 2007; State Services Commission/New Zealand Treasury, 2008; Verbeeten, 2008), but they can be summarised into three broad categories (Table 2.2).



**Table 2.2 Purposes of performance measurement**

Category	Key questions	Orientation
Learning	What works (or does not) and why (or why not)? How can policy or management be improved?	Change and the future
Steering & control	Are policies and programmes on track?	Control and the present
To give account	Can performance be justified?	Survival and the past

Source: Adapted from van Dooren et al (2015)

Productivity measures (in the sense of technical efficiency) are most typically used for steering and control goals, but can be used for all three purposes. “Steering” and “give account” uses of productivity measures tend to have a desired end state in mind (eg, achievement of a specific cost-per-output level, reductions in productivity variations between similar providers, downward movements in costs-per-output), whereas “learning” uses are generally less prescriptive about end states. Performance can be tracked over time, or compared between service providers.

There can be tensions between some purposes of a performance measure – especially accountability/control and learning. Gill & Schmidt (2011) note that a “focus on accountability and control tends to punish deviations from standards rather than providing an opportunity to learn.” (p. 16) Cooley (1983) argues that “indicators will be corrupted more readily if rewards or punishments are associated with extreme values on that indicator, than if the indicator is used for guiding corrective feedback” (p. 9).

Of course, performance information is often used to achieve multiple objectives, and the actual uses of performance information may vary from their intended purpose. Most information created by the government can be accessed through the Official Information Act 1982. A measure intended to learn or steer may end up being used by external parties for accountability purposes. This is largely unavoidable, and should not be seen as a reason to avoid the development of productivity measures. But it does point to the importance of ensuring that state sector agencies have a clear understanding of what can be inferred from productivity measures, so as to avoid being distracted by alternative interpretations developed by third parties.

In many cases, performance measures are just one input to a wider suite of decision-making tools and data. Where the primary objective of a productivity measure is to promote learning and improvement, it is worth thinking about:

- how directly the results of the measure will lead to a decision (eg, will other information be used to “triangulate” or understand the results?);
- the consequences of any decision that are based on the measure (eg, how significant will the impact on funding levels, managerial flexibility or team reputations be?); and
- where productivity measures present an incomplete or misleading picture of performance (eg, because of other, extenuating factors), how will seemingly contradictory decisions be explained?

## Measurement at what level?

The OECD (2017) distinguishes four levels at which productivity analysis can be conducted (Table 2.3). The sector, institution and activity levels broadly align with the levels and decision-makers discussed in the Commission’s terms of reference.

**Table 2.3 Levels at which public sector productivity analysis can be conducted**

Level	Examples	Terms of Reference level	Potential clients for results
Government	The entire public sector.		Minister of Finance, Cabinet, Parliament, the public.

Level	Examples	Terms of Reference level	Potential clients for results
Sector	The health system, the education system, judicial system, defence.	Meso (sector)	Vote Minister, Minister of Finance, Ministry Chief Executive, the public.
Institution	Hospitals, ministries, law courts, schools.	Micro (function or service)	Institution Chief Executive, Ministry Chief Executive, Vote Minister.
Activity	Tax administration, pension or benefit administration, surgery.		Agency/institution Chief Executive, Line Manager.

Source: Adapted from OECD, 2017.

In general, productivity measures are likely to be of most use where they fit the locus of control of the decision-maker. For example, a line manager will be most interested in assessing the outputs they are responsible for, and the inputs they control. Similarly, ministers could be expected to be interested in understanding the efficiency with which their vote is being used. Examples of how the levels of information required, and the sorts of productivity measures that might be developed, might differ between decision makers are outlined in Table 2.4.

**Table 2.4 Different decision-makers and types of productivity measures**

Decision maker	Locus of control / decision rights	Possible areas of interest	Potential measures	Examples
Vote Minister	Overall funding levels, mix of outputs.	Efficiency of overall spend.	Weighted average of total outputs / inflation-adjusted Vote expenditure	Weighted average of Vote Health outputs / real Vote Health expenditure.
Chief executive	Allocation and distribution of agency inputs.	Efficiency of key agency outputs.	Weighted average of gross agency outputs / inflation-adjusted agency expenditure.	Weighted average of Vote Health outputs / real Vote Health expenditure. Weighted average District Health Board outputs (eg, discharges, treatments) / real DHB expenditure.
		Drivers of efficiency trends.	Weighted average of agency outputs / labour and capital inputs.	Weighted average of Vote Health or DHB outputs / labour and capital expenses.
		Impact of changes in input mixes.	Specific agency output / labour and capital inputs.	Primary health outputs / primary health labour and capital expenses.
		Comparative performance.	Weighted average agency costs / real agency input costs.	Comparative efficiency of hospital case-weighted discharges.
Unit manager	Allocation and distribution of unit inputs	Efficiency of key unit output/s.	Unit output(s) / total real unit expenditure.	Real average cost of key output/s (eg, operations, diagnostic tests).

Decision maker	Locus of control / decision rights	Possible areas of interest	Potential measures	Examples
		Drivers of efficiency trends.	Unit output(s) / unit labour and capital inputs.	Number of operations, diagnostic tests / labour and capital expenses.
		Impact of changes in input mixes.	Substitutable outputs / total real unit expenditure.	Real unit cost of treatment outputs (eg, cataract treatments).

## F2.2

Productivity measurement can occur at a number of levels (eg sector, institution, service). Measures are likely to be most useful where they align with the decision rights or scope of control of decision makers.

## Setting measures to answer the right productivity questions

There are choices about the types of outputs and inputs that should be used in constructing productivity measures. Which types should be measured depends very much on the question you are trying to answer.

- For inputs, a key choice is whether to use *total inputs* (eg, expenditure); or, to include only *one production factor* (eg, labour, capital).
- For outputs, a key choice is whether to use *gross output* or *value added*. Value added measures remove intermediate inputs from gross output.

Isolating specific production factors in the inputs allows the impact of changes in capital or labour on overall productivity to be assessed. A total input measure permits assessments of the efficiency with which all resources are being used (“multi-factor productivity”). Because productivity gains often occur as a result of input substitution (eg, greater use of technology, such as the cataract treatment example cited in Chapter 1), or new governance and management techniques (eg, the introduction of PHARMAC as a purchasing agent, discussed in Chapter 1), the domain of measurement should be designed to capture such changes.

Gross output measures are useful for understanding the *total* output of a sector or organisation, while value added measures are useful for assessing the *marginal additional value* added. Statistics New Zealand (2010) illustrates this difference with a health sector example:

...if interest lies in understanding the marginal extra value added by the health system (for example, the fact that medications are typically bought in and not produced by the government health sector, so are not part of its value added), then a value added single or multifactor productivity methodology should be constructed. If interest lies in understanding the total output of the health system, then a productivity measure based on gross output should be constructed. (p.19)

If the objective of measurement is to discover or learn whether changes in input mixes or activities have led to greater efficiency, outputs should be defined at a sufficiently high level so as to capture the effect of changes in underlying inputs and activities.<sup>9</sup> Statistics New Zealand (2010) cites the example of “treatments that have the same outcome for patients”:

...grouping together psychotherapy and drug therapy where they are substitutes should mean that any gains in efficiency through substitution will be captured in the productivity index. (p.14)

<sup>9</sup> This technique has been used in some health productivity studies (Frank et al, 1999; Shapiro et al, 2001).

**F2.3**

Different types of output and input measures will answer different questions. For example, gross output measures are useful for understanding the total output of a sector or organisation, while value added measures are useful for assessing the marginal additional value added (or removed).

## 2.5 Selecting a measurement technique

The key measurement techniques covered in this chapter are simple ratio analysis and frontier analysis (data envelopment analysis and stochastic frontier analysis).

### Ratio Analysis

Ratio analysis is a widely-used approach to productivity measurement. In this approach, productivity is the ratio of outputs over inputs. Two key measures are often calculated: labour productivity (outputs over labour inputs) and multi-factor productivity (outputs over total inputs). These ratios are most useful when tracked over time or compared across organisations or governments. Tracking a measure over time can show increases and decreases in the productivity of an organisation or part of an organisation. Comparisons in the productivity measures of different organisations (or different parts of the same organisation) can show how relative performance is changing.

### Frontier analysis

A limitation of ratio analysis is that it does not illustrate the distribution of performance across organisations within a sector. Ratio analysis can tell the researcher about average performance, but a change (or not) in the average can mask a variation in performance. Frontier analysis can, in contrast, provide a relatively rich picture of public sector productivity and help explain the performance of organisations. For example, it can explain whether relatively poor performance of a sector is due to a lack of productivity growth among the best performing organisations (the frontier), or best practices failing to diffuse throughout a sector (eg, from the best performers to the laggards). Yet while potentially valuable, frontier approaches can be relatively data and resource intensive.

Gemmell, Nolan and Scobie (2017) identified the following general stages in a frontier analysis:

- *Define the decision making units (DMUs)* – DMUs are a central feature of frontier analysis and can refer to an individual, firm, public agency (eg, a school or hospital), region or country.
- *Calculate the efficiency frontier* – the efficiency frontier (sometimes called the reference set) is made up of DMUs whose input levels are the lowest for any given level of output; this becomes the set against which the efficiency of all DMUs can be assessed.
- *Estimate the distance of DMUs to the efficiency frontier* – each DMU receives an efficiency score that is determined by their performance relative to that of the best performers.

There are two broad approaches to estimating frontiers: non-parametric and parametric.

Non-parametric approaches make no allowance for “random noise” such as measurement errors or other random shocks. As a result, any observation falling within the frontier is treated as technically inefficient. The most widely used non-parametric approach is Data Envelopment Analysis (DEA).

Parametric approaches, on the other hand, do not attribute all of the observed differences between DMUs to differences in technical efficiency, as they allow for measurement error and other random noise. As a result, no DMUs necessarily need to lie on the efficiency frontier (Gabbitas & Jeffs, 2008). For parametric methods, a widely used approach is stochastic frontier analysis (SFA).

Box 2.2 sets out some considerations that should influence the selection between these two frontier techniques.

**Box 2.2 Factors affecting the choice between frontier techniques**

- Cases with less heterogeneous samples are more suited to DEA. SFA is more suited to more heterogeneous samples. DEA is more sensitive to heterogeneity in the sample (influenced by outliers) and will tend to give lower average efficiency scores (although not consistently). The regression approach of SFA means outliers are given less weight.
- Cases where output supplied is subject to variable and/or unpredictable client demand are less suited to DEA. Unpredictability of client demand can introduce a source of variance in outputs and weaken the relationship between inputs and outputs. SFA is better suited to coping with unpredictable demand.
- Both methods can deal with cases where exogenous variables influence operating environments. Where these variables could be an important consideration, a DEA approach of restricting the comparison set (to DMUs that have similar or less favourable operating environments) is likely to be less suitable. Other DEA approaches or an SFA approach based on regression analysis would be better.
- SFA requires the parameters of the production function and the random error term to be estimated. DEA is more suitable for cases where these parameters cannot be feasibly estimated, such as where there are a limited number of observations available for robust regression analysis.

Source: Gemmell, Nolan & Scobie (2017).

Further detail about the DEA and SFA can be found in the Commission's working paper *Public sector productivity: Quality adjusting sector-level data on New Zealand schools* (Gemmell, Nolan & Scobie, 2017). SCRCSSP (1997) and Gabbittas & Jeffs (2008) also provide practical guidance regarding the use of DEA and SFA respectively.

## Matching the question to the technique

As noted earlier, an important early step in undertaking productivity measurement is to consider what questions are to be answered and for whom. Table 2.5 shows possible questions that an organisation may have and how these relate to potential measurement techniques.

**Table 2.5 Identification of possible measurement techniques**

Productivity question	Measurement option	Data requirements	Suggested interpretation
What has happened to an organisation/sector's overall productivity?	Ratio analysis or unit cost, over a time period.	Outputs will need to be appropriately weighted and combined to create an agency/sector index.	Ratio analysis shows the change in productivity levels for an organisation or sector.
Have changes in the input mix changed output levels?	Ratio analysis, over a time period.	Input data should clearly identify input categories (ie, labour, capital) and quality adjust for significant changes (eg, skills, technology).  To provide a base, set an analytical time period before the input mix change occurred.	Ratio analysis results in an index which allows tracking of changes in productivity over time.  Productivity indices show change in productivity relative to a base year, and allow calculation of changes between years.

Productivity question	Measurement option	Data requirements	Suggested interpretation
How do the different parts of an organisation/service provider compare? How much more efficient could divisions/service providers be?	Ratio analysis, comparing decision-making units.	If the aim is to understand the drivers of different performance levels, input data should identify input categories and quality adjust appropriately.	Ratio analysis shows the change in productivity levels in each part of the organisation.
	Frontier analysis.	Advanced techniques require information on inputs and outputs for each of the different parts of the organisation.	More advanced techniques enable comparisons of overall productivity levels, as well as changes, in each part of an organisation.

As an example, the Commission's case study into Early Childhood Education (Green, 2017) sought to understand what has happened to the overall productivity of the ECE sector over the past decade. Using publicly available data, the study presents an index of labour and multi-factor productivity, which show changes in the sector's productivity over time.

The measures show that multi-factor productivity in the ECE sector has fallen by an average of 3.4% per year between 2001/02 and 2014/15. However, this result needs to be treated with caution, given the absence of good data on parents' financial contributions to ECE. Between 2004/05 and 2012/13, labour productivity grew by an annual average of 0.4%. When adjusted for changes in the qualifications of the teaching workforce, labour productivity growth was either flat or negative (depending on the technique used).

## 3 Measuring outputs

### Key points

- Outputs are the final goods or services that are produced for use by an individual or other organisation. After a decision has been made that the state sector should produce certain outputs, productivity measures show how efficiently those outputs are produced.
- Ideally, outputs should comprehensively cover the range of functions performed by the agency. Where this is impractical, it is important that the range of outputs capture the fundamental goods and services produced by the agency.
- The state sector produces a diverse range of outputs. Measuring outputs for some services, for example, processing benefit payments, is likely to be relatively straightforward owing to their routine, high-volume and often automated nature.
- But for other state sector services, output measurement can be more challenging owing to factors such as joint production, co-production, and collective consumption. However, these challenges, along with data availability issues, are not insurmountable.
- Where a state sector entity produces more than one output, the outputs need to be weighted so that they can be aggregated into a single measure. This can be done using several different methods – however, it is widely recommended that outputs are weighted based on their volume and their cost of production.

### 3.1 Introduction

Much of the discussion around measurement of productivity in public services relates to the accurate measurement of outputs. In the market sector, the measurement of outputs is more straightforward and can be calculated based on their volume and market price. Prices (and the assumption of functioning markets) can act as a proxy for value, or willingness to pay, and as a means of aggregating diverse outputs. For public services that are provided without charge, or below the cost of production, alternative approaches to output measurement are required.

This chapter examines the steps involved in measuring outputs and presents guidance on approaches for state sector agencies to measure output as part of a productivity measure. The chapter begins by discussing what constitutes an output, and how outputs should be differentiated from activities and outcomes. Section 3.3 considers two issues to consider when selecting which outputs to measure; ensuring that outputs are appropriately specified in terms of their level of aggregation; and appropriate output coverage. Section 3.4 sets out how outputs vary depending on the type of service that is under examination. The chapter concludes (Section 3.5) by setting out options for weighting different outputs in order to create a single output metric.

### 3.2 Defining outputs

An initial step in measuring outputs, is to define what an output actually is. Outputs are the final goods or services that are produced by the entity under examination for use by an individual or other organisation.

#### Outputs vs outcomes

Outputs are distinct from *outcomes*. Outcomes are the changes in society, the economy or the environment that result from outputs (State Services Commission and New Zealand Treasury, 2008). Outputs and outcomes are easily confused, and the labels are sometimes misapplied both inside and outside government. Table 3.1 provides examples of the differences between outputs and outcomes.

**Table 3.1 Outputs and outcome in the state sector**

	Outputs	Outcomes
<b>Hospitals</b>	Hospital discharges for different diagnosis related groups.  Number of treatment courses for specific medical conditions.	Healthy population through - for example, recovery from injury or illness and reduction in preventable diseases.
<b>Schools</b>	Number of student places provided.	Well-educated population.  Young people who are confident, connected, actively-involved lifelong learners.
<b>Courts</b>	Number of cases heard, hearings held, mediation sessions conducted and fines collected.	Cases are resolved in a procedurally fair and just manner.  Safe communities.  Public trust and confidence in the justice system.
<b>Work and income services</b>	Number of individuals that move off support and into sustainable employment.  Number of young people that move off benefits and into education, training, or work-place learning.  Number of people accessing emergency housing.	More people are in sustainable work and out of welfare dependency.  Fewer people commit fraud and the system operates with fairness and integrity.  More people are able to participate and contribute positively to their communities and society.

Although achieving a certain outcome or outcomes is usually a primary objective for state sector agencies, outcomes are not included in the measurement of productivity. The main reason for this is that outcomes are typically influenced by numerous factors, including those that are outside the control of the state sector. For example, overall health outcomes, such as life expectancy, are affected by factors including lifestyle choices, environment, and education (Sharpe, Bradley & Messinger, 2007). Measuring outcomes is conceptually different from measuring efficiency (Chapter 1).

Several inquiry participants raised concerns about focusing exclusively on outputs, without sufficient consideration of outcomes:

Efficiency alone is an insufficient focus for productivity measures. It is simple to measure the number of services delivered to clients, but if these services are not meeting the client's needs or resulting in improved outcomes, can the organisation be said to be productive? (Youth Horizons, sub. 4, p. 1)

Technical efficiency in producing outputs that are ineffective in producing the desired outcomes can do harm if it encourages a shifting of resources away from less technically efficient processes that produce outputs that are more closely tied to desired outcomes. (New Zealand Initiative, sub. 7, p. 2)

These submissions emphasise that technical efficiency alone is not a complete measure of state sector performance – but this does not diminish its importance. Technical efficiency is measured after a decision has been made that the state sector should produce certain outputs; after this decision has been made, productivity measures play the important role in ensuring that outputs are being produced in the most efficient way possible.

The fact that an efficiency measure is not an appropriate tool for evaluating decisions about what outputs to produce does not mean that efficiency should not be measured or cannot be used.



**F3.1**

Productivity measures show how efficiently outputs (the final goods or services produced by an entity for individuals or organisations) are produced, and hence are not a complete measure of state sector performance. Other performance information including measurement of outcomes should be captured through separate processes.

## Outputs vs activities

Outputs are also distinct from *activities*. Activities describe the individual tasks that public service agencies perform that contribute to the delivery of an output. Typically, numerous activities will be involved in the production of an output. Dunleavy (2016, p. 5) notes that outputs need to be restricted to “complete activity packages (not parts of operations) or to finally-delivered services (akin to end products in firms)”. When considering the productivity of a hospital for example, an output might be the complete treatment of a patient.

## 3.3 Selecting which outputs to measure

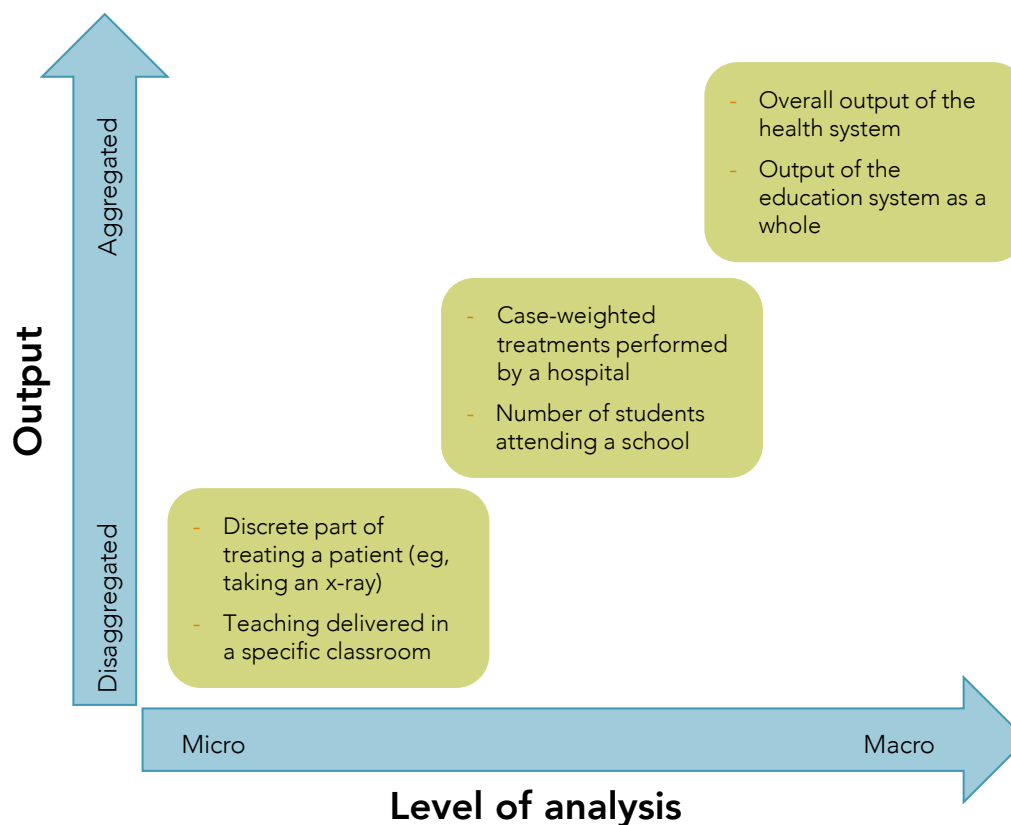
### Disaggregation of outputs

As noted above, activities are the individual tasks that state sector agencies perform that contribute to the delivery of an output. In practice, this distinction is not clear-cut. Whether something is classified as an activity or an output will depend on the level of the analysis. In the case of the police, Hatry (1975) suggests that at least six levels of productivity analysis are possible:

1. the productivity of the individual police officer or police employee;
2. the productivity of police units - for example, different police districts or different stations;
3. the productivity of different kinds of police units - for example, investigative units, maritime units, armed offender’s squads, police dog teams;
4. the productivity of the police as a whole;
5. the productivity of the crime control system including both police activities and private activities to reduce crime; and
6. the productivity of the total criminal justice system, including the police, the courts, corrections and social service agencies.

Different levels of output disaggregation will be needed depending on the type of analysis being undertaken. For an emergency department, an output might involve an initial diagnosis and course of treatment, even if the patient was then transferred into another hospital department for further treatment. This represents the complete activity of the entity (the emergency department) under examination. However, when examining the productivity of a hospital as a whole, the services performed by the emergency department might be considered as one activity in the overall output of treating a patient. Similarly, an examination of productivity in the police maritime units will measure a different set of outputs than those measured in an analysis of the New Zealand Police as a whole.

Figure 3.1 shows examples of different types of outputs that might be appropriate depending on the level of productivity analysis.

**Figure 3.1 Matching outputs with the level of analysis**

Importantly, outputs should be appropriately disaggregated so that they are “controllable by the agency, or at least include only those aspects of performance that the agency has reasonable influence over and is therefore able to meet” (OECD, 2001, p. 7). A productivity measure that is based on outputs over which the relevant agency has little control is unlikely to be helpful in terms of accountability, steering, or learning for performance improvement. Equally, outputs should not be so finely disaggregated that the agency has little discretion or management flexibility in how they perform their functions (OECD, 2001).

## Output coverage

In addition to ensuring that the aggregation of outputs is aligned with the function of the entity that is under examination, the wide range of outputs that the state sector produces raises questions about the breadth of information that should be collected and how best to use that information to construct output measures (Simpson, 2009).

Some argue comprehensive coverage of outputs for the monitored body is ideal, since incomplete measurement may lead to misleading or inaccurate results (Statistics New Zealand, 2010). For example, if only a selected range of outputs are used to measure productivity it may create incentives for agencies to focus on measured outputs at the expense of unmeasured outputs (Simpson, 2009). The New Zealand Council of Trade Unions (sub. 9) raised similar concerns:

Elevating any subset of measured outputs to ‘core’ status runs counter to the flexible approach described above. It risks distorting the operations of the organisation or programme if more effort is devoted to improving that indicator at the expense of its complete set of objectives. (p. 6)

However, given the range of public sector outputs and limitations in data availability, comprehensiveness may not be feasible. Atkinson (2005) notes that a balance has to be drawn between coverage and cost of implementation. There are likely to be diminishing marginal returns from attempting to measure all outputs, particularly at the sector or agency level. One submitter noted this point:

It is neither necessary nor desirable to measure every single output of any sector, service or function. The Pareto principle states that, for a lot of events, roughly 80% of effects come from 20% of the causes... This 80/20 rule can be applied in this context by focusing on the critical 20% of functions of any sector which would produce roughly 80% of the outputs. This would maximise the cost/benefit ratio for

the project, deliver the most gains in productivity, and avoid wasting time dealing with problems which are trivial. (Grobler, sub. 5, p. 6)

While there may be good reasons that make it impractical for a productivity measure to capture the complete coverage of an agency's outputs, it is important that the fundamental goods and services produced by the agency are captured (SSC and Treasury, 2008; Dunleavy and Carrera, 2013). Dunleavy and Carrera (2013) suggest that a key question to ask about a state sector agency is "what its broad mission is, and what few main outputs capture that mission and can be cost-weighted in a reasonably accurate manner" (p. 36).

Similarly, Statistics New Zealand (2010) suggest that if comprehensive coverage is not possible, then the second-best option is to aim for representativeness. They note that the growth rate for some outputs might reasonably be considered to be representative of growth rates in other areas where data is not available. But it is unclear what constitutes good grounds to assume that productivity growth in unmeasured services is the same as measured services. Hence, a cleaner approach is to aim for the most comprehensive coverage possible, and to be clear about any outputs that are not captured within the measurement.

### Output coverage and data availability

The choice of which outputs to measure is likely to be influenced by existing data availability, and the ease with which any additional data might be gathered. Statistics New Zealand (2010) notes that output coverage is typically based on what information and classifications are already available, rather than on purity of concept. Atkinson (2005, p. 47) argues against this approach, and notes that services provided to households and firms should be the starting point for outputs:

...the procedure of defining direct output indicators within a government function should start by seeking to identify the services provided by government to households and firms, and attempts made to find data to reflect these services as comprehensively as possible, with appropriate allowance for quality change. The services should be the starting point, not the available indicators.

While the service provided to households and firms is a sensible starting point in terms of outputs to measure, there is a need to be pragmatic given that state sector cost data is frequently incomplete. The absence of complete data should not necessarily prohibit measurement efforts. Instead, the limitations associated with measures should be clearly articulated, and measurement efforts should be undertaken in parallel to efforts to improve data collection.

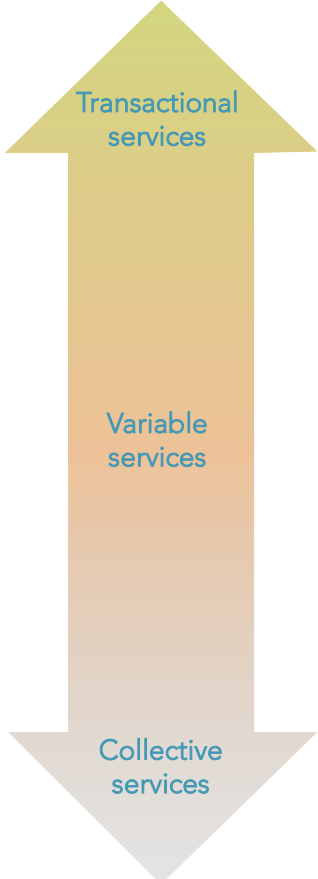
#### F3.2

Data limitations should not necessarily be a barrier to productivity measurement. Any data limitations should be clearly articulated, and measurement efforts should be undertaken in parallel with work to improve data collection.

## 3.4 Challenges in measuring outputs

The state sector produces a diverse range of services. For some services, outputs are relatively easy to identify and straightforward from a measurement perspective, while others are more challenging. Figure 3.2 sets out three broad classifications of services and some defining features of their outputs.

**Figure 3.2 Typology of state sector outputs**

	Features	Examples
 <p>Transactional services</p>	<p>Outputs tend to be:</p> <ul style="list-style-type: none"> <li>- Subject to a clearly defined production process</li> <li>- High volume</li> <li>- Standardised and repetitive</li> <li>- Delivered with relatively little involvement of the consumer</li> </ul>	<p>Payment of benefits</p> <p>Taxation</p> <p>Issuing passports</p>
<p>Variable services</p>	<p>Outputs tend to be:</p> <ul style="list-style-type: none"> <li>- Measureable (outputs can be counted)</li> <li>- Variable in terms of the production process</li> <li>- Often co-produced (involving input from the service recipient) or jointly produced (by more than one agency)</li> </ul>	<p>Teaching</p> <p>Healthcare</p> <p>Criminal investigations</p>
<p>Collective services</p>	<p>Outputs are difficult to specify because services tend to be:</p> <ul style="list-style-type: none"> <li>- Consumed collectively</li> <li>- Intangible (often they will relate to prevention of something occurring)</li> </ul>	<p>Threatened species conservation</p> <p>Fire prevention</p> <p>Crime prevention</p>

### Transactional services

Outputs from transactional services will tend to be:

- standardised, high volume and repetitive
- delivered with relatively little interaction with or involvement of the consumers [when done well]...
- relatively easy to specify in advance, with actual performance also relatively easy to measure. (OECD, 2001, p. 28)

An example of a transactional service within the scope of this inquiry is the payment of benefits – measuring productivity for income support benefit payments is examined in the Commission’s case study on using administrative data to develop productivity measures for the Ministry of Social Development (Moore & Hayward, 2017).

Transactional outputs are generally less challenging to measure. These outputs are usually “supported by clearly articulated, comprehensive procedure guides and operating manuals that contain the rules to be followed and standards to be met during the various stages of the production process” (OECD, 2001, p. 28). As such, the quantity of completed units will implicitly address qualitative dimensions (OECD, 2001). Dunleavy and Carrera (2013) suggest that for these types of outputs, quality can be assumed to be relatively constant. They recommend that researchers take note of any substantive failures of quality control when presenting productivity data, rather than seeking to quality adjust the output numbers.

### Variable services

A second category of services, described here as “variable”, includes services such as teaching and individual healthcare. For these services, outputs can be readily defined and counted, but they are subject to much greater variation than transactional services in terms of their production process. They are often delivered with a high degree of interaction with, or involvement of, the service consumer. The variability in

production process introduces much greater scope for differences in quality making quality adjustment more important (Chapter 5).

Many of the services included in this category are what Wilson (cited in Gregory, 1995) refers to as “craft” and “coping” services. Craft services are “dependent on the activities of highly trained professionals exercising a large degree of autonomy from day-to-day managerial supervision” (Gregory, 1995, p. 173). Examples include criminal detection work by police and the provision of medical care. Provision of coping services also involves considerable discretion by practitioners and attempts to change aspects of human behaviour. Examples include probation work and social work.

Outputs can generally be measured for craft and coping services; therefore, productivity measures can be developed. However, due in part to the high level of discretion involved in performing these tasks, staff involved in service delivery will usually be best placed to assess what matters for productivity and performance more broadly. This reiterates the importance of including staff in the development and design of performance metrics (Chapter 6).

Another characteristic of some outputs included in this category is that they are co-produced (meaning their production involves the input of both the state sector agency and the service recipient), or jointly produced (the service is produced by two or more state sector agencies).

### **Outputs that are co-produced by an agency and the service recipient**

Many services delivered by the state sector involve co-production by the government agency and the service recipient. Simpson (2009, p. 260) notes that this co-production can play out in two ways:

First, the characteristics of the individuals using the service will make a difference to the measure of gross output, and these individuals can in a sense be thought of as (intermediate) inputs themselves. For example a pupil’s ability will affect a measure of gross output such as examination results.

Second, the effort expended by an individual in conjunction with the inputs of the service provider, such as the time and effort a pupil devotes to homework and their behaviour in the classroom, may also affect output.

In the latter case, the amount of effort expended by a school student is something that a school can influence. An excellent teacher who is able to motivate students to apply themselves in the classroom and to complete their homework, is likely to be able to produce higher quality outputs, and hence it is appropriate that this is reflected in any productivity measure. Similarly, different health professionals may have more or less success in co-producing improvements in health, such as stopping smoking.

By contrast, state sector agencies cannot easily influence the innate characteristics of the population they are co-producing a service with (although they can make efforts to tailor their services to different population groups). The extent to which the characteristics of service recipients affects the ability of state sector agencies to produce outputs, depends on the extent of the input from the customer (Linna et al, 2010). In some cases, the impact is significant, for example, Engler (2010) found that prior school achievement was the most important factor associated with academic performance in tertiary education. The likelihood of students with below-average school achievement passing all of their first-year tertiary courses ranged from about 0.3 to 0.8. For above-average students at school, the likelihood of passing most first-year courses ranged from 0.8 to near certainty.

Where consumers play a significant role in the co-production of outputs, productivity measures can be improved by applying adjustments that take account of the characteristics of the individuals using the service – enabling measurement of the value added through the contribution of the state sector (Chapter 5).

### **Measuring outputs that are jointly produced by multiple agencies**

Outputs that are jointly produced by multiple agencies will often be apparent when taking a macro-level view of state sector productivity. For example, in the health sector, an output might be “number of patients treated”. However, this treatment might traverse primary and secondary care, which complicates measurement efforts:

...how does one distinguish between the value added by the (private sector) general practitioner, the (public sector) outpatient appointment, and the (public sector) inpatient day care or inpatient stay? (Statistics New Zealand, 2010, p. 23)

Statistics New Zealand (2010) notes that in the case of the health sector, costing information is not always available to distinguish which part of the sector is responsible for what proportion of the care. In these cases, they suggest using whatever information is available to approximate proportions. Where this is not possible, another solution is to use qualitative assessments, "drawing on expert opinion to arrive at a set of reasonable assumptions" (Statistics New Zealand, 2010, p. 23).

## Collective services

Collective services are provided to society as a whole (eg, defence). Collective services are provided for the benefit of the population at large, rather than to individuals separately; no one person can be separated out to receive the service or be excluded from receiving the service. This makes it very difficult to determine the volume of output of collective services that are consumed in the economy (ONS, 2009).

Collective services often relate to efforts to avoid adverse events from occurring - for example, crime prevention or conservation of a threatened species. In these cases, counting outputs may not be detailed enough to capture all outputs of a service provider. For example, counting the number of fires successfully extinguished does not adequately cover the full range of services performed by a Fire Service.

If a fire service in one area were more proficient at fire prevention they would consequently need to tackle a lower number of fires, and if only the latter activity were used then the fire service would be incorrectly assigned a relatively low level of output. (Simpson, 2009, p. 254)

Similar problems could emerge when measuring outputs for any state sector agency that performs preventative functions.

Owing to the difficulty of measuring outputs for collectively consumed services, most countries have retained the outputs = inputs approach. This means that for collective services, which typically account for around 30% of public sector spending in OECD countries, any potential changes in productivity levels are not captured (OECD, 2017).

The Atkinson Review (2005) presents several options for measuring collective outputs.

- Assume that the productivity growth of collective services is equal to that of measured services – ie, assume that the productivity of public health campaigns is equal to measured health services. However, as noted earlier, it is not clear what a reasonable basis would be for making such an assumption.
- Divide expenditure into two, applying direct output measures where possible, and reverting to an input method for the remaining collective services. In effect, this means that collectively consumed services are excluded from the productivity measure.
- Measure outputs where possible, and use activity indicators for collectively consumed services. In the case of fire prevention services, the total number of hours spent delivering prevention activities might be an appropriate substitute for outputs (Office of the Deputy Prime Minister, 2005).

Of these options, the use of activity measures is the most promising. In the case of defence for example, counts of activities such as training exercises, could be used as a substitute for output. Another approach is to use an agency's "readiness" as an output measure. For example, in the case of defence services, output measures could include the ability to deploy a certain capability within a certain timeframe. Although such a measure does not fit the ideal definition of an output, as it is not a service that is delivered to a final customer, it is possible to use it to develop a productivity measure, by recording the level of inputs required to maintain that level of response capability. Given that a readiness output is not actually delivered to an external customer, an agency using such a measure would need to develop ways to test their readiness. Without this, there may be an inclination to simply rely on input measurement (ie, assuming that the expenditure of a certain level of inputs = a certain level of readiness).

### 3.5 Calculating a total output metric

A final step in the process of measuring outputs is to calculate a single output metric. Where a productivity measure includes more than one output, a simple count of the total number of outputs produced is unlikely to give an accurate picture of how productive an entity is. For example, when the police conduct a criminal investigation, the process they follow and the resources they deploy will vary significantly depending on the nature of the case in question.

A large part of police work consists of investigating crimes committed. But not all crimes are the same: on average, a violent crime has much more time and resources spent on it than do other types. (Pritchard, 2003, p. 37)

Dunleavy (2016, pp. 5–6) notes that for most private firms, the presence of sales volumes and prices makes the process of calculating a total output metric relatively straightforward:

... suppose a firm has two products, the first X priced at \$5 and selling 20,000 units and the other Y priced at \$10 and selling 5,000 units. Its total output is thus:  $(\$5 * 20,000) + (\$10 * 5,000) = \$150,000$ . Price is important here in two ways. First, it allows us to easily price-weight across completely dissimilar products. Second, in competitive markets with consumer sovereignty, we can make welfare implications about the sales patterns observed – in this case that consumers would not freely pay \$10 for product Y compared with \$5 for product X unless they were getting commensurate benefits from it.

This approach is not feasible for the state sector (except for a small number of special cases), because outputs are not priced, and many outputs must be consumed whether citizens or enterprises wish to do so or not (Dunleavy, 2016).

Diewert (2017, p. 2) suggests three methods for valuing state sector outputs, ranked in order of their desirability:

- *First best:* valuation at market prices or purchaser's valuations;
- *Second best:* valuations at producer's unit costs of production;
- *Third best:* output growth of the public sector production unit is set equal to real input growth and the corresponding output price growth is set equal to an index of input price growth.

The third best option (assuming that inputs=outputs) has been used historically as a convenient way to overcome measurement difficulties in the public sector. But, by definition, this approach will result in productivity growth of zero; therefore, is of little value. The following sections examine the advantages and disadvantages of Diewert's (2017) first and second-best approaches.

#### Valuation at market prices or purchaser's valuations

One method for valuing outputs is to obtain price information from comparable services provided in the private sector. Atkinson (2005, p. 89) provides two examples of how this approach could be used in practice:

- In the case of road use, "we may attach value weights to passenger miles and to freight tonne miles, based on the alternative costs of using rail".
- Another example would be the provision of personal care by social services, where there is a parallel private market. The price that people are willing to pay for daily care in the private sector can be used for the marginal valuation.

Simpson (2009, p. 255) also notes that comparable prices in the private sector can be used to value state sector services, but notes the following caveats:

... private sector alternatives might differ in their scope and characteristics; private healthcare might offer reduced waiting times and higher quality accommodation. In addition the characteristics of individuals using private sector alternatives, for example their underlying health, may differ from those using public sector provision and may also affect the price. Hence in each case questions would remain about how reliably these methods would capture the relative valuations of different goods.

In addition, the sheer size of the public sector as a provider of certain services (eg, healthcare) can skew the price of parallel services provided in the private market. In other cases, (eg, police, fire service) there is no comparable private market (Parker, Waller and Xu, 2013).

Willingness to pay (WTP) methodologies can also be used to estimate the value that the public places on certain outputs. WTP seeks to establish in advance what somebody would be prepared to pay in order to receive a good or service - for example, a certain health intervention. There are two broad approaches for estimating WTP (Accent and Rand Europe, 2010):

- *Revealed preference methods* – these approaches involve the observation of preferences revealed by actual market behaviour and represents real-world evidence on the choices that individuals exercise. Examples include the premium that individuals are willing to pay to live in the catchment area of particular schools, how long individuals are prepared to wait for a certain service, or how far they are willing to travel to access a certain hospital (Simpson, 2009).
- *Stated preference methods* – these approaches make judgements about potential impacts in the absence of real-world evidence on how individual consumers may respond. In an environmental context this might involve asking how much an individual would agree to pay for avoiding a degradation of the environment or, alternatively, how much they would ask for as compensation for the degradation. Alternatively, people can be asked to make trade-offs among different alternatives, from which their willingness to pay can be estimated.

The New Zealand Initiative (sub. 8) supported the investigation of a wide range of techniques for assessing the value of non-market outputs including revealed preference techniques. In particular, the submission notes that the “risk-premium that obtains for particular jobs, for example, can provide a reasonable benchmark for the value that workers place on avoiding relatively small risks” (p. 3). The New Zealand Initiative suggests that such data could be used to standardise the valuation of interventions that offer a similar reduction in risk across different policy areas. Other submitters were less optimistic about willingness to pay – for example, New Zealand Council of Trade Unions (sub. 9) and the Public Service Association (sub.14) argue that “public services are not market goods and there is no value in a subjective measure based on the assumption that they could be treated as such”.

One potential way that outputs could be weighted in the health sector, is through the use of quality-adjusted life years (QALYs). QALYs give an estimate of the total benefit of a health intervention to the patient in terms of added life years. However, in New Zealand’s case, comprehensive sets of QALY information are lacking (Douglas, 2006).

While some studies prefer market prices or purchaser valuations as the most appropriate means of valuing outputs, many (for example, Douglas, 2006; Atkinson, 2005; Schreyer, 2010) note it is difficult in practice to consistently use these approaches, and instead advocate for the use of cost weights.

## Valuations at producer’s unit costs of production

This method applies weightings to different outputs depending on the cost of providing that output. In doing so, the per-unit production costs act as a proxy for the per-unit value to the service recipient. If a state sector agency has three core outputs (F, G, and H), the calculation would be defined as:

$$(\text{units of F} * \text{unit costs for F}) + (\text{units of G} * \text{unit costs for G}) + (\text{units of H} * \text{unit costs for H}). \text{ (Dunleavy, 2016)}$$

Box 3.1 sets out a worked example of the application of cost weightings for an agency with three outputs.

### Box 3.1 An example of cost weighting outputs

#### Volume of outputs

	Year 1	Year 2	Year 3	Year 4	Year 5
Output A	3 800	4 500	5 000	6 000	7 000



Output B	2 000	2 200	1 900	2 100	2 150
Output C	6 000	6 000	6 200	6 300	6 400

### Unit cost of outputs

	Year 1	Year 2	Year 3	Year 4	Year 5
Output A	\$10	\$12	\$15	\$16	\$17
Output B	\$19	\$25	\$22	\$23	\$20
Output C	\$50	\$55	\$53	\$60	\$62

### Expenditure (output volume \* unit cost)

	Year 1	Year 2	Year 3	Year 4	Year 5
Output A	\$38 000	\$54 000	\$75 000	\$96 000	\$119 000
Output B	\$38 000	\$55 000	\$41 800	\$48 300	\$43 000
Output C	\$304 000	\$330 000	\$328 600	\$378 000	\$396 800
Total	\$380 000	\$439 000	\$445 400	\$522 300	\$558 800

### Expenditure shares for the base year of the series:

- Output A = 10%
- Output B = 10%
- Output C = 80%

### Cost-weighted output (volume \* expenditure share in the base year)

	Year 1	Year 2	Year 3	Year 4	Year 5
Output A	38 000	45 000	50 000	60 000	70 000
Output B	20 000	22 000	19 000	21 000	21 500
Output C	480 000	480 000	496 000	504 000	512 000
Total	538 000	547 000	565 000	585 000	603 500

### Index of cost weighted total output

	Year 1	Year 2	Year 3	Year 4	Year 5
Total	100	101.7	105	108.7	112.2

In contrast to the use of market prices or purchaser valuations (which attempt to attribute societal value to public services), cost weights reflect willingness of government to pay for certain goods or services. This simplifies the process of deriving weights, particularly where organisations face multiple competing objectives. Smith and Street (2005, p. 408) note “ultimately the selection of objectives in the public services is a job for the politicians who are charged with reconciling conflicting claims on public resources”. Similarly, Schreyer (2010, p. 11) notes that “the purpose of output measurement is not to provide estimates of societal value, so the use of cost weights does not constitute a major drawback”.

Box 3.2 provides examples of how cost-weightings have been applied in various measures of state sector productivity.

### Box 3.2 Examples of cost weightings in public sector productivity measures

- In their study of the evolution of productivity of the UK customs service Dunleavy and Carrera (2013) measured two key outputs: the total number of import and export declarations processed per year. These declarations were weighted by the relative unit costs in each year to create a total outputs data series.
- The UK Office of National Statistics (2012) calculates output of the education sector using the number of students in nine different learning services, including schools, the higher education training of teachers and health professionals, and further education. Student numbers are weighted by their share in aggregate education expenditure and converted into a single volume of education output series.
- Administrative costs can be a useful proxy where reliable data on per-unit costs are not available. In their study of productivity in UK tax administration, Dunleavy and Carrera (2013) used the share of administration costs for different taxes collected to weight the different tax volumes. Weighted tax volumes were added together to create a total index of tax output.
- In an analysis of police productivity in England and Wales, Pritchard (2003) applied weightings to categories of crimes investigated based on the costs involved. Their results show that although the total number of recorded crimes reduced between 1995 and 2001, the weighted output of investigations actually increased. This was due to a sharp increase in violent crime (which is the costliest type of crime to investigate) and a reduction in several types of crime that are less expensive (such as thefts from vehicles and burglaries).

## 3.6 Conclusion

Productivity measures seek to measure how efficiently outputs (the final goods and services produced by the entity under examination for external users) are produced. What constitutes an output will differ depending on the agency under examination, but importantly, outputs should be controllable by the agency in question. By contrast, outcomes (the changes in society, the economy, or the environment that an agency aims to effect) are typically influenced by factors outside of the agency's control, and are not part of productivity measures.

Some of the outputs produced by the state sector are relatively straightforward to measure – these include “transactional” services that are standardised, high-volume, and delivered with little interaction with the service recipient. Measuring outputs for other state sector services can be more challenging due to factors such as joint production, co-production, and collective consumption. However, these issues, along with data availability issues, are not insurmountable.

Where a state sector entity produces more than one output, the outputs need to be weighted so they can be aggregated into a single measure. This can be done using several different methods – however, it is widely recommended that outputs are weighted based on their volume and their cost of production.

After the measurement and weighting of outputs, the next step in developing a productivity measure is to record the inputs involved in their production. Guidance relating to this step is set out in the next chapter.

# 4 Selecting and costing inputs

## Key points

- It is critical to have transparency about the costs (inputs) involved in producing outputs.
- Inputs fall into three broad categories – labour inputs, capital services, and intermediate goods and services.
- There are two main ways to calculate inputs used in the production of outputs – volume measures and expenditure based measures. In some circumstances, volume and expenditure data may need to be combined to provide more accurate input measures.
- Attribution of inputs to outputs can become more complex as the scale and complexity of outputs increase. Some agencies have developed systems that manage these challenges. A detailed understanding of agency business models is important for accurately attributing inputs to outputs in large and complex organisations.
- Data limitations, missing inputs and costs, and the co-financing of some public services can complicate the process of identifying and attributing inputs to outputs. Better data collection systems, “second best” approaches to input calculation, and quality-adjustment of outputs can help overcome these challenges.

## 4.1 Introduction

This chapter considers issues in selecting and costing inputs as part of the construction of productivity measures. The first part of this chapter covers the identification of inputs and how these relate to the outputs produced by an agency as part of a production (or service provision) process. The remainder of the chapter covers the issues of co-financing of outputs by private individuals as part of the production or service provision process, and treatment of missing inputs and costs.

## 4.2 Types of inputs and attribution

It is important to have accurate and transparent information about agency costs (inputs) if productivity is to be measured correctly. Inputs represent the direct and indirect factors involved in the production of outputs – the final goods and services that are produced by an organisation.

There are three broad categories of inputs:

- **Labour:** people involved in the production of outputs, both directly (eg, teachers for school outputs) and indirectly (eg, administrative staff, who contribute to the functioning of an agency and the production of multiple outputs).
- **Capital:** the use/consumption of capital in the production of the output (eg, buildings, vehicles, information technology). This may need to be estimated through proxies (eg, depreciation, which measures the reduction in the value of the capital stock, and the capital charge which reflects the opportunity cost).
- **Intermediate goods and services:** other goods or services consumed as inputs in the production of the output.

Table 4.1 extends the table from Chapter 3 on the types of outputs to include examples of the different types of inputs that could be attributed to them.

**Table 4.1 Outputs and possible inputs in the state sector**

	Outputs	Inputs that contribute to production
<b>Hospitals</b>	<p>Hospital discharges for different diagnosis related groups.</p> <p>Number of courses of treatment for specific medical conditions.</p>	<p><b>Labour:</b> doctors, nurses, and other staff directly associated with producing the output (eg, ward clerks) and indirectly associated with its production (eg, laundry staff).</p> <p><b>Intermediate goods and services:</b> products and materials used as part of procedures or treatments (eg, bandages, scalpels, medicines).</p> <p><b>Capital:</b> contribution to capital costs related to the hospital (eg, internal charging for space and overheads).</p>
<b>Schools</b>	<p>Number of student places provided.</p>	<p><b>Labour:</b> teachers, principals, teacher aides, administrative staff.</p> <p><b>Intermediate goods and services:</b> teaching materials.</p> <p><b>Capital:</b> depreciation, capital charge.</p>
<b>Courts</b>	<p>Number of cases heard, hearings held, mediation sessions conducted and fines collected.</p>	<p><b>Labour:</b> judges, adjudicators, stenographers, court security staff.</p> <p><b>Intermediate goods and services:</b> law books, software, electricity.</p> <p><b>Capital:</b> capital costs relating to the operation of courts (eg, depreciation and capital charge).</p>
<b>Work and income services</b>	<p>Number of individuals that move off support and into sustainable employment.</p> <p>Number of applications processed in a given period.</p> <p>Number of young people moved off benefits and into education, training, or work-place learning.</p> <p>Number of people accessing emergency housing.</p>	<p><b>Labour:</b> case management and administrative staff.</p> <p><b>Intermediate goods and services:</b> cost of training courses for clients.</p> <p><b>Capital:</b> internal charging for space; or depreciation and capital charge.</p>

Identifying the different input types is important where partial factor productivity measures are being developed (ie, labour productivity and capital productivity<sup>10</sup>) or where questions are being asked about the impact of particular changes in input mixes (eg, hiring more or better-qualified staff, or buying new technology). Where it is not possible to differentiate input costs into labour, capital or intermediate categories, multi-factor productivity (MFP) analysis can be used. MFP analysis measures the amount of output produced from each unit of resource employed. All resources (ie, labour, capital, and intermediate goods) are used for the analysis, with total expenditure frequently used as the input. Examples of TFP analysis can be seen in the Commission's case studies on developing productivity measures of early childhood and school education (Gemmell, Nolan & Scobie, 2017b).

<sup>10</sup> Labour productivity measures the amount of output produced from each unit of labour employed, while capital productivity measures the amount of output produced from each unit of capital employed.

Identifying, calculating and allocating inputs depends on having a good understanding of the “business model” or processes by which outputs are produced. At the level of government agencies (or “sector”), inputs should already be allocated and costed into outputs as a requirement of the Public Finance Act. However, measuring productivity at a service or micro level may require the identification or creation of input data.

### 4.3 How to calculate inputs?

There are two main ways to calculate the inputs used in the production of outputs – volume and expenditure. Volume measures changes in the *number* of inputs (eg, staff numbers, hours, or full-time equivalents for labour); expenditure measures changes in *spending* on the different types of inputs. Volume is generally preferred by statisticians (eg, Atkinson, 2005), on the grounds that it more directly captures changes in the inputs. In practice, volume and expenditure may need to be combined to provide more accurate input measures.

#### Labour

For labour inputs, Statistics New Zealand lists the following sources of data in ascending order of their ability to provide an accurate, consistent, and continuous time series (Table 4.2).

**Table 4.2 Options for calculating labour input**

Labour input measure	Issues
Employment count (ie, number of workers).	Most straightforward to collect, but gives all workers the same weight regardless of whether they are full or part-time. Will not capture changes in input mix (and hence productivity) arising from changes in the full-time/part-time mix.
Full-time equivalent.	Takes into account the mix of full and part-time employment. However, often requires assumptions to be made about the relative contribution of each (eg, part-time workers are often given a weighting of 0.5). This may not reflect actual labour contributions.
Hours paid.	Does not require assumptions to be made about the relative input of full to part-time workers. However, may not fully capture changes in actual labour inputs as workers are often paid for a set number of hours, but change the number of hours worked each week.
Hours of work / actual hours.	More accurate than hours paid, but treats hours worked by all individuals as equal, regardless of their “quality”/skill level/seniority.
“Composition-adjusted” hours worked (ie, hours worked weighted by skill level. Higher skill labour units are given greater weight).	The most representative measure of labour volume, as it explicitly recognises differences between workers. Allows changes in labour composition that affect output to be reflected as change in labour contribution, and not as a change in productivity.

Source: Statistics New Zealand, 2014.

In many instances, available data may not be sufficient to develop the most sophisticated input measures. In such circumstances, “second best” approaches can suffice. For example, in the Commission’s case study on measuring productivity in the early childhood education sector, the available data on teaching staff was largely an employment count, broken down by full and part-time status and qualification status. To establish a more useful labour input, the Commission:

- weighted the part-time headcount numbers to distinguish their contribution to output production more clearly (using a range of weights, from 0.25 to 0.75); and
- used sector collective contract rates and Statistics New Zealand average income data for early childhood teachers to weight the qualified headcount numbers. This allowed the labour input measure to reflect changes in the quality/composition of the teaching cohort (Green, 2017).

The calculation of labour inputs also needs to account for the indirect costs associated with the production of goods or services, such as labour provided through administrative or support services. Agency finance divisions will have accounting rules for calculating or attributing the overhead contributions to outputs. Alternatively, if information on overhead costs needs to be collected directly, techniques such as time recording or resource accounting can be employed. The Treasury has guidance on these approaches (see, for example, New Zealand Treasury, 1994).

For some services, a significant proportion of the “labour” contribution to the production of an output may be carried out through outsourced or contracted staff. Therefore, these volumes or expenses will feature in the “intermediate goods and services” category. If the aim of measurement is to assess how productive the labour used in the delivery of services has been, excluding outsourced staff from the calculation will create misleading results. As Dunleavy & Carrera (2013) note, the:

interpretation of labour productivity analyses in the public sector always needs to be rather carefully carried out at an organizational level. TFP [total factor productivity] measures (including all forms of input costs) are generally preferable. (p.40)

## Capital

“Capital” refers to the fixed assets owned by the service provider used in the production of outputs, such as buildings, computers and infrastructure.

The measurement of capital can be difficult, as the capital used in the production of a service output is not generally observable. Therefore it must be approximated (Atkinson, 2005, p.49). Two approaches have been suggested in the literature. Dunleavy and Carrera (2013) suggest that “a good proxy of capital consumption is capital depreciation, which is published in most public organizations’ annual reports” (p.43). Depreciation measures reductions in the value of the capital stock over the useful life of assets. Depreciation rates often vary, depending on the asset in question and its expected useful life.

On the other hand, Atkinson (2005) argues that:

Capital consumption refers to the depreciation of fixed capital, not to the opportunity cost of the capital being employed in the public sector, rather than in another, use. For any given type of asset, there is a flow of productive services from the cumulative stock of past investments. To illustrate, take the example of an office building. Service flows of an office building are the protection against rain, the comfort and storage services that the building provides to personnel during a given period.

We recommend that the appropriate measure of capital input for production and productivity analysis is the flow of capital services of an asset type. This involves adding to the capital consumption an interest charge, with an agreed interest rate, on the entire owned capital. It would be this total cost of the capital which would influence the leasing charge or the user charge under the Private Finance Initiative (PFI) where these alternative routes to securing capital services for the public sector were employed. Conceptually, capital services reflect a quantity, or physical concept, not to be confused with the value, or price concept of capital. (p 49)

In general, measurement of productivity in the public sector should, as far as possible, mirror the approaches used in the measured sector. It follows from this that both depreciation and an interest charge should be included in the calculation of capital services. Leasing charges in the private sector incorporate a rate of return on the investment (including a risk premium for holding the asset), and an additional margin to cover maintenance and depreciation costs. Incorporation of an interest charge and depreciation in the calculation of capital services estimates the leasing charge that would be observed in the private sector.

In the New Zealand context, the equivalent for this interest rate would be the capital charge applied to departmental assets. The capital charge is “an expense derived from the capital cost of the Crown’s investment in each department.” (New Zealand Treasury, 1996, p. 42). It is designed to ensure that prices for government services reflect full production costs, allow comparisons of production costs with the private sector, and create incentives for departments to dispose of surplus fixed assets (ibid).

Where productivity analysis is being conducted at a service level, the capital charge or interest payment would need to be apportioned in relation to the use of the asset in producing the output (eg, on a per metre basis for space used). Internal charging regimes used by public service entities, such as hospitals, may

already incorporate these costs. For agencies that are not subject to the capital charge, the current Treasury discount rates should be used.<sup>11</sup>

#### F4.1

The capital charge (or equivalent Treasury discount rate) is a reasonable proxy for inclusion in the calculation of capital inputs.

One issue which can arise with calculations of capital productivity is the lag between investment and utilisation. There can be several years between an investment in new technology or infrastructure and any impact on the production of outputs. As a result, measured productivity in the short run may fall (as the capital input expenses rise without an offsetting increase in outputs). Having a sufficiently long data time-series is important for assessing the full impact of capital investments on productivity. However, the absence of a productivity impact from capital investment may not necessarily indicate mismeasurement; as Gordon et al (2015) comment, “if the new capital investments fail to be fully utilised...productivity will remain below the potential” (p.18).

### Intermediate goods and services

Intermediate goods and services refer to inputs procured and consumed in the production of outputs; this includes outsourced or contracted staff, repairs and maintenance, and charges for assets that are rented or leased by an organisation. Intermediate goods and services can be further disaggregated into sub-categories for more granular levels of analysis. The KLEMS framework, for example, breaks down intermediate goods and services into three categories – energy, materials and services – along with capital (K) and labour (L).

Issues can also arise in establishing the boundary between maintenance and repair of an asset, and costs associated with improving the asset. United Nations et al (2008) recommend including maintenance and repairs in intermediate good and service input measures be limited to activities that:

- an owner of an asset must take periodically over the life of an asset, and will otherwise drastically shorten the assets expected service life if they are not undertaken; and
- do not change the asset’s performance, such as replacing defective parts with new parts of the same kind.

In contrast, “major renovations, reconstructions or enlargements that go considerably beyond what is required simply to keep the fixed assets in good working order” adds to the stock of fixed assets (United Nations et al, 2008). Costs associated with renovation and changes in capital stock should be reflected in the interest charges and depreciation in the calculation of capital services.

The nature of a lease can have an impact on the ownership of an asset, and how it should be categorised as an input. Operating leases involve rental of buildings or other assets, whereas financing leases are a form of asset financing that substantially transfer the risks and rewards of owning the asset (OECD, 2009). There may be different treatments of operational and financing leases, and the distinction is often not clear-cut. To address this, the OECD suggests:

for analytical purposes such as productivity measurement it may be useful to classify all assets according to the kind of activity of the user of the asset, without attempting to distinguish between operational leases on the one hand and financial leases and rents of land and sub-soil assets on the other hand. This is certainly a pragmatic way forward (p. 174).

## 4.4 Systems to capture and allocate inputs

Some inputs are easily understood, such as labour costs that are part of an output. For small decision-making units or services, only a few people may be involved in production and the input can hence be easily calculated. Labour costs are attributed to periods of months and years, often as a matter of course within

<sup>11</sup> Note that Treasury discount rates typically vary, depending on the nature of asset in question. For example, the rate applied to general purpose office and accommodation buildings has historically been lower than the default rate or the rate applied to IT equipment.

standard budgeting, contract, or financial management. This can enable simple attribution of labour costs to activities and outputs.

As the scale and complexity of outputs increases, or as staff are expected to contribute to multiple outputs, attribution issues can become challenging. However, a number of agencies have already developed systems to manage this challenge. Two examples are described below. Both models identify a significant amount of input data and allocate it to outputs, based on administrative processes used for daily business. Both systems also illustrate the importance of having a detailed understanding of agency business models when productivity measurement is being conducted in large or complex organisations.

### MSD's individual Cost Allocation Model (iCAM)

The Ministry of Social Development's (MSD) individual Cost Allocation Model (iCAM) was designed to help understand trends in the efficiency and cost-effectiveness of MSD interventions and services, and adequacy of targeting to different groups. iCAM uses activity information from MSD's administrative datasets to allocate expenditure on inputs to the various outputs of its Service Delivery group; this is the group responsible for assessing and processing entitlements and assistance (eg, income support, student loans and allowances, New Zealand superannuation, employment and training programmes).

The cost allocation model is based around four main principles:

- **Include all financial costs for Service Delivery:** these are the departmental costs of delivery, rather than the benefit payments processed by the group.
- **Reconcile allocated expenditure to financial totals:** for each appropriation, iCAM reconciles allocated expenditure to the appropriation amount in each financial year.
- **Disaggregate costs down to the individual output level:** costs are broken down to the individual level.
- **Apply the same approach over successive financial years.**

The iCAM model identifies the activities ("components") that contribute to each of Service Delivery's outputs (eg, booking a client to attend a seminar is a component in delivering an employment seminar output). Staff time data is collected automatically through MSD computer systems as Service Delivery employees carry out their activities, thereby avoiding the need to use time-sheets. Quantitative information (metrics) is then used to allocate costs to each component, based on a scale. Total output costs depend on the sum of the costs of each component. Examples of the Service Delivery cost components, and their metrics, are outlined below (Table 4.3).

**Table 4.3 iCAM Service Delivery components**

Component	Definition	Metric
Appointment	Scheduling an appointment with a client.	Staff time
Benefit administration	Assessing and maintaining entitlement to income support assistance.	Staff time
Benefit payments	Bank fees for payment of income support benefits	Pay weeks
Client contact	Contact with clients to help them plan and move into employment or updating their records.	Staff time
Contract administration	Administration of contracts including tendering, negotiation, payment, and managing the performance of contracted providers.	Contract amount
Placement Opportunity	Time spent by contact centre staff and work brokers to identify and establish vacancies with employers.	Starts
Seminar	Staff time in administering and running seminars.	Staff time



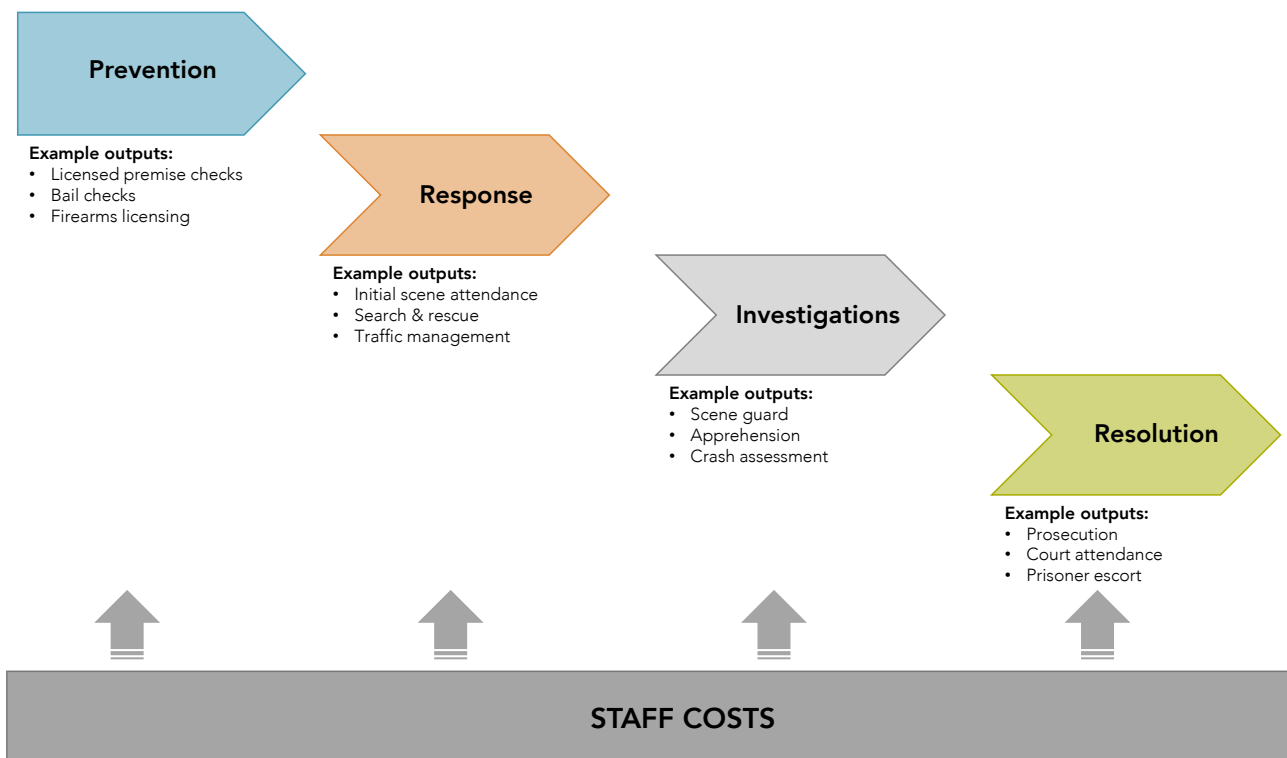
Component	Definition	Metric
Indirect costs	IT, corporate services, property, and support staff costs.	Departmental cost of each output

Source: Ministry of Social Development

## New Zealand Police's "Know Your Business" model

New Zealand Police's input allocation system is based off a business model for the entire agency, which was designed in part to enable police leadership to understand how its activities and outputs related to each other, and communicate these to external stakeholders. "Know Your Business" groups the agency's 40 outputs into four main categories, reflecting the flow of police work (Figure 4.1). Because front-line staff may carry out multiple outputs in the same day, cost attribution is a potentially significant undertaking.

**Figure 4.1** NZ Police "Know Your Business" model



Like MSD's iCAM, the New Zealand Police system collects and allocates a significant proportion of labour input information to outputs through existing administrative processes. Staff hour information is derived from the police central dispatch system, which tasks police officers. This records with a relatively high level of accuracy the time a police officer responds to a certain incident, and the time they report completing the incident. Where dispatch information is not available, estimates of staff time are used for cost allocation.

The Know Your Business model "provides a consistent view of how much effort is allocated to activities, and a framework for discussion and decisions about future allocations" (NZ Police, 2016, p. 46). It can also be used to analyse specific areas of police activity (eg, family violence), including assessing trends in performance and calculating the cost of service delivery.

## 4.5 Co-financing, missing inputs and costs

### Co-financing

Many public services are provided free of charge. But a large number also have co-payments. These are usually monetary, but are sometimes donations of labour (eg, volunteers in the Department of Conservation or parents volunteering in schools). Where there is a financial charge, this can be compulsory (eg, in the case of most tertiary education) or voluntary (as in schools). It can be a universal charge, or it may be targeted on

the basis of means. In primary healthcare, co-payments can vary significantly within and across providers, or for care provided outside business hours.

If productivity measurement does not account for this co-financing:

- A government agency could appear more productive than it is (ie, the cost of producing its outputs will be artificially low and its productivity will appear high).
- It may create perverse incentives. Agencies could improve apparent productivity by cost-shifting to the public (eg, by increasing the proportion of costs covered by co-payments) without any real improvement in efficiency.

One approach in the UK has been to effectively split outputs directly in proportion to the share of inputs. Where a service is 70% government-funded, then 70% of the outputs are attributed to government (Statistics New Zealand, 2010). But this approach relies on having good information about the share of private financing.

Another approach is to include the contribution made by private individuals to the production of a product or service, essentially treating it as jointly produced. This may be particularly important when the inputs are not financial, such as direct parent involvement and time contributed to schools, or the contribution of neighbourhood watch and other voluntary patrols to crime prevention. For “in kind” contributions, such as volunteer time provided to an early childhood centre, estimates could be developed based on the hours of voluntary labour provided and average wage rates.

## Accounting for gaps in input information

The inability to fully measure, or the exclusion of, some inputs may lead to inaccurate productivity results. For example, the Australian Productivity Commission has concluded that the failure to take account of reductions in the quality of the natural resources as inputs in calculations of mining productivity led to unduly low multi-factor productivity results for the industry (Box 4.1).

### Box 4.1 Accounting for changes in the quality of inputs in the mining sector

In their analysis of Australian mining sector productivity, Topp et al discuss the role that declines in the quality of inputs (natural resources like ore, in the case of mining) can have on productivity measurement:

While natural resources are obviously a major input into mining production, changes in their quality are not generally taken into account in standard measures of productivity. This omission would not be a problem if natural resources were in infinite supply and of homogenous quality – that is, available without constraint at the same unit cost of extraction. But neither is the case: resource deposits are non-renewable, and depleted by ongoing extraction. And as mineral and energy deposits are depleted, the quality and accessibility of remaining reserves generally decline...As the quality and accessibility of deposits decline, greater commitments of capital and labour are generally needed to extract them. When deposits are deeper, more development work is needed to access the desired resources. If there are greater impurities, greater costs may be incurred in extracting and processing the material into saleable output. In short, more ‘effort’ is needed to produce a unit of output. The additional capital and labour per unit of output show up as a decline of productivity. (2008, pp. xvii-xviii)

To take account of changes in the quality of inputs, Topp et al constructed an index of mining “yield”, using data such as average ore grades in metal mining, the ratio of saleable to raw coal in coal mining and “the implicit flow-rate of oil and gas fields in the petroleum sector” (2008, p.xviii). Over the period studied by Topp et al (1974/75 to 2006/7), the average yield fell substantially, reflecting the declining quality and increasing depletion of natural resource inputs. Once this effect is removed from the calculations, “mining MFP [multifactor productivity] grows at any average rate of 2.5 per cent per year, compared with 0.01 per cent per year in conventionally measured mining MFP” (ibid).

Source: Topp et al, 2008

In the context of public services, two potential missing inputs or costs are worth keeping in mind – technological improvements and the characteristics or behaviours of people receiving the services.

### Technological improvement

Technological progress is a key driver of productivity improvement and economic growth, yet can sometimes be difficult to discern in data. This is not an issue exclusive to public services. Problems reflecting technological improvements in the market sector have led to debates as to whether national statistics, and the recorded slowing in labour and multi-factor productivity in developed countries, are in fact under or mis-measuring the benefits from information technology (Syverson, 2016; Byrne et al 2016).

Difficulties capturing technological improvements at the level of a service can be seen in the case of computers and the internet:

- While the price of computers has fallen dramatically in recent years (down 69.2% between the second quarters of 2007 and 2017, according to Statistics New Zealand), the processing power has grown significantly. Gordon et al (2015) note that in 2009 “average memory of computers on the market was about 2 gigabytes, four years later it was 8 gigabytes.” (p.17)
- The average price for telecommunication service fell by 60% between June 2012 and June 2016, while the number of broadband connections over the same period grew by 12.8%. Furthermore, the share of broadband connections with no cap increased from 5% in June 2012 to 49% in 2016 (Statistics New Zealand, 2016; n.d.).

In brief, people are getting more processing power and connectivity for lower prices. This can be particularly important for knowledge-intensive agencies and services. Simply measuring the expenditure on these intermediate goods and services would not pick up these improvements to input quality and volumes. Some of these issues can be resolved through “quality adjustment”, which is discussed in more detail in Chapter 5.

### Inherent characteristics of clients / co-production

One of the challenges of measuring productivity in services is taking into account pre-existing attributes or the contributions of clients to service production. As noted in Chapter 3, one often-cited example is the knowledge and competencies a young person brings to school. The learning the student gains from school will be the combined result of the teaching services received, and their own effort and talents. Failure to take this into account can lead to overly-positive productivity estimates. Another example is the pre-existing conditions that a patient brings to a medical treatment, which may affect the success of any subsequent intervention. This may have positive or negative impacts on productivity results.

Technically, both the pre-existing competencies of school children and conditions of medical patients can be thought of as inputs to the production of outputs and could theoretically be included in the input calculation. In practice it has proven easier to deal with these issues by quality adjusting the outputs through such factors as “casemix” for health, and value-add for education (eg, progress over the course of the year against the curriculum or standards).

#### F4.2

Challenges measuring the impact of technological improvement on inputs, or the contribution of client characteristics to service production, can often be managed through the quality adjustment of outputs.

## 4.6 Summary

Inputs are conceptually simpler to categorise than outputs. However, data limitations, missing inputs and costs, and the co-financing of some public services can complicate the process of identifying and attributing inputs to outputs. Better data collection systems, “second best” approaches to input calculation, and quality-adjustment of outputs can help overcome these challenges.

## 5 Adjusting for operating environments, quality and price

### Key points

- A raw measure of productivity – the ratio of inputs to outputs – is not particularly useful by itself: it is only meaningful as part of a comparison. In making comparisons it is important to account for differences that organisations (or parts of organisations) may face in their operating environments or how operating environments may change over time.
- There is a sizeable literature on addressing differences or changes in the operating environment. Approaches can include limiting the range of providers studied to those from similar environments or measuring the outputs related to different population subgroups separately (treating them as different outputs).
- The quality of services may also differ between organisations or change over time. There are three general approaches that can be taken to accounting for quality:
  - *implicit adjustment*: this groups outputs so that only products and services that have similar specifications are compared;
  - *explicit adjustment based on proximate outcomes*: this adjusts outputs by the resulting change in status directly attributable to the services; and
  - *explicit adjustment based on final outcomes*: this adjusts outputs by a state that consumers value, such as health status, without necessarily relating the change to a policy intervention.
- As productivity is a volume measure it is necessary to remove any effect that changes in prices may have on comparisons over time. The two main methods for developing price deflators are to use publicly available market price information (such as the Consumer Price Index) or through calculating direct volume indexes.

### 5.1 Introduction

A raw measure of productivity – the ratio of inputs to outputs – is not particularly useful by itself: it is only meaningful as part of a comparison (Statistics New Zealand (2010) in Productivity Commission (2017, p. 5)). These comparisons can be made over the productivity levels or growth rates of different organisations (or parts of organisations). They can also cover the productivity growth rate of a particular organisation (or part of an organisation) over time. This chapter discusses ways in which data on the volumes of inputs and outputs can be adjusted for differences in operating environments and for quality and price changes so that useful comparisons can be made.

### 5.2 Why adjustments for operating environments, quality and price matter

In making comparisons it is important to account for differences that different organisations face in their operating environments and for any changes in operating environments over time. The quality of services may also differ between organisations or change over time. And, as productivity is a measure of the volume of output over the volume of inputs, it is necessary to remove any effect that price changes may have on comparisons. Examples of how different operating environments, or changes in quality or prices, can affect productivity measurement are outlined below

- *Differences in operating environment:* Differences in the performance of schools may, for example, reflect the socio-economic status of their students as well as the performance of their staff. Failing to account for these differences in operating environments could mean measures overstate the performance of schools that largely draw students from advantaged backgrounds.
- *Changes in quality over time:* Suppose that the number of patients treated in a public hospital grew at a slower rate than labour and capital inputs. Measuring productivity on this basis would tell a story of falling productivity. But also suppose that the quality of care increased and readmission rates fell. In this case the change in measured productivity would be missing an important part of the story.
- *Effect of price changes:* Consider a service where the direct volume of output cannot be measured and where price changes have meant that over time it has become cheaper to provide the service. In this case it is possible that expenditure on the service could remain the same but for the volume of output to increase over time. Failing to “deflate” expenditure (eg, remove the effect of price changes) would understate any productivity improvements.

The following sections discuss approaches for accounting for these differences and changes in productivity measures.

### 5.3 Differences in operating environments

Differences that organisations face in their operating environments can be seen in the example of two hospitals that produce the same number of operations for the same quantity of inputs. These hospitals may appear to have equal productivity but if one is treating patients with more complex conditions then the value it is adding is higher. It is therefore important to account for differences in the complexity of activity in measuring the output of state services such as hospital care (Box 5.1).

#### Box 5.1 The changing operating environment for hospital outputs

Fraser and Nolan (2017) investigated changes in two aspects of hospitals’ operating environment:

- the age profile of the population; and
- the complexity of treatments (reflected in average length of stay and cost weights).

Understanding these changes in the environment is important for illustrating structural changes in the health system (eg, ageing of hospital patients, burden of chronic conditions, and moves to more preventative and community based care) and provides context for productivity measures.

**Table 5.1 Age of Inpatient Events, 2002-2014**

Year	Mean	Median	Year	Mean	Median
2002	42.41	40.77	2009	44.05	44.45
2003	42.57	41.17	2010	44.42	45.14
2004	43.18	42.37	2011	45.05	46.34
2005	43.19	42.66	2012	45.30	46.90
2006	43.55	43.39	2013	45.79	47.80
2007	43.55	43.43	2014	45.81	47.82
2008	43.75	43.83			

Source: Fraser and Nolan (forthcoming), *Understanding health sector productivity*.

Note: Age is calculated as days between date of birth and date of admittance divided by 365.2

Based on data for hospital inpatients (the National Minimum Dataset) for 2002 to 2014, the mean age of the inpatient population rose from 42.4 to 45.8, a trend of 0.28 years annually. The median age rose

from 40.8 to 47.8, a trend of 0.58 years annually. There was also some evidence of increasing complexity.

Although the data used provides only a partial view of the health system, the findings raise more general questions. For example, does the ageing of patient discharges tell us something about the “healthy aging” hypothesis or does it reflect the interaction between hospital care and other parts of the health system (such as primary and residential care)? To fully answer these questions more system-wide data are required.

Key differences in operating environments that can be useful to account for include:

- the characteristics of the clients of the services (eg, age, socio economic background, pre-existing status, support networks, etc.);
- the size and scope of the organisations (eg, whether hospitals have specialist units);
- market structure (eg, presence of other suppliers/competitors); and
- overall performance of the economy.

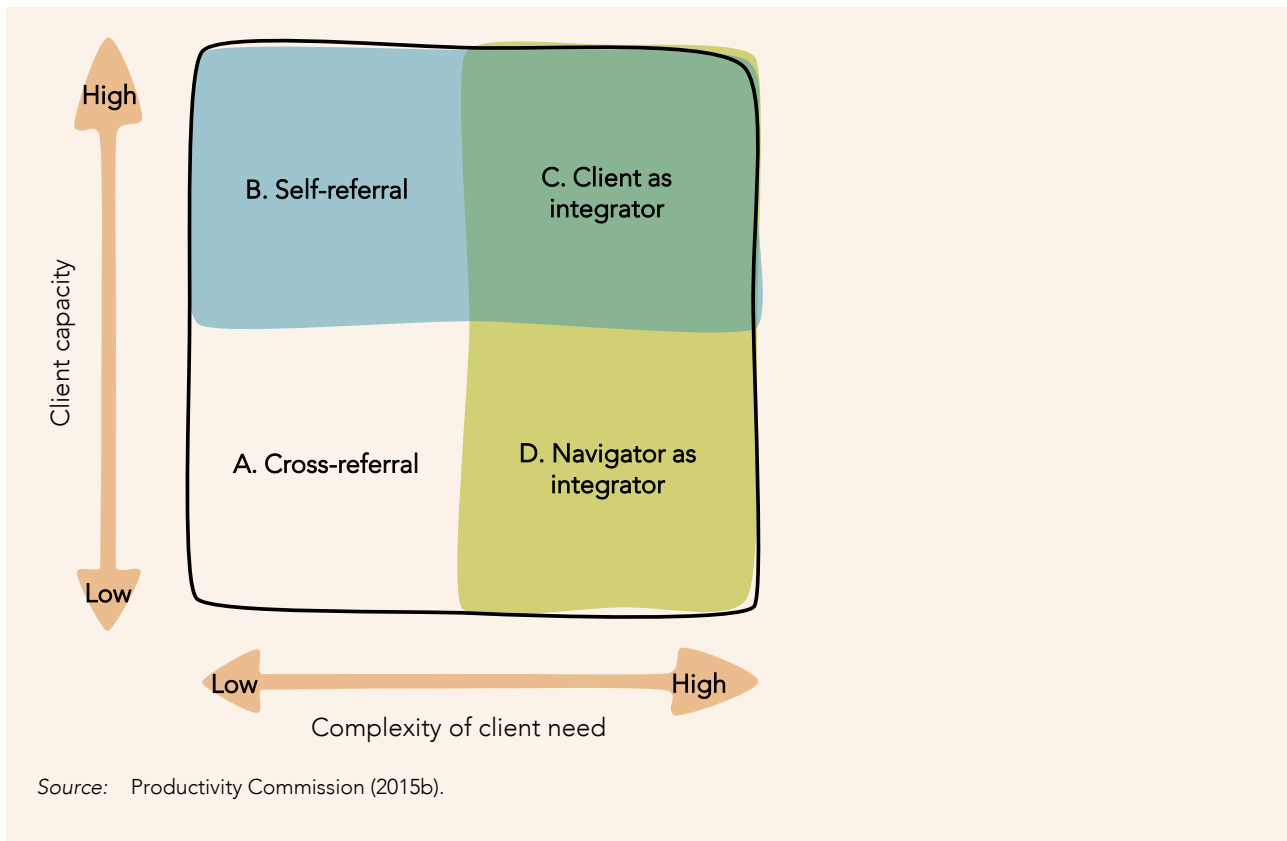
There are a number of approaches to accounting for these factors. One way is to measure the outputs related to different population subgroups separately (segmenting the population) and treat them as distinct outputs. Similarly, the range of providers studied could be limited to those from similar environments. An example of segmenting a population is how the Productivity Commission distinguished between users of social services depending on how they access the system and their reasons for doing so (Box 5.2). Under this approach the productivity of the services received by clients in the four quadrants would each be measured separately. However, studying different population groups separately (or limiting the range of providers studied) can limit the scope of the analysis, unless methods for combining the results for different quadrants are used (such as cost weighting).

#### **Box 5.2 Example of segmenting a population: the characteristics of people interacting with the social services system**

In *More effective social services* (2015a), the Productivity Commission highlighted how clients access the social services system in different ways and for different reasons. For some, their main interaction with the system is through their local school or childcare centre. On occasions, they may need to visit their local general practitioner or perhaps a hospital if the issue is more serious. For these people, coordinating services to meet their needs is relatively straightforward, and in many cases they prefer to coordinate their own interactions with the social services system.

The Commission thus segmented service users according to the complexity of their needs and their capacity to extract the services they need from the system. The Commission found it useful to group clients under four headings:

- people with relatively straightforward needs who require assistance to access services (quadrant A);
- people with relatively straightforward needs who have the capacity to access services for themselves (quadrant B);
- people with complex needs who have the capacity to access services for themselves (quadrant C); and
- people with complex needs who require assistance to access services (quadrant D).



Another approach is to adjust the volumes of outputs for differences in the operating environment. For example, in the health sector casemix adjustments, which account for the characteristics of patients, are used to allow for comparisons across settings or time. As Rouse and Swales (2006) noted, volumes can be adjusted for case severity, typically using cost weights. These weights “group together treatments that are clinically similar, consume similar quantities of resources and are likely to be similar in cost”. In principle, a type of casemix adjustment might be applied to productivity measures for other public services.

### F5.1

Productivity measures should account for differences or changes in an organisation’s operating environment. Key differences that it can be useful to account for include the characteristics of the clients of the services, the size and scope of the organisations, and market structure.

These differences can be reflected in measures by:

- measuring the outputs related to different population subgroups separately (segmenting the population) and treating them as distinct outputs;
- limiting the range of providers studied to those from similar environments; and
- adjusting the volumes of outputs for differences in the operating environment (eg, severity of treatments).

## 5.4 Quality adjustment

Quality can have many dimensions, as consumers may value a wide range of characteristics when consuming public services. Schreyer (2010) outlines three general approaches to adjusting for changes in the quality of the output of public services.

- *Implicit quality adjustment (stratification)*: This approach groups outputs so that only products and services of the same specification are compared over time or in space (Schreyer, 2012). For example, if there are two services that differ in quality, the changing consumption of these two goods could indicate

quality change. Yet this method only captures the changes in quality that are associated with compositional shifts (Sharpe, Bradley & Messinger, 2007). It fails to capture changes in the quality of individual goods or services (or activity packages). It requires an assumption that the quality of each of the stratified activities are visible and remain unchanged over time.

- *Explicit adjustments*: Explicit approaches to quality adjustment are based on measures that adjust outputs for changes in outcomes. There are two broad approaches to explicit quality adjustment.
  - *Explicit adjustment (proximate outcomes)*: The first approach is based on a resulting change in status directly attributable to the services received (Schreyer, 2012). In this case, quality adjustments could be based on factors like examination scores or attainment levels for education outputs (O'Mahony and Stevens, 2009) or the change in health status associated with an intervention (Schreyer, 2012).
  - *Explicit adjustment (ultimate outcomes)*: A second approach is based on a broader definition of outcomes as "a state that consumers value, for example the health status without necessarily relating the change in this state to the medical intervention" (Schreyer, 2012). In this case, quality indicators could include the population's education level, life expectancy, or level of crime. Proxies for these indicators have included future earnings as a measure of the underlying population education level or prices of houses as a measure of school quality (Black, 1998; Cannon et al., 2015; Gibson and Boe-Gibson, 2014).

Further, as Gemmell, Nolan and Scobie (2017a and 2017b) illustrated, it can also be possible to use similar techniques to quality adjust the inputs (eg, staffing) into public sector production. Different approaches are illustrated below with examples from the education and health sectors.

## Quality adjustments in education

In education, key quality adjustments can relate to inputs (eg, teacher quality or pupil-staff ratios), proximate outcomes (eg, performance in school inspections or student attainment) or final outcomes (eg, impact on human capital or house prices (due to school zoning)). Possible approaches to quality adjusting data on schools are shown in Table 5.2. For examples related to tertiary education see Gemmell, Nolan and Scobie (2017a).

**Table 5.2 Possible approaches to quality adjusting data on schools**

Concept	Variable(s)	Measure(s)	Key challenges
Labour inputs (resources used in production)	Labour	Labour force (employment count, FTEs, hours paid, actual hours worked, quality adjusted hours (eg, based on wage rates or qualifications))	Combining non-commensurate inputs into an index  Informal inputs (such as student attributes)
Total inputs (resources used in production)	Labour, Capital, Intermediate (eg, teaching aids, electricity usage and building maintenance)	Total real operating allowances	Implicitly assumes expenditure weights are appropriate
Proximate outcomes	Acquisition of skills and qualifications/ Transfer or increase in knowledge	Pupil based: pupil numbers (hours v. EFTS), educational attainment (milestones, credits, degrees)  Teaching based: eg, number of lessons, class size, school inspections	Combining non-commensurate outputs into an index  Attribution (eg, informal inputs)  Teacher quality  Grade inflation  Teaching to the test



Concept	Variable(s)	Measure(s)	Key challenges
Final outcomes (direct)	Human capital	Additional lifetime earnings	Lags and attribution to expected earnings
Final outcomes (indirect)	Social network	Housing value approach	Accounting for more general neighbourhood effects

Source: Based on Howell (2016) in Gemmell, Nolan, and Scobie (2017b)

The examples in the table illustrate a number of general considerations when quality adjusting productivity data. For instance, in relation to labour inputs, teacher quality has been identified as the primary influence on pupils' educational progress (eg, Hanushek & Rivkin, 2006). (See also the discussion in Chapter 3 on teachers and standardised assessments.) Yet measuring teacher quality can be complex and there are a variety of approaches that could be used. Qualifications could be used as a proxy for teacher quality and measured directly. Alternatively, labour inputs (eg, actual hours worked, hours paid or FTE number of employees) could be segmented into categories of employees (eg, teacher, teaching assistant). Where data on qualifications or the composition of the labour force is not available then proxies for quality (such as salary data) could be considered (Gemmell, Nolan & Scobie, 2017b).

Another potential indicator of quality are measures like class size (pupil-staff ratios). These adjustments are relatively simple to make and are used in a number of countries. However, Gemmell, Nolan and Scobie (2017b) did not support the use of pupil-staff ratios as a proxy for quality as the evidence on their impact was mixed. They cited a range of studies that drew different conclusions regarding class size and argued that class size is one of a number of components that impact on teaching quality and it is unclear at what number of students any negative effects of congestion start to take effect. Another approach would be to use school inspection reports, such as those by the Education Review Office. But inspections are largely qualitative and may be subject to personal biases of particular inspectors (Schreyer, 2010). The criteria for defining good or bad performance may also change over time and it is necessary to find some approach to weighting individual inspections so they are comparable over time and across schools.

Finally, Gemmell, Nolan and Scobie (2017b) identified a range of outcomes that could potentially be used as a proxy for the quality of schools. This included the use of:

- internationally comparable standardised test scores (Leigh and Ryan (2008), Hanushek et al., 2010);
- higher expected future earnings (Barslund & O'Mahony, 2012; Hanushek, 2010; Murray, 2007; O'Mahony & Stevens, 2009; O'Mahony et al., 2012); and
- differences in prices for equivalent houses belonging to different school zones (Black, 1998, Gibson & Boe-Gibson, 2014).

There are limitations to using outcomes for quality adjustment. Some studies use expected earnings as a quality adjustment for school or tertiary education. The method can be influenced by selection bias, where students enrolling in additional education are self-selecting (ie, they do not constitute a random sample). Secondly, the historical earnings profiles for different levels of qualification are typically used as the basis for future earnings, but the past is not always a good predictor of the future. Thirdly, any earnings premium is often attributed to education when some part of this may have been due to a person's innate abilities, their family backgrounds, health status, the influence of their peers, or employer provided training, and so on (Hanushek, 2015).

## Quality adjustments in health

Approaches to quality adjustment can also be illustrated with the case of healthcare. There is sizeable literature on applying quality improvement approaches to healthcare and researchers such as Professor Martin Connor of the Centre for Health Innovation, Griffith University, have illustrated the real potential of hospital performance data. As Martin Marshall (2009) noted in the *British Medical Journal*, statistical approaches first developed in the manufacturing sector by researchers like Deming could support quality

and reliability in healthcare. When it comes to defining quality, as with education, it is possible to think about inputs (eg, wage rates; the qualifications of clinicians), proximate outcomes (eg, variations in care; quality and safety markers) and final outcomes (eg, patient experience data; measures of whether people are being treated in the right setting). Some examples are summarised in Table 5.3 and are discussed below.

**Table 5.3 Possible approaches to adjusting health sector data**

Concept	Variable(s)	Measure(s)	Key challenges
Labour inputs (resources used in production)	Labour	Labour force (employment count, FTEs, hours paid, actual hours worked, quality adjusted hours (e.g., based on wage rates or qualifications))	Combining non-commensurate inputs into an index  Informal inputs (such as patient characteristics)
Total inputs (resources used in production)	Labour, Capital, Intermediate (eg, electricity usage and building maintenance)	Total real operating allowances	Implicitly assumes expenditure weights are appropriate
Proximate outcomes	Avoidance of direct harm  Avoidance of excessive variation	Quality and Safety Markers  Atlas of Healthcare Variation	Whether markers based on priority areas reflect system performance more generally  How to define appropriate variation
Final outcomes	Direct	Patient experience data	Currently limited to adult inpatients, but being extended to primary care
Final outcomes	Indirect	Changing population shares in levels of chronic care	Sampling bias and attribution issues

Source: Productivity Commission

In relation to proximate outcomes, the Health Quality and Safety Commission (HQSC) compiles an Atlas of Healthcare Variation in New Zealand (see, for example, Fraser & Nolan (forthcoming)). This concentrates on individual conditions and clinical groups and highlights differences in practice and the improvements required to reduce unwarranted variation. The HSQC also produce a series of Quality and Safety Markers. These markers are a mix of process and outcome measures focused on driving improvement for four key safety priorities: falls, healthcare associated infections, surgical harm, and medication safety. They set expected levels of improvement, publicly report progress against these thresholds, and support links to accountability mechanisms.

In relation to ultimate outcomes, one possible approach to quality adjustment would be to adjust complexity-weighted events with Patient Experience Data. In New Zealand the HQSC collect two sources of patient data:

- *Adult Inpatient Survey*: a 20-item adult inpatient survey in hospitals beginning in August 2014. The survey runs quarterly in all district health boards and covers four key domains of patient experience: communication, partnership, coordination and physical and emotional needs. A selection of adult patients who spent at least one night in hospital are sent an invitation via email, text or post inviting them to participate in the national survey. Responses are anonymous unless patients choose to provide their contact details.
- *Primary care patient experience survey*: an online patient survey being developed by the HQSC to find out what patients' experience in primary care is like and how their overall care is managed between their

general practice, diagnostic services, specialists, and/or hospital staff. A number of practices across New Zealand began using the survey in February 2016.

A further approach to quality adjusting for ultimate outcomes would be to segment the treatment population into different groups, such as supported self-care (the large majority of people with chronic conditions), disease management (high risk patients), and case management (highly complex patients). The proxy for quality could then be changes in the distribution of the population in these different groups over time. However, undertaking this analysis would be a relatively data intensive exercise and it would be necessary to address potential sampling bias issues and challenges in attributing changes in shares to health system performance.

## Quality adjustments in other areas

As Atkinson (2005, p. 143) notes, the output of prison services is often measured by numbers of nights spent in prison by: any prisoners on remand; prisoners under sentence; non-criminal prisoners; and prisoners in police cells. But this fails to quality adjust for overcrowding, reoffending, and achievements during incarceration such as educational attainment or drug rehabilitation. It also fails to weight according to cost – for example, high risk/low risk prisoners. Atkinson (2005) argued that overcrowded cells could be given a lower weight in output, although developing a precise weight requires robust evidence. As Atkinson (2005, p. 171) also noted, similar concerns can also be raised regarding the standard approaches to measuring output in social services (eg, admissions) and benefit administration (Table 5.4).

**Table 5.4 Possible approaches to adjusting data in other sectors**

Sector	Unadjusted Output	Limitations	Possible quality adjustment(s)
Prisons	Number of nights	Fails to reflect overcrowding, reoffending, rehabilitation, and prior risk	Proportion of cells that are overcrowded
Social services	Admission rates	Fails to reflect quality of life and independence of vulnerable older people	Proportion of older people living in their own homes
Benefit administration	Raw activity numbers	Fails to reflect whether recipients receive a high-quality service	Measures of timeliness and accuracy of payments

Source: Productivity Commission

## Pros and cons of different approaches to quality adjustment

The choice of how to account for quality changes can have a significant impact on estimates of productivity growth. Work in New Zealand and Australia has shown how sensitive estimates of productivity can be to the approach taken to control for the quality of inputs and outputs.

- Gemmell, Nolan and Scobie (2017b) tested a range of quality adjustments to productivity estimates for New Zealand schools based on sector level data. They found that although most adjustments provided a broadly (though not completely) consistent picture of flat or declining productivity, in one case the change of method led to the measured productivity trend reversing.
- Likewise, when the Australian Productivity Commission (Lovell & Baker, 2005) developed experimental productivity estimates for 10 government services drawing on data contained in the Report on Government Services they found that the estimates of productivity were sensitive to the approach taken to control for the quality of inputs and of outputs.

In the UK, the Office for National Statistics (ONS) takes a case-by-case approach to quality adjustment. Approaches used include stratification of services (implicit adjustment) and explicit adjustments (based on

the attributable contribution of the activity to outcomes). The ONS has noted the greater degree of subjectivity involved in quality adjustment compared to volume measures, which means there needs to be a higher standard for judging their use.

In their work, Gemmell, Nolan and Scobie (2017b) tested a range of approaches to quality adjusting productivity estimates. They highlighted the need for caution in making adjustments to labour inputs (such as wages as a proxy for labour quality) given the important caveats on the use of salaries as a proxy for quality of inputs. This reflects the nature of public service labour markets (eg, whether a change in salaries reflects quality or compositional changes) and the importance of missing inputs such as the previous performance of students (needed for measures of value added).

Likewise, they also argued that approaches based on final outcomes (such as the impact of the education system on earnings) raised attribution issues. They showed how the decline in measures based on ultimate outcomes was likely to reflect changes (at least partly) in unemployment and real wage growth following the Global Financial Crisis. Changes in these measures reflect differences in the economic context facing different cohorts of school leavers, not just the performance of schools.

Gemmell, Nolan and Scobie (2017b) argued that explicit adjustments should be based on the attributable contribution of the activity to outcomes (such as student performance in tests). This was similar to the conclusion reached by the ONS in the UK. However, even in this case there can be scope for ambiguity. For example, student performance can be measured against performance in domestic or international tests, and in recent years the performance of New Zealand students in domestic tests has contrasted markedly with their performance in international ones.

While their work pointed to challenges in quality adjusting state sector productivity measures, this was by no means supporting a case for only relying on unadjusted measures. Quality adjustment remains important but it is unlikely that a single measure would capture the full richness of the productivity story and a range of approaches may be needed. Further, different approaches will be relevant for different stakeholders and questions. Any approach will need review and should change over time as data availability and analytical techniques improve.

A final challenge in quality adjusting state sector productivity is the proportionality problem; this arises when comparing indexes of quality with indexes of inputs and outputs. It is necessary to establish a “factor of proportionality” between the change in quality scores and the change in output (Schreyer, 2010). Quality-adjusted output will not necessarily reveal how much extra quality is valued by the users of services. Should, for example, a 5% increase in student test scores mean the value of school output should be 5% higher? This consideration is especially important when comparing productivity outcomes across industries, for example, whether a 10% increase in the quality of education should be equivalent to a 10% increase in quality of healthcare.

## F5.2

Properly measured, productivity captures changes in both the volume and quality of outputs. Quality changes can be reflected in productivity measures by:

- grouping outputs so that only products and services of the same specification are compared;
- adjusting outputs based on the resulting change in status directly attributable to the services received; or
- adjusting outputs based on final outcomes, for example the population’s education level or life expectancy, without necessarily relating change to any intervention.

## 5.5 Accounting for price changes

There are two reasons to account for changes in prices when measuring state sector productivity.

The first is that productivity is a volume measure. This means it should compare how the volume of outputs changes in response to a change in the volume of inputs. Yet in some cases measures of outputs and inputs reflect both changes in prices as well as changes in volumes. For instance, a change in expenditure on staff would reflect both changes in the volume of the labour input (eg, hours of work or numbers of FTE employees) and changes in salaries. These “price changes” need to be removed.

A second, related, reason is that in some cases direct measures of the volume of outputs and inputs are unavailable. In these cases, it can be necessary to deflate measures of expenditure to estimate the volume of outputs (Atkinson, 2005, p. 71). However, the approach to deflating expenditure can have a material impact on results and requires explanation.

## Characteristics of a good price deflator

Atkinson (2005, p. 71) presents criteria against which the adequacy of price deflators could be assessed. These are set out in Table 5.5. These criteria cover the quality of the deflator (eg, comprehensive, full coverage), data availability (eg, sustainability, timeliness, periodicity, availability of cost weights), and the capacity of the deflator to illustrate the questions under consideration (eg, relevance, homogeneity, and quality change).

**Table 5.5 Quality Criteria for Deflators for Government Services**

Label	Short description	Examples/explanation
Comprehensiveness	The set of deflators should cover all components of expenditure to be deflated	There should be full geographic and sector coverage of the expenditure being deflated (eg, health deflators should cover the whole of the health system not just hospitals)
Coverage	The individual deflator should relate to all expenditure on the item to be deflated	Deflators for labour expenditure should cover all aspects of employee compensation (eg, all direct taxes and social securing contributions, and pensions as well as earnings)
Relevance	The deflator should correspond to the expenditure item to be deflated	For example, expenditure on books should be deflated using an indicator of the price change in books
Sustainability	The deflator should be available for the foreseeable future, and for a reasonable number of periods in the past	Micro studies of changes in price for only a single year have limited use: long time series are preferable
Homogeneity	Deflation should be carried out at a level of disaggregation that maximises homogeneity of items within a category	For example, significant difference in the movement of pay between staff grades would suggest that separate deflators are needed
Timeliness	The deflator should be available in good time after the end of the reference period	Estimation for missing periods may introduce bias
Periodicity	The deflator should be available on a quarterly basis	Annual figures may be satisfactory, but only where there is evidence of insignificant short-term change
Quality change	Where changes in characteristics of a good/service occur, price indices should reflect pure price changes only	Improvements in composition and consequently effectiveness of a drug should be distinguished from pure price change

Label	Short description	Examples/explanation
Availability of cost weights	Corresponding weights (of the same periodicity) for deflators should also be available	

Source: Atkinson (2005), p. 72

## Calculating deflators

As Schreyer (2010) notes, in some cases it is possible to draw on market price information (such as the Consumer Prices Index) for purposes of deflating values of non-market production. Alternatively, direct volume indexes can be constructed. These two approaches are explained in more detail below. As noted above, the approach to deflating expenditure can have a material impact on results and requires explanation.

### Publicly available market price information

Where the data required to estimate direct volume indexes are not available it can be possible to use publicly available Statistics New Zealand deflators. The key data that Statistics New Zealand publish that could be used are:

- **Consumer Price Index (CPI):** The CPI measures the changing price of a fixed basket of goods and services. This basket needs to be representative of the spending habits of New Zealand households and remains a fixed quantity so that changes in the CPI represent only price changes. As the quantity needs to remain fixed, changes in the size, performance, or functionality of products is taken into account. The basket of goods is also updated for changes in consumption over time (entering and exiting products). The goods and services covered by the CPI are classified into nine groups, 21 subgroups and 73 sections.
- **Subgroups of the CPI:** A number of the CPI groups and subgroups include data on state services (such as primary and secondary education). These subgroups reflect consumers' spending in these specific areas while the CPI reflects price movements more generally. The full CPI provides a "common numeraire" as a basis for real comparisons – so it indicates a common average real basket of goods that the funds in question could alternatively buy.
- **PPI:** Another possible deflator is a subgroup of Statistics New Zealand's Purchasing Price Index (PPI). Yet this index only covers the "productive sector" and measures changes in the prices of outputs that generate operating income and inputs that incur operating expense. It does not include prices for items related to capitalised expenditure, non-operating income, financing costs, or employee compensation. The subgroups are also not published at a further disaggregated level (eg, for primary and secondary schools).

As Schreyer (2010) noted, when considering the use of these data it is necessary to first ensure they are suitable for the purpose of deflating non-market production. In particular:

- The services supplied by the market provider have to be sufficiently similar to those supplied by the non-market provider – this has to be true for each type of service and for the mix between different services.
- The scope of the data has to match the scope of non-market production. This may not be the case when, for example, the CPI is designed to reflect prices for out-of-pocket expenditures and when consumers only pay part of the full price for the medical good or service. In this case, the CPI is not an appropriate tool for deflation of non-market production which relies on a concept of measuring production at its full cost.

For instance, Gemmell, Nolan and Scobie (2017b) compared quality adjustments based on the full CPI with those based on the CPI level 2 subgroup for primary and secondary education. This subgroup reflects consumers' spending on schooling and the CPI reflects price movements more generally. Between 2002 and 2015 the CPI grew by an average of 2.2% while the subgroup grew by an average of 5.2%.

Although nominal total teacher salaries grew at an annual average of 4.4%, this was less than the annual average growth in the level 2 CPI subgroup (of 5.2%). The result was that if salaries were deflated by the subgroup then real salaries would appear to have fallen, and productivity estimates would be higher than otherwise. For instance, using the subgroup as the deflator would mean:

- Salary-based measures would grow faster than unadjusted productivity measures, even though the growth in total salaries (4.4% nominal, or 2.1% real when deflated by the full CPI) has been much faster than the growth in teacher FTEs (1.2%) or price growth more generally (the CPI at 2.2%). Higher salary growth should (all else being constant) lead to lower productivity growth so the results of these two measures would be inconsistent.
- Similar issues arise in relation to school revenue, where nominal school revenue (average annual growth of 4.0% between 1997 and 2015) has risen much faster than student numbers (average annual growth of 0.4% for the same period).

Consequently they argued for using the full CPI to deflate teacher salaries and school revenue. The full CPI provides a “common numeraire” as the basis for all real comparisons – so it indicates a common average real basket of goods that the funds in question could alternatively buy. Another possible deflator was the education and training subgroup of Statistics New Zealand’s Purchasing Price Index (PPI). Yet this index only covers the “productive sector” and measures changes in the prices of outputs that generate operating income and inputs that incur operating expense. It does not include prices for items related to capitalised expenditure, non-operating income, financing costs, or employee compensation. The education and training subgroup is not published at a further disaggregated level (eg, for primary and secondary schools).

### Direct volume indexes

As Schreyer (2010) noted a “direct volume index is the weighted average of the volume indexes of different types of treatments, where the cost share of each type of treatment constitutes the weight”. There are a number of different approaches to producing these indexes (such as Paasche, Laspeyres and Tornquist indexes).<sup>12</sup> The choice between approaches largely reflects the data that is available and how volatile these data are likely to be. In particular, while a Tornquist index may be a theoretical ideal it is relatively data intensive and in most cases the data are unlikely to be volatile enough to justify its use. The simpler approaches of the Laspeyres (holding base prices constant) or Paasche (using current prices) index is usually preferred. Both the Laspeyres and Paasche indexes can be used for prices and quantities (the same index should be used for both) but the Paasche requires more recent data while the Laspeyres uses historic prices. Therefore, the Laspeyres index is likely to be most useful.

In a world with multiple inputs and outputs it is necessary to have a method for combining individual indexes. Simply averaging the change in indexes could be misleading as the volumes of different goods are likely to vary. For example, consider an economy with two outputs (eg, A and B), and where initial production is 1,000 units of A and 10,000 units of B, and over a year production of A increases by 5% and production of B increases by 1%. Also assume that one unit of A has the same value as one unit of B. Averaging the two growth rates would give a result of 3% while the actual growth in output would only be 1.4% (from 11,000 to 11,150).

There are a number of ways that different indexes can be combined. One option is to use a fixed base approach. This approach implicitly assumes that the value shares of the different goods does not change over time. However, when goods are changing in importance in the economy the assumption of them having fixed values (relative to other goods) is flawed. For this reason, weights are often rebased or adjusted regularly (eg, every five years or annually). Chain-linking is an approach where the weights are adjusted annually. This simply means “that, for each period, the base used is the weight from the previous period” (Goodridge, 2007). As Goodridge (2007) noted chain-linking provides a number of advantages:

<sup>12</sup> These are briefly explained below. As Goodridge (2007) noted “the Paasche calculates the expenditure needed to buy current year quantities, and is expressed as a percentage of what the expenditure would have been in the base period if the quantity consumed had been at current levels.” It divides spending on a basket of goods and services in the current period (eg, the sum of price multiplied by quantity for each product) by how much the same basket would cost in a base period. More formally this can be expressed as:  $(\sum(P_{tn} * Q_{tn})) / (\sum(P_{t0} * Q_{tn}))$ , where  $P_{tn}$  and  $Q_{tn}$  are prices and quantities at time  $n$ , and  $P_{t0}$  is the price in the base period. The main feature of the Laspeyres is that the weights used are taken from the base period (Goodridge, 2007). This can be expressed formally as:  $(\sum(P_{tn} * Q_{t0})) / (\sum(P_{t0} * Q_{t0}))$ , where  $P_{tn}$  is the price at time  $n$ , and  $P_{t0}$  and  $Q_{t0}$  are the prices in the base period.



- The first main benefit of chain-linking is that new items can be added to the “basket” every year. If the index is non-chained, new items can only be added to the base year.
- The second benefit is that by chaining the series, the comparison is with the previous year, rather than the base year. This is obviously relevant when most of the interest is in the annual change such as with prices, output and most economic indicators.
- A third benefit of chain-linking is that it removes substitution bias. This is a problem that is encountered when there are large shifts in both the weight and in the actual variable that is being indexed.

However, the ability to chain-link depends on the timeliness of the data used for the weights. And if the relative values of goods do not shift over time then chain-linking is unlikely to provide additional useful information. Further, the effect of chain linking will vary between the Paasche and Laspeyres indexes. In the case of the Paasche, chain-linking has the effect of reducing the index because growth is not calculated as a percentage of expenditure in the base period but instead is backward-looking. This means substitution is less pronounced when the index is chained together.<sup>13</sup>

### F5.3

Removing the effect that prices may have on comparisons over time helps ensure productivity measures accurately reflect changes in volumes. However, there is scope for the introduction of significant error from using the wrong deflators. Price effects can be removed using market information (eg, the consumer or producer price indices, or subgroups of these indices) or direct volume indices.

## 5.6 Conclusion

This chapter, and the preceding three, have aimed to explain how productivity measures for public services can be designed. The high-level message from these chapters is that, while there are some challenges in identifying and adjusting inputs and outputs, measuring state sector productivity is both feasible and desirable. Furthermore, the more productivity measurement that takes place, the better-placed agencies will be to overcome measurement challenges, by improving data sets and adjusting outputs appropriately for operating environments and changes in quality or prices. The remaining chapters in this report look at how the capability of agencies to measure productivity, and their incentives to use productivity information to improve public services, can be increased.

<sup>13</sup> In contrast a chain-linked Laspeyres index would rise by a greater amount than the standard Laspeyres.



## 6 Putting productivity measures in place

### Key points

- Productivity measures have an important part to play in ensuring scarce resources are put to their best use. Such measures complement outcome and effectiveness indicators.
- While the state sector has increasingly focused on effectiveness and outcomes, equivalent attention has not been paid to efficiency. For large sections of the state sector, there is little measurement or understanding of productivity.
- This relative lack of interest in productivity is the result of:
  - weak incentives within the public sector financial management system for agencies to seek efficiencies in existing programmes;
  - the concepts of “productivity” and “efficiency” having negative connotations in some parts of the public sector; and
  - low demand from ministers, and insufficient demand from senior public sector leaders for the necessary information and analysis.
- In order to expand and integrate productivity measurement into state sector practice, agencies should:
  - design productivity measures to complement outcomes;
  - design measures with the involvement of staff who deliver services;
  - collect productivity data as part of business-as-usual activity;
  - use productivity information primarily as the basis for conversations and learning about service improvement;
  - ensure agency leaders actively support the use of productivity measures;
  - develop measures that enable comparisons to be made between similar organisations, business units or outputs, to identify and promote sharing of good practice; and
  - treat productivity measures as one input into performance decisions, rather than the sole factor with high stakes impacts.

### 6.1 Introduction

Earlier chapters in this report have provided advice on how to measure productivity. This chapter, and the following two, discuss how to put productivity measures in place and ensure their appropriate use. In particular, the current chapter:

- discusses the place for productivity measures in New Zealand’s performance measurement and management system;
- explores why productivity measures are so rare in current practice; and
- lays out some principles for integrating productivity measures into state service practice and performance systems.

## 6.2 Productivity is a companion to effectiveness

As noted in Chapter 1, New Zealand's state sector management system has put an increasingly strong emphasis on outcomes. The most recent manifestation of this is the Better Public Services results, which set ten intermediate outcomes around which social agencies have organised their activities.

This shift occurred over the past twenty-five years, reflecting a desire across multiple administrations to improve coordination between agencies, focus the efforts of the state sector on key government priorities, and increase the value achieved from public funds. There have been a number of benefits from this growing focus on outcomes, including greater clarity about the government's goals, more attention paid to customer/citizen satisfaction with service quality, increasing use of data and more sophisticated analytical tools to inform decisions, and growing opportunities to shift resources toward more effective interventions.

However, effectiveness measures and targets on their own are not enough to ensure value for money and the sustainability of state services. There is a gap in the current performance management system. As the New Zealand Initiative comments in its submission:

Measures of efficiency added to existing performance measures would be helpful. The opportunity costs of achieving different BPS [Better Public Services] targets may vary across regions, and it would be useful to know whether BPS targets were met through massive investment that might have drawn resource from other non-targeted areas, or whether they were met at more reasonable cost. Similarly, where different agencies may have BPS targets that work towards similar outcomes, and where the agencies' activities are substitutes rather than complements, efficiency measures may assist in overall resource and task allocation. (sub.8, p. 6)

A summary of the key findings from the first round of the Performance Improvement Framework reviews of government agencies observed that, of the performance areas assessed:

Core Business Efficiency is the weakest result area, with only half of the core business areas rated Strong or Well Placed on efficiency. The large number of Unable to Rate ratings in the earlier reviews was a response to a lack of evidence on efficiency; something that is since reflected in a Needing Development rating at best. (SSC, 2013b, p. 13)

Productivity measures have an important part to play in ensuring that scarce resources are being put to their best use, and are a critical complement (rather than a substitute) to outcome and effectiveness indicators. As British experience illustrates, a focus on effectiveness without equivalent attention to efficiency can lead to waste:

Ministers were accused of devoting too much time and attention to delivering outcomes and too little time to controlling inputs: of throwing money at targets rather than focusing on whether that money was being well-spent. This criticism was not totally unfounded: ONS [Office of National Statistics] data shows public sector productivity in 2005 was 2% lower than it had been in 2001. (Barber, 2017, p. 13)

## 6.3 Much to build off, but also significant gaps

While the state sector management system has increasingly focused on effectiveness and outcomes, some pockets of the state sector have continued to give attention to efficiency. Through the course of this inquiry, the Commission has seen evidence of measurement and management techniques in public agencies that capture and use information on efficiency (eg, the Inland Revenue Department). The Commission has also been impressed by the data collection processes in place in a number of sectors, notably the New Zealand Police, Ministry of Social Development, and district health boards. The data collected by these agencies lends itself to a wide range of efficiency analyses, as well as other important uses.

However, it is clear that for large sections of the state sector, there is little measurement or understanding of productivity. In a number of sectors the Commission was asked to explore, little or no work had been done for some time to assess productivity. The Commission conducted its own analysis for some services (courts, schooling, early childhood education), based on publicly-available information. These analyses revealed gaps in the quality and coverage of data, and highlighted the need to capture better information on changes in the quality of inputs and outputs.

In the health sector, an absence of good data about primary care undermines comprehensive productivity analyses. Given long-standing policies over consecutive governments to move more health services into primary care, this is a significant gap that risks distorting measures of overall health system productivity. A visiting American health expert, reviewing New Zealand's primary care system, noted that while PHOs throughout the country collected data, "only visit data is submitted to the Ministry of Health" and "[m]inimal synthesis of these data occurs to strategically inform public policy" (Downs, 2017, p. 6). To address this, Down recommended that the

Ministry of Health and Treasury should collaborate on a more extensive data collection and monitoring programme for the primary care sector. While the government provides over \$900 million to subsidise primary health services, it has little information regarding quality, utilisation and outcomes of different types of primary care services. Data submitted to and analysed by the Ministry of Health should include utilisation of different services, diagnostic codes, quality metrics and demographic characteristics of patients (Downs, 2017, p. 6).

The Australian Productivity Commission has noted similar problems in primary health care in that country. Reliable health data can lead to improved health outcomes, including through assisting providers to self-evaluate their relative performance. Yet this requires ensuring "the data are of the right quality, the setting in which the data have been collected is divulged, the risks of unintended negative outcomes from misinterpretation or mismeasurement is assessed and remedied, and the measurement of performance is regularly refined" (APC, 2017b, p. 97). Yet, the "difficulty of accessing information forgoes opportunities for richer analysis, including of causal analysis of the factors that affect population health, benchmarks for performance at the regional level, and a greater capacity for testing the efficacy of some health promotion initiatives" (APC, 2017b, pp. 100-101). Similar opportunities are being missed in New Zealand.

In some cases the Commission was unable to access the data required for analysis. For example, most useful information about the performance of the court system is owned by the judiciary. There are some legal barriers and constitutional considerations that make it difficult to access this data. The Commission has therefore undertaken a case study of court productivity using publicly available data. However, there are significant limitations with the available data and the productivity measures constructed should be viewed as the best that can be produced within those limitations. The Ministry of Justice has undertaken to work with the Commission to discuss with the judiciary how measurement of productivity in the court system can be progressed. A recent study by Toy-Cronin et al investigated delays in High Court civil cases, and illustrates the useful analysis that can be done with good access to data. Those authors also concluded with a plea for better access to data:

There is an urgent need to improve data about who uses our courts, whether or not they are represented, and how their cases proceed. Without this information, we are unable to design a civil justice system that responds to the needs of those using the court and that protects its important public function. ...

The larger issue is that New Zealand continues to lack basic information about its civil justice system. This study makes a contribution to increasing this knowledge but it has also highlighted the lack of reliable data that is readily available about who is using our courts, why, whether the litigants are accessing with or without representation, and how cases progress once they are in the system. (2017, pp. ii, 117)

### F6.1

For large sections of the state sector, there is little measurement or understanding of productivity. In some sectors covered by this inquiry, the Commission was unable to access the data needed to conduct productivity assessments.

### F6.2

The lack of available data from parts of the health sector and from parts of the justice sector limits meaningful measurement of the efficiency of the New Zealand justice and health systems.

## 6.4 Why so little interest in or analysis of productivity?

There are numerous reasons why productivity has relatively little cachet in many public service agencies, chief amongst which are the incentives within the state sector financial and performance management system, negative connotations in some sectors, and the absence of demand for such measures.

### Financial and performance management incentives

Interviews with State sector leaders (Pickens, 2017) and previous work undertaken by the Commission (2014b) on regulatory institutions and practices suggests that the current public sector financial and performance management systems may discourage agencies from measuring and taking steps to raise productivity. Issues that have been raised, which are relevant to both Crown Entity Boards and state sector departments, include:

- **A focus on bringing in new revenue, rather than on better use of existing funds:** For example, former and current state sector leaders have commented that the desire to attract more funding, and competition between agencies for more funding, is top of mind for many state sector leaders. It is also worth noting that this attitude is incentivised because budget and department size are viewed as de facto sources of state sector leaders' prestige (Pickens, 2017; Barber, 2017).
- **Little attention in performance management on the quality or efficiency of service delivery:** Board Chairs of regulatory Crown entities interviewed for the Commission's regulatory institutions and practices inquiry frequently mentioned "that their [departmental] monitor was quite strongly focussed on the entity's financial performance" and some Chairs felt that "financial performance was given undue weight" relative to the organisation's core business and service delivery. As a result, plenty of time was wasted in low-value reporting and meetings (Spencer, 2014, p. 7). Nor does there appear to be much pressure from Select Committee for greater efficiency (Hitchener & Gill, 2011; Scott, 2001). Frequent changes to output classes and performance indicators limit the ability of external parties to assess the productivity of state agencies, and efficiency measures in output performance indicators have been "almost totally absent" (Lonti & Gregory, 2007, p.477).
- **Limited incentive to review existing programmes (vis a vis new ones):** There is significant scrutiny in the annual budget process of new initiatives and programmes, but little re-examination of existing activity, in particular, of existing services that are directly delivered by agencies or crown entities. Some senior public service leaders interviewed by Pickens noted that similar incentives applied to ministers – "that is, to simply get more money to do more things" (2017, p. 21). Formal review mechanisms – baseline and output price reviews – were introduced in the late 1990s in an effort to examine more closely the costing of specific "core" government services. However, these have since fallen into abeyance, with the last "output price review" of a department held in 2004.<sup>14</sup>

Other research into the state sector has reached similar conclusions. Norman (2003) found in interviews with public servants that "outputs have become a formality rather than a central feature of accountability" and that a "new form of incremental budgeting has been institutionalised".

The only serious analysis by Ministers is focused on a small pool of funding available for allocation on top of the baseline system of expenditure control. This small pool of funding is like a 'sandpit' that public servants are quite happy to let Ministers play in as long as the central baseline is not touched. (p. 170)

Norman's interviewees also argued that the inducements for departments to seek efficiencies within their own baselines were weak:

...in financial terms, the system points only in the direction of fiscal control. As described by one CE: 'if you overspend, it's a hanging offence; if you deliver outputs at least cost, tough luck – the [saved] funds will be taken away'. (p. 204)

Schick (1996) noted that, while New Zealand's budget guidelines are designed to encourage ministers to seek efficiencies before seeking new funds, reallocation within baselines "is not as common in New Zealand

<sup>14</sup> A baseline review was, however, commissioned for a Crown entity (the Commerce Commission) in 2008.

as in some other countries that have baseline systems” (p. 60). He attributed this, in part, to an “inattention to programme evaluation” in New Zealand (ibid). Other reviews of New Zealand’s public sector management system have similarly concluded that there is a need for more evaluation (SSC, 1999a; Advisory Group on the Review of the Centre, 2001; Cook, 2004; OAG, 2011).

### F6.3

Current public sector financial management systems provide weak incentives for agencies to seek efficiencies within existing programmes.

## Negative connotations within some sectors

Concepts such as “productivity” or “efficiency” have negative connotations for some public service organisations or sectors, especially those that provide non-transactional or more personalised services. Productivity is perceived by some to be in conflict with desired social outcomes, or with professional values. These objections are evident in some submissions made to this inquiry, public commentary from the judiciary, research on the health sector conducted for the Commission, and interviews of public sector leaders conducted by Pickens (2017).

A number of submitters argued that a focus on efficiency could have a distortionary effect, particularly on desired social outcomes:

Productivity in the technical sense defined by the present inquiry should not be used as a proxy measure for public sector performance and would have a significant distorting effect if adopted as such by government. In general terms, increases in technical productivity only contribute to better performance where measured outputs are well matched to desired outcomes. The methods of increasing measured productivity must also be carefully managed, in close consultation with frontline workers, so as not to increase waste or negatively impact on other activities that are necessary for achieving outcomes, but which may not be captured as measured outputs. (New Zealand Council of Trade Unions, sub.9, pp. 2-3)

NZNO’s main concern is to ensure that state services’ productivity is measured against indicators developed through coordinated, long term policy and planning, not only specific programmes. Meaningful results in the state provision of health services are those demonstrated over time at population level showing: improved equity, improved health and safety; and long-term sustainability. (New Zealand Nurses’ Organisation, sub.14, p. 10)

Output measures tend to exclude the intangible aspects of health services that contribute to overall effectiveness and public trust, including the social and emotional labour performed by health professionals. In education, many of the outputs and outcomes of quality teaching are likewise intangible and difficult to measure. Attempting to measure quality can be counter-productive, as too much focus on measurable indicators such as test scores and standards can have a distorting effect on education. In the justice system, efficiency measures do not capture fairness and may be counter-productive to quality outcomes if there is too much emphasis on fast turnover of cases. (New Zealand Public Service Association, sub.11, p. 6)

Chief Justice of New Zealand, the Right Honourable Dame Sian Elias has observed that the growing emphasis in justice policy on “more instrumental ends such as efficiency, cost-effectiveness, proportionality and...wider government objectives such as inter-agency co-ordination and information-sharing and relentless attention to reducing cost” risk undermining the independence of the courts and fundamental values and principles such as the right to a fair trial (Elias, 2017, pp. 1-2).

Research conducted for the Commission has also found a distrust of the motives behind productivity measurement, and perceptions of clashes between the objectives of such measures and professional values. In interviews of senior public service leaders, Pickens (2017) found that efficiency often carried negative connotations of cost-cutting and compromised service delivery (colourfully described as “the sound of chainsaws” by one interviewee) (p.4). Similarly, in her study of efficiency measurement by the New Zealand health sector, Knopf concludes that “technical efficiency/least cost production/productivity is the language of economist/accountants...without a health outcome or health service quality anchor, health practitioners will fail to engage” (2017, p. 29).

Concern with efficiency measurement is not universal, and is greater where personal relationships are important to achieving desired social outcomes, and where services need to be tailored to individual circumstances. In comparison, comfort with output measurement appears higher in units that directly enforce the law or regulations (Gill, Kengama & Laking, 2011), and with transactional services. This can be seen in the examples of MSD, New Zealand Police (see Chapter 4 and sub.12) and Inland Revenue (sub.7), where internal process changes have been made to improve productivity or data collection mechanisms introduced, which enable greater efficiency.

#### F6.4

Concepts such as “productivity” or “efficiency” have negative connotations for some parts of the public sector. These connotations include concerns that the pursuit of efficiency may compromise important social outcomes and values, or lead to reduced service quality.

### Little demand for productivity measures?

New Zealand’s public management system was explicitly designed to promote responsiveness to ministers (Palmer, 1988). The fact that there are few efficiency or productivity measures currently in place in public sector performance mechanisms suggests that there is little demand for them. This may partly reflect the factors outlined above (incentives within the public sector financial system and negative connotations). However, literature on performance management also suggests that executive politicians and public servants may see little benefit in measuring and reporting on productivity performance.

The absence of productivity measures can be seen in the Estimates of Appropriations for the 2017/18 financial year. These include many hundreds of outputs and performance indicators, only a handful of which are directly or indirectly aimed at raising efficiency. Only one was targeted at the reducing average service costs (Vote Revenue); the remainder sought improvements in the quality or quantity of outputs within constant or falling funding levels (Table 6.1). Most other cost-related performance measures sought to maintain unit expenses in real or nominal terms, although one output (Vote Environment policy outputs) envisaged an increase in per-hour costs.

**Table 6.1 Budget 2017 outputs with efficiency-related performance measures**

	2016/17	2017/18	
<b>Vote Revenue: Services to Process Obligations and Entitlements</b>			
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
Average cost of processing income tax returns, GST returns and employer monthly schedules	\$5.00 or less	\$2.17	\$4.00 or less
<b>Vote Transport: Accident or Incident Investigation and Reporting</b>			
	<i>Final Budgeted \$000</i>	<i>Estimated Actual \$000</i>	<i>Budget \$000</i>
Total appropriation	5 639	5 639	5 530
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
For domestic inquiries completed, the proportion open for fewer than 440 working days increases	50%	33%	60%

	2016/17		2017/18
For domestic inquiries in progress at 30 June, the proportion open for fewer than 440 working days increases	50%	80%	60%
<b>Vote Justice: Administration of Legal Services</b>			
	<i>Final Budgeted \$000</i>	<i>Estimated Actual \$000</i>	<i>Budget \$000</i>
Total appropriation	34,158	34,158	31,768
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
Number of new criminal legal aid applications administered	57 000-63,000	64 200	63 000-69 000
<b>Vote Justice: Provision of Protective Fiduciary Services</b>			
	<i>Final Budgeted \$000</i>	<i>Estimated Actual \$000</i>	<i>Budget \$000</i>
Total appropriation	2 627	2 627	2 627
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
Number of actions undertaken for individuals who have Protection of Personal Property Rights	5 744	6 950	7 000
<b>Vote Social Development: Management of Student Loans</b>			
	<i>Final Budgeted \$000</i>	<i>Estimated Actual \$000</i>	<i>Budget \$000</i>
Total appropriation	15 545	15 545	15 218
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
The percentage of initial entitlement assessments for a Student Loan completed within five working days will be no less than	90%	99.7%	95%
<b>Vote Social Development: Management of Student Support</b>			
	<i>Final Budgeted \$000</i>	<i>Estimated Actual \$000</i>	<i>Budget \$000</i>
Total appropriation	17 356	17 356	16 459
Assessment of performance	<i>Final Budgeted Standard</i>	<i>Estimated Actual</i>	<i>Budget Standard</i>
The percentage of initial entitlement assessments for a Student Allowance completed within five working days will be no less than	90%	100%	95%

Source: New Zealand Treasury (2017)



The most recent (2016/17) annual reports of the key agencies covered by this inquiry (Education, Health, Justice, Social Development) were equally light on productivity measures. One possible exception was the “contracting” output in Vote Health, which sought to reduce the ratio of departmental expenditure on procurement against relevant non-departmental spending (Ministry of Health, 2017, p. 62). Beyond this, the only other cost-based measure consistently used in all the reports was the “hourly cost of policy advice” indicator. None of these policy advice cost indicators sought to reduce average costs.

The other systemic performance tool that had a clear efficiency lens – the State Services Commission’s (SSC) Performance Improvement Framework (PIF) reviews of government agencies – has since abandoned explicit efficiency measures. In its earlier rounds, PIF reviews assessed how well agencies delivered their “core business”. These assessments were made in terms of both effectiveness and efficiency. However, in 2015, the PIF model was “refreshed”, with the questions posed to agencies:

upgraded to make explicit that the Core Business of State sector agencies is to deliver value for customers and New Zealanders, in accordance with each agency’s purpose. The focus and ratings for each Core Business areas have changed from ‘effectiveness’ and ‘efficiency’ to ‘value to customers and New Zealanders’ and ‘increased value over time’. (SSC, 2016a, p. 1)

It is possible that there is productivity measurement taking place within agencies that is not reported on, but beyond the examples already cited in this report, the Commission has not been made aware of any.

## The costs and risks of measurement

The theory behind performance management systems suggests that executive politicians such as ministers should want to have and use performance indicators to steer public agencies towards government goals (eg, Hood, 2002). In practice, however, performance information plays a fairly limited part in executive decision-making, including in New Zealand (Gill, 2011; Hitchener & Gill, 2011; Norman & Gregory, 2003; ter Bogt, 2004).

For politicians, performance information can be a double-edged sword – providing support for political objectives at times, and fodder for opposition at others. Van Dooren (2005) observes:

Performance information does not make politicians win or lose an election. However, the data that the performance measurement system yields may show weak performance. For politicians, everything they measure can be used against them, so they claim the right to remain silent and not to measure. Politicians may indeed have a disincentive to collect data. (p. 374)

Performance information appears to be of most salience to executive politicians when things go wrong. Pollitt (2006b) examined the use of performance information to steer agencies in four European countries and concluded that, on the whole, “politicians do not take much interest, and neither do citizens – unless and until disasters, scandals, or breakdowns come along” (p. 41). Nielsen and Moynihan’s (2017) study of Danish local councils similarly found that “only among jurisdictions where performance is relatively low do elected officials – when they are provided school test score data – become more willing to attribute causal responsibility to school principals for student outcomes, suggesting a negativity bias in attribution of responsibility to bureaucrats” (p. 270). This may reduce the inclination of public servants to propose productivity measures, especially in an environment where the sector may poorly receive them. A summary of the Performance Improvement Framework (PIF) reviews noted that there is “a lack of demand from senior managers for the sort of information and analysis they need to improve efficiency and, unsurprisingly, this is associated with a weak supply response” (SSC, 2013b, p. 34). This was due in part to “a lack of value placed by senior executives on strategic financial analysis that informs options and decision-making” (ibid).

### F6.5

The absence of productivity or efficiency measures likely reflects low demand from ministers, and insufficient demand from senior public sector leaders for the necessary information and analysis.



## 6.5 How to make progress

Encouraging greater measurement and improvement of productivity requires a number of systemic changes, including improvements to specific capabilities and the development of processes to collect and analyse data. These, and other changes, are discussed more fully in the following chapters. The remainder of this chapter deals with the principles for expanding and integrating productivity measures into public service practice and performance systems.

The principles outlined below reflect some of the lessons outlined above, and the need to manage undesirable effects from performance measurement. Smith (1995) cites eight main undesirable consequences from the use of performance measures (Box 6.1). Several submitters expressed concern about the potential for such negative impacts to arise as a result of unsophisticated productivity measures (New Zealand Council of Trade Unions, sub. 9; New Zealand Public Service Association, sub. 11; New Zealand Nurses' Organisation, sub. 14; New Zealand College of Critical Nurses, sub. 15).

### Box 6.1 Undesirable consequences from the use of performance measures

- **Tunnel vision** – “an emphasis by management on phenomena that are quantified in the performance measurement scheme, at the expense of unquantified aspects of performance.”
- **Measure fixation** – “an emphasis on measures of success rather than the underlying objective.”
- **Sub-optimisation** – “the pursuit of narrow local objectives by managers, at the expense of the objectives of the organisation as a whole.”
- **Myopia** – “the pursuit of short-term targets at the expense of legitimate long-term objectives.”
- **Misinterpretation** – misreading the “signals emerging from [performance] data” and sending “the wrong policy signals to the agent.”
- **Misrepresentation** – “the deliberate manipulation of data so that reported behaviour differs from actual behaviour.”
- **Gaming** – “the deliberate manipulation of behaviour to secure strategic advantage.”
- **Ossification** – “organisational paralysis brought about by an excessively rigid system of performance evaluation.”

Source: Smith, 1995

At the heart of the challenge with performance measurement is that the information can be “corrupting” or “corruptible”. Performance data is “corrupting” when its use encourages providers to change their behaviour in undesirable ways, and “corruptible” when it is possible to “affect the indicator without affecting the underlying phenomena that the indicator is attempting to reflect” (Cooley, 1983, p. 9). However, all forms of performance measures face these problems, and productivity indicators are no more prone to them than any other type. The right response to these problems is not to dispense with productivity measures, but to design them intelligently and use them appropriately.

### Design productivity measures to complement outcome measures

As noted above, productivity measures are an important complement to outcome indicators and help provide a balanced suite of information for decision makers. While the Commission acknowledges the concerns expressed by some submitters and other inquiry participants, the pursuit of greater productivity is not mutually exclusive with the achievement of social outcomes. Rather, it will generally support better outcomes when done well. A policy or intervention that achieves a desired outcome (eg, better health, greater safety) at a lower cost is clearly preferable to one that hits the goal but is more expensive. Tolerance

of inefficiency also limits the ability of governments to shift scarce resources towards more important outcomes.

Nor is efficiency incompatible with the maintenance of fundamental values, such as justice. As Ostrom and Hanson (1999) conclude in their assessment of the performance of US state courts:

[t]imeliness and the quality of justice are not mutually exclusive either in theory or in fact. Expeditious criminal case resolution is found to be associated with court systems in which the conditions also promote effective advocacy. Because effective advocacy underlies due process and equal protection of the law, it is an integral aspect of the broader concept of quality case processing. The evidence from this study suggests that well-performing courts should be expected to excel in terms of both timeliness and quality. (p.xiii)

Indeed, greater confidence around productivity can help protect core values and institutions. Banks (2002) observes:

While the separation of powers is fundamental to judicial independence, it is difficult to see how requirements to report on such matters as court delays, finalisations or lodgements could involve a significant conflict. *Indeed, without adequate transparency and accountability, the stature of the courts could ultimately be eroded.* (pp.13-14, emphasis added)

## Involve the staff who deliver services in the development of measures

Managers and staff are the most frequent users of performance information. Measures developed with staff are more likely to reflect the reality of service delivery, will be more accurate, trusted, sustainable, more effectively implemented and used, and will be more useful for identifying opportunities for improvement (Berman & Wang, 2000; Bourdeux & Chickoto, 2008; Moynihan & Pandey, 2010; New Zealand Council of Trade Unions, sub. 9; New Zealand Public Service Association, sub. 11; Yang & Hsieh, 2007). As Knopf (2017) notes in her review of efficiency measurement in the health sector, the “workforce has strong views and most of the expertise on the best way to provide services. They are critical to implementing service improvements” (p. 14). Involving staff in the development and implementation of performance measures can help manage the various forms of undesirable effects outlined above (Smith, 1985). Employee engagement in the development of organisational measures and strategies is also important for promoting higher performance, innovation, and staff wellbeing (OECD, 2016a).

One example of collaboratively-developed performance measures, which support better and more efficient service delivery, comes from the health sector – the Care Capacity Demand Management System (Box 6.2).

### Box 6.2 Care Capacity Demand Management (CCDM)

CCDM is a programme that helps district health boards better match their capacity to care with patient demand. It was developed by the Safe Staffing Health Workplaces Unit, a joint DHB-New Zealand Nurses Organisation (NZNO) initiative established following the 2005 multi-employer collective agreement negotiation round, where the NZNO had sought to have nurse to patient ratios included within employment contracts. Agreement was instead reached to pursue a more sophisticated mechanism.

CCDM has three main elements:

1. An evidence based method for setting the base staffing model in wards (numbers, skill mix and schedule) utilising validated patient acuity data.
2. Developing and supporting a system of multiple response strategies (Variance Response Management) within DHBs to manage short & midterm variance, so that demand can be met and safety and quality maintained.
3. Developing technical and social processes around a core set of data that is meaningful from the floor to the board, to ensure real time feedback and monitoring over time the demand/capacity match. (TAS, nd)

The core data set captures information against 24 indicators, which are grouped into three main categories:

- **Quality work environment:** this includes factors such as overtime, hours worked over contract, staff incidents, staff satisfaction/engagement and professional development.
- **Best use of health resources:** this includes factors such as casual use, excess accrued leave, and personnel costs.
- **Quality patient care:** this includes factors such as patient incidents, patient acuity and acute staffing shortage incidents.

As noted above, core data set information can be used at all levels of the DHB – from ward staff and warm managers, up to the executive team. Two key uses are to identify workload peaks and troughs, and support staffing and scheduling decisions. Other uses include informing unit improvement plans, and strategic planning with DHBs. Improvement is a key focus of data use.

CCDM was initially trialled in three DHBs in 2009 (Counties Manukau, Bay of Plenty and West Coast) and then progressively rolled out more widely. The Safe Staffing Health Workplaces Unit is now funded by all 20 DHBs, and in 2017, the intellectual property underpinning CCDM was vested in the Ministry of Health, in recognition of the “partnership approach to the programme’s development.” (TAS, 2017, p.37)

## Collect productivity information as part of business-as-usual activity

Ideally, data for productivity measurement should be collected as part of ordinary administrative processes. Examples of such collection mechanisms were discussed in Chapter 4 (MSD, New Zealand Police). Automatic or low-cost data collection is important for minimising errors and opportunities for misrepresentation. It also helps gain staff support for such measures by reducing the compliance burden on them.

## Use productivity information primarily as the basis for conversations and learning about service improvement

As discussed in Chapter 2, there are three main purposes of performance measurement – to learn, to steer and to give account. New Zealand’s system arguably gives the greatest weight to accountability, and much less to steering and learning (Hitchener & Gill, 2011). While accountability is clearly important, the main benefit from productivity measurement is the potential to obtain greater value and quality from existing resources. This is most likely to occur where productivity information forms the basis for conversations and learning about service improvements. As the House of Commons Public Administration Select Committee comment:

Public services need to be seen as learning organisations, with learning aimed at improvement. This puts the apparatus of measurement, including targets and league tables, into its proper context. A target may be missed, but if learning takes place in the process then that is a gain...Asking the right questions is, indeed, the key point about the proper use of targets, and performance measurement generally. Whereas some have seen measurement as the answer to public service problems, good managers see it as a means of asking the right questions. (2003, p. 33)

Data, openness to internal debate, flexibility, and adaptability to change matter for performance (NZPC, 2014; 2016). Noting generally poor performance from New Zealand state agencies in “core business efficiency”, a summary report on the first round of PIF reviews commented that agencies rated better:

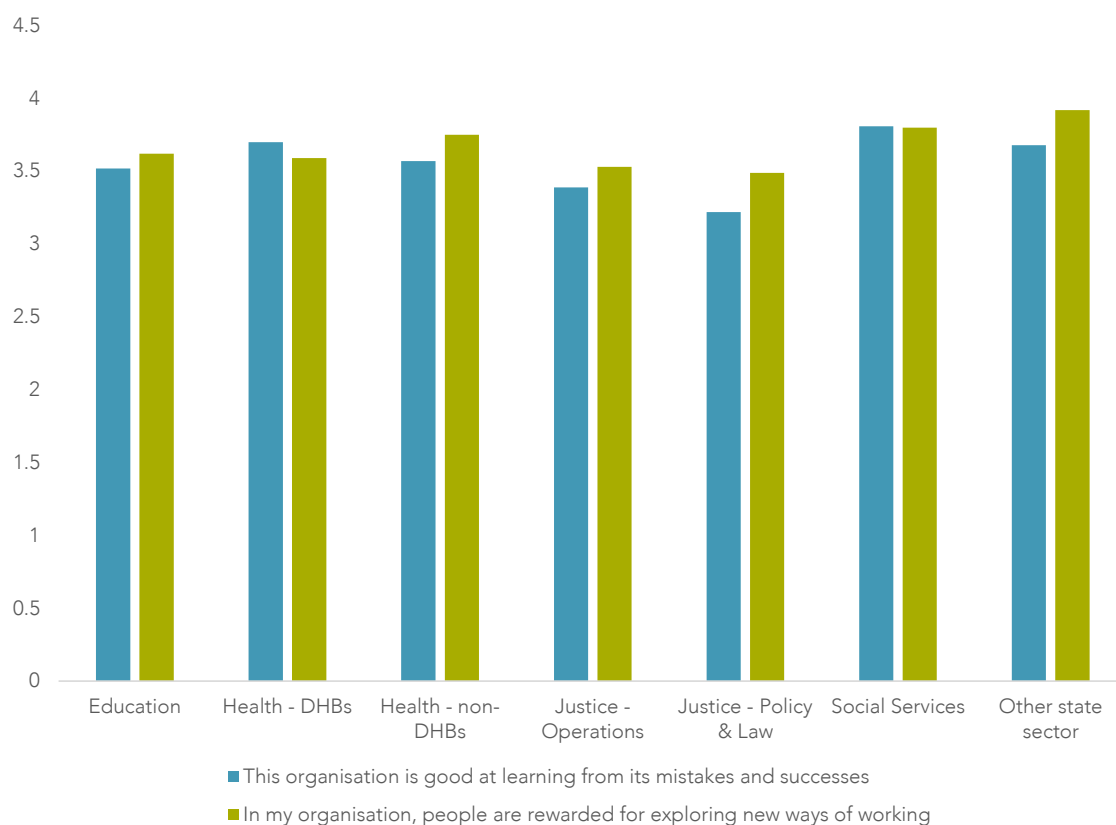
when they have good measures of relative efficiency...and can demonstrate they are managing to improve efficiency, eg, understand and manage their unit costs; actively consider outsourcing; have well managed procurement programmes. All of that requires investment in collecting and analysing relevant data...Discovering ‘what works’ may require an investment in research and analysis, as well as a

willingness to question existing practice, try new things and review the effectiveness of different approaches. (SSC, 2013b, p. 13)

Laking (2011) points to Work and Income and the Department of Corrections as agencies that have used performance information “as a diagnostic [tool] rather than a target” (p. 433). However, results from the Public Service Association/Victoria University of Wellington survey of public service workplace dynamics (Plimmer, Cantal & Qumseya, 2017) indicate that, for some agencies within the scope of this inquiry, having genuine conversations and learning about service improvements will require cultural change. The survey used a 7-point scale, where 1 equalled “strongly disagree”, 4 equalled “neither agree nor disagree”, and 7 equalled “strongly agree”.

Key findings from the survey suggest that “core” state sector workers do not see their organisations as being especially open to learning. Officials working in these sectors (education, health, justice, and social services) tended to somewhat disagree that their organisation was good at learning from its mistakes and success. “Core” state sector officials similarly somewhat disagreed that people in their organisation are rewarded for exploring new ways of working (Figure 6.1).

**Figure 6.1 Mean scores for selected learning culture questions by sector**



Source: Plimmer, Cantal & Qumseya, 2017.

Notes:

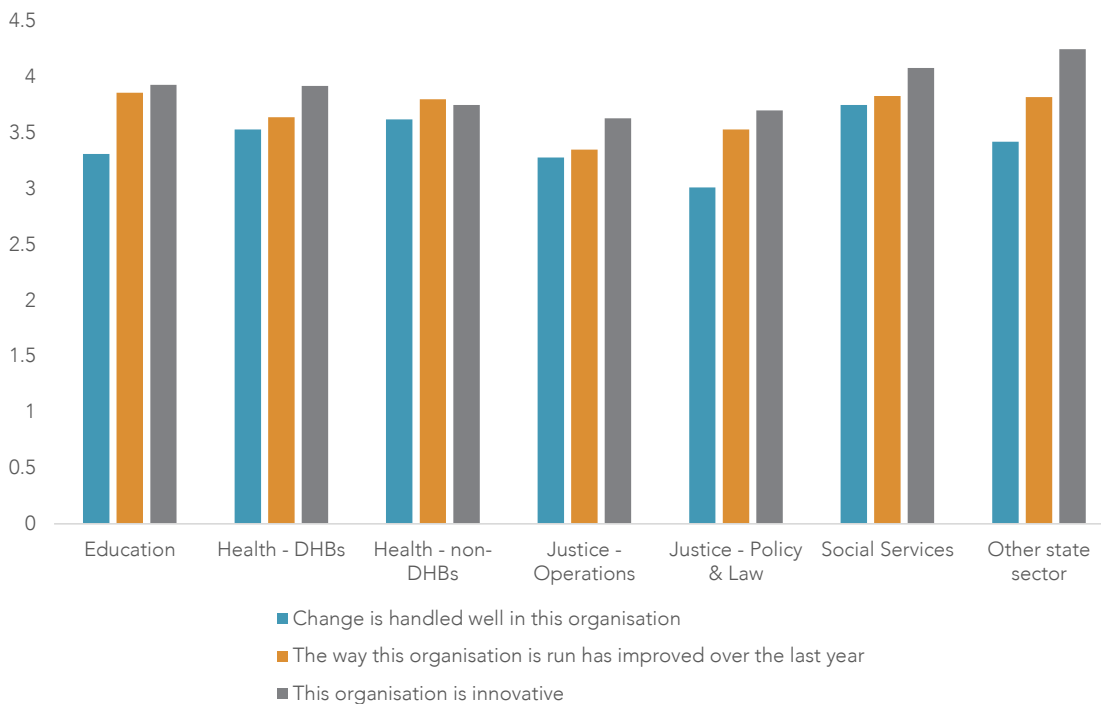
1. Scores less than 4 indicate disagreement with the statement. All results are statistically significant.

The VUW/PSA survey also explored perceptions of workplace innovation (ie, the extent to which organisations adapt to change, are well-run, and are innovative in improving their services). Studies have related higher levels of workplace innovation to higher levels of organisational performance and productivity (Langford, 2009). On average, results from the survey indicate that “core” state sector workers do not see their agencies as being particularly innovative or adaptable:

- officials across the state sector somewhat disagreed their organisation handles change well;
- officials across the state sector somewhat disagreed that the running of their organisation has improved over the last year; and

- officials working outside the “core” state service organisations tended to neither agree nor disagree that their organisation are innovative. However, officials working within the “core” state sector tended to somewhat disagree that their organisations were innovative (Figure 6.2).

**Figure 6.2 Mean scores against selected innovation questions by sector**



Source: Plimmer, Cantal & Qumseya, 2017.

Notes:

- Scores less than 4 indicate disagreement with the statement. All results are statistically significant.

Similar findings about low staff engagement in decisions and ability to innovate were evident in the 2017 New Zealand Police Workplace Survey. The Survey asked police staff for feedback “on a range of key organisation and workplace features such as vision, communication, teamwork, the job itself, development opportunities, as well as respect and integrity within the organisation” (New Zealand Police, 2017, p. 3) Police staff results were then compared against a state sector benchmark. Noteworthy results were that:

- 48.3% of 2017 police staff agreed with the statement “I am sufficiently involved in decisions that affect the way I do my job” (12.9% below the state sector benchmark);
- 50% agreed with the statement “I am encouraged to try new ways of doing things” (12.7% below the benchmark); and
- only 38.7% agreed with the statement that “NZ Police is interested in the views and opinions of its staff” (13.4% below the benchmark). (NZ Police, 2017, pp. 16-17).

The results above are consistent with the Commission’s observations in previous inquiries and with other government reviews. For example, the Savings Working Group (2011) found that many (but not all) public sector organisations have “no culture of, or commitment to, productivity and continuous improvement” (p.68) and that the public sector in general “has not been particularly successful in developing the systems, processes and culture which enable successful improvement and productivity initiatives” (p. 69). In reaching this conclusion, the Savings Working Group drew off a wide range of evidence, including reports of the State Sector Standards Board, work from the Managing for Outcomes initiative, PIF reviews and academic research. Similarly, the Commission’s 2014 inquiry into regulatory institutions and practices found the culture

of some New Zealand regulatory agencies places significant weight on managing risks to the organisation, at the expense of the efficient management of social harm<sup>15</sup>.

## F6.6

“Core” public service agencies do not appear to have the openness to change and innovation required to drive service and productivity improvements.

### Ensure agency leadership actively supports the use of productivity measures

The actions and attitudes of state sector leaders send powerful cultural signals to officials. A survey of state sector chief executives (CEs) undertaken by the Commission in 2014 found that 91% (21 out of 23 respondents) agreed that the senior leadership teams drive corporate culture, and that corporate culture and values influence how front-line staff operate (NZPC, 2014b).

There is considerable evidence that a leader’s attitude towards performance information has a direct impact on the use of performance data. For example, Ho (2006) surveyed mayors in the American Midwest and found that mayoral interest in performance measurement and benchmarking was one of the most critical factors in determining data usage. Similarly, based on surveys of Australian state government organisations, Taylor (2011) reported that managers’ attitude towards performance indicators were an important determinant of how data is used. And drawing off surveys of managers in US federal agencies, Dull (2009) argues that officials only truly engage in the use of performance data if they perceive it to be a priority for organisational leaders.

Therefore, a receptive culture is one in which leaders demonstrate credible commitment to the use of productivity measures in decision-making (eg, establishing systems and processes that encourage the use of productivity measures, rewarding staff that take steps to raise efficiency). When a leader publicly supports the use of productivity indicators, they signal that efficiency is important for individual and organisational success. Kroll (2013) notes:

Only if managers continuously signal that performance measures are important to them, will their staff be willing to devote their scarce resource to an improvement of the collection, analysis, and visualisation of data. (p. 6)

### Find comparators, to identify and encourage sharing of good practice

Comparative performance information (eg, between similar types of provider or output) is important for a number of reasons. It can help identify and enable sharing of good practice, promote discussion about the reasons for differing levels of performance, and find areas where improvement is needed. As will be discussed later (Chapter 8), a key failure of the state sector management system is the ability of good practice to spread.

The use of comparative information is also associated with better organisational performance and higher efficiency, as the SSC’s Performance Improvement Framework reviews have found:

...it is unsurprising that agencies tend to score better on results when they are clear about the goals they are trying to achieve... assess progress towards those results, such as using comparative performance data. (SSC, 2013b, p. 11)

Agencies rate better when they have good measures of relative efficiency, such as relative to public and/or private sector comparators. (ibid, p. 13)

A wide range of comparative performance information is already collected in the health sector through TAS (Technical Advisory Services), the Health Quality & Safety Commission and the Health Roundtable (a private benchmarking body, which many DHBs have joined). The Commission used some of this information in its own health productivity case studies.

<sup>15</sup> For example, the Commission’s survey of chief executives found that 48% (11 out of 23 respondents) of responding chief executives agreed with the statement “Agencies are often too risk averse when enforcing regulations” (26%, or 6 respondents, disagreed).

Other sectors covered by this inquiry either already have comparative performance measurement systems or arrangements on to which productivity measures could be attached.

- The Corrections Department has a Prison Performance Table, which tracks prisons against a number of metrics (eg, number of escapes, serious assaults, wrongful releases, prisoners in rehabilitation programmes).
- Communities of Learning (where groups of schools cluster together and use data to develop better learning pathways for students) provide another forum for productivity discussions.
- “Tribal” is a non-public benchmarking tool developed for the public tertiary education sector.

Examples of what comparative measures could look like in public services are available in the Australian Productivity Commission’s annual *Report on Government Services*. The Commission has recommended in earlier inquiries (APC & NPZC, 2012; NZPC, 2015b) that the New Zealand government should look for opportunities to benchmark public services, such as through participating in the APC’s annual report.

## Avoid attaching high stakes or incentives to productivity measures on their own

Performance measures that have high stakes or incentives attached to them are poorly-suited to learning objectives (Chapter 2). The risks of misrepresentation and gaming are heightened by the extent to which performance information is used to reward or punish (Smith, 1995; van Dooren et al, 2006; Cooley, 1983). There are many international examples of high-stakes performance measures leading to cheating or diversion from outcomes in such sectors as health (Bevan & Hood, 2006; Hood, 2006) and education (Berliner, 2011; Jones et al, 1999; Lewis & Hardy, 2015; Supovitz, 2009). Productivity measures are one input to wider performance management and decision-making frameworks, and should be used within that wider context. Ideally, productivity measures should form part of an agency’s internal management framework, linked to the organisation’s output and outcome measures, with regular data updates to inform decisions. Senior public sector leaders told Pickens (2017) that even imperfect measures can contribute to better performance:

This is because going through the process of producing that information helps officials think about and understand how best to improve outcomes - their intervention logic.

Reporting on variance against a target, even if that target is imperfect, can lead to valuable conversations which in turn feed into strategy and more effective interventions. (Pickens, 2017, p. 12)

### F6.7

The following principles are useful for expanding and integrating productivity measures into public service practice and performance systems:

- design productivity measures to complement outcomes;
- design measures with the involvement of staff who deliver services;
- as far as possible, collect productivity data as part of business-as-usual activity;
- use productivity information primarily as the basis for conversations and learning about service improvement;
- ensure agency leaders actively support the use of productivity measures;
- develop measures that enable comparisons to be made between similar organisations, business units or outputs, to identify and promote sharing of good practice; and
- treat productivity measures as one input into performance decisions, rather than the sole factor with high stakes impacts.

## 6.6 Conclusion

Productivity measures have an important role to play in complementing outcome measures. Their relative absence in current performance management systems creates a significant information gap, and limits the ability of agencies and government to achieve greater value for money. This relative absence is due to a number of factors – incentives within the state sector financial management system, negative connotations with the concepts of “efficiency” and “productivity” within some parts of the state sector, and limited demand from ministers.

To expand and integrate productivity measures into public service practice and performance systems, such measures should complement effectiveness indicators, be designed with affected staff, collect data based on daily administrative activities, form the basis for discussions and learning about service improvement, and be supported by agency leaders. This lines up with SSC’s (2013b) summary of lessons learned from PIF reviews about what it takes to improve the efficiency and effectiveness of public service agencies:

The key elements of good practice highlighted by this assessment...are:

- a. A value for money framework with good measures of effectiveness and efficiency which show how the components of value for money fit together
- b. A clear focus on what the agency is trying to achieve and the development of an operating model, and ongoing business process improvements, that support that goal (including benchmarking business process)
- c. A culture of review and continuous improvement, including in business processes
- d. Useful, timely and quantitative performance information that is closely linked to the clearly specified outcomes the agency is trying to achieve. (p.32)

Achieving these good practice elements across the state sector is likely to require more active engagement by central agencies with public sector leaders, to ensure they understand their operating models, and have a clear and compelling performance story for their organisations. Fully integrating productivity measurement into state service practice and performance also requires the right data sets and collection systems, and the right capability to use the data and incentives to seek productivity improvements. The following chapters discuss these matters in more detail.



# 7 Capability and systems needed to measure productivity

## Key points

- Agencies need to have a good understanding of their business flow, processes, and goals if they are to make the most of productivity measures; this knowledge is needed to identify what to measure and what questions to ask, to find ways to improve efficiency and add value.
- Public sector leaders (at all levels) need to lift the understanding of productivity, and build the capability to measure it. Leaders themselves need a sound conceptual understanding of productivity and measurement techniques, and the commitment to put measures in place.
- A range of other capabilities, systems and structures support good measurement. These include technical expertise in quantitative disciplines, and institutional systems and structures (for governance, strategic and business management, financial management, IT systems and data management).
- Many productivity measurement techniques are quite simple and do not require sophisticated skills and technology, while others are more complex. The key is in the interpretation and application of measures - to understand what a simple measure is useful for, and where something more complex is needed.
- Agencies could leverage their current emerging data analytics capability to measure productivity, or build the capability to measure it. As they become more familiar with productivity measurement techniques, that capability will grow. Although most agencies already have many of the skills required, few are using them to measure efficiency.
- Good financial management and accounting systems are largely already in place. However, in order to add the greatest value, financial processes need to shift from their current largely transactional and operational focus, to supporting strategic decisions. Expertise in agencies' finance teams could be leveraged to help measure productivity.
- Good IT systems and relevant and accurate data collection are major assets for developing useful productivity metrics. The New Zealand state sector has plenty of data, but it is not always in a very "usable" form, it is not always well linked (even within some parts of the same sector) and there are still some gaps in data collection that need to be filled.
- Wider data management issues such as who "owns" the data, and trust in how data will be used, are also important in embedding productivity measurement as "business-as-usual".
- The Performance Improvement Framework reviews could be used to better promote and support productivity measurement.

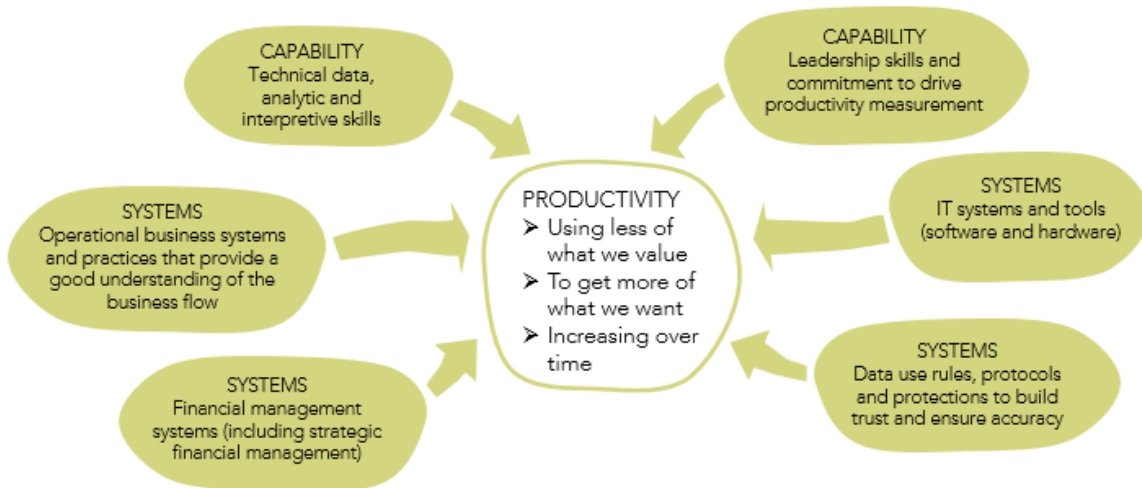
## 7.1 Introduction

A goal of this inquiry is to embed the systems and infrastructure that make it easy and useful to measure productivity, and to improve the awareness of the capability needed to measure it. In this context,

- "capability" refers to the range of skills and technical expertise needed to measure and track productivity over time; and
- "systems" describe the infrastructure, tools, processes, and protocols that are also needed.

Figure 7.1 below shows the range of capabilities and systems discussed in this chapter.

**Figure 7.1 Core capability and systems required to measure productivity**



This chapter discusses what capabilities and systems are needed, what the state sector currently has, and how it could move forward. In particular, it discusses:

- The operational “business flow” – the systems, processes, methods, tools, standards and routines that ensure consistent, quality, work flow practices that are agreed and understood by staff. A good understanding of these is essential before relevant and useful productivity metrics can be developed.
- The technical data analytic and interpretive skills and capabilities that state sector agencies need to have in their staff to measure productivity, and the current level of these skills and capabilities in state sector agencies.
- The leadership capability and commitment required to drive the measurement of productivity and ensure the measures are used.
- Financial management systems and processes that allow funding to be tracked, inputs and outputs accurately costed, and broader views of “value” (which might not be measurable) conceptualised.
- IT systems and the software and hardware tools to collect, match, collate and store data.
- Data management processes that support and proscribe the use of IT systems – the rules, requirements, and protocols for how data is treated that protect privacy and ethics.
- Making greater use of existing state sector performance tools to support agencies to measure and improve their productivity.

## 7.2 Productivity measures do not have to be sophisticated or complete to be useful

Measuring productivity need not be rocket science. As described in Chapter 2, productivity measurement techniques can vary in complexity from a simple ratio analysis to more complex approaches based on frontier techniques. Sometimes complex analysis is necessary, but in many cases, a simple form of measurement will be quite adequate to answer some questions, and prompt valuable performance discussions. As the Treasury (2015) have commented in their guide to cost benefit analysis:

A systematic method does not need to be complex, detailed and expensive. Even a rough back-of-the-envelope calculation can be logical and methodical. (p. 6)

The key is to “reduce the degree of uncertainty that would otherwise exist” (p. 6).

In developing productivity metrics, it is also necessary to consider the existing state of systems and capability levels within agencies. Some agencies will have a much stronger ability to start developing and using productivity measures than others. Less well-developed systems or missing capability will mean that a simpler productivity measure may need to be used – at least initially. Sometimes it will be easier to start with a partial measure, a measure of a service, or a part of a service. Thus, agencies could begin with simpler productivity measures and then over time build their capability to use more sophisticated techniques, if these are required to answer productivity questions.

### 7.3 Understanding an agency's operational business systems and practices

A good understanding of state sector agency's business processes and system flows is key to measuring productivity. In particular, work flows need to be documented and measurable so it is clear what work is done, how outputs are produced (with what inputs) and how quality is maintained. Understanding these work practices, routines and flows allows business processes to be disaggregated and productivity measures and indicators developed.

Data must be systematically recorded and tracked over time and the work flow processes should be sufficiently clear and consistent so organisations know whether apparent changes represent real changes in efficiency or not. Such data will also enable organisations to identify pressure points and bottlenecks, so they can find ways to improve efficiency (eg, through activities like re-engineering business processes or introducing new information technology – often using “lean” type processes).

Lean processes focus on identifying which parts of an organisation's business systems and processes are adding value to their desired business goals, and eliminating the rest. These can have a big impact, as shown in an effort undertaken by a network of North American hospitals and clinics (known as the Leadership Alliance) to explore the types and magnitude of waste like the kind identified by Zoe Radnor in her “bicycle book” example (see Box 7.1).

#### Box 7.1 The “bicycle book”

Professor Zoe Radnor, Dean of the School of Business at Leicester University, tells a story of an old practice she discovered at an English hospital. Staff who arrived at work by bicycle were required to sign a register when they arrived at work which recorded who travelled to work by bicycle each day. The books, when filled, were collected and stored. On examination, it was found that the “bicycle books” had been collected and stored since the 1940s. During rationing in the Second World War, people who travelled to work by bicycle could claim extra food rations as credit for saving on fuel. Decades later, the data was still being collected and a completely time-wasting routine was still in place for staff who used bicycles.

Source: Professor Zoe Radnor, School of Business, University of Leicester (presentation to the Productivity Hub, 23 August 2017)

The Leadership Alliance exercise identified 342 unhelpful rules or barriers to better hospital care. These were divided into three categories: habits, myths (organisation specific requirements that local leaders could change) and actual statutory regulations or requirements. Although they found that “wasteful statutory and regulatory barriers existed, the majority (78%) of obstructive, wasteful, rules they identified were fully within the administrative control of health care executives and managers to change” (Berwick et al., 2017).

A good understanding of the business systems and flows will also force agencies to take a whole of system perspective. This can mitigate the ongoing pressure for ministers and chief executives to be repeatedly pulled into the specific – to incidents and events - and away from the wider system impacts.

Standards and rules that ensure consistency and a specific level of quality allow outputs to be readily compared; this may make measurement easier, as it can reduce the need for complex adjustments for differences in quality, as discussed in Chapter 5.

Models that track agencies activities in this way are common for many transactional/low discretion services (in both the public and private sectors) but some models have also been developed for tracking the complex/high discretion services, such as the activities of the NZ Police, discussed in Chapter 4.

## The current understanding of business systems and practices is uneven

Within the sectors covered by this inquiry (health, education, justice, and social support), most agencies have some understanding of the flow and components of their businesses – particularly for specific parts of their sectors or businesses, or for specific service functions. However, the Commission found few agencies that had mapped their entire business from end-to-end, and no sectors that had a broad picture of their business flow across an entire sector or system.

### Police

The New Zealand Police “Know Your Business” model is a very good end-to-end map of an entire business, and one of the more complete models the Commission found (for a description of the model, see Chapter 4). However, although it is a good model of the whole of the police business, any productivity measures developed from it will not include the impact of the wider justice sector activity on police productivity (for example, the impact of activity in the Courts system); therefore, it is not a “whole of sector” model.

### Social support

MSD also has a good understanding of the work flow across the benefits and employment systems, and this is augmented by its actuarial (forward liability) investment approach. MSD’s iCAM model, discussed in Chapter 4, unpacks these activities in some detail, which has great potential to help develop productivity measures. But MSD has not articulated the connections between the benefits and employment services and the services purchased from the community services sector in any detail (this is discussed in more detail in the Productivity Commission’s 2015 report on *More Effective Social Services*).

### Health

The health sector does not routinely take a system-wide view of the whole of the health system either, but it has a very detailed and thorough understanding of some parts of it (such as the hospital sector). Its 2012 PIF notes that the Ministry of Health:

Appears to still respond to saving pressures on a largely incremental basis, rather than thinking more broadly about the changes to its operating model and modes of service delivery that may be required to deliver value-for-money. (p. 21)

Similarly, the Commission’s research note on Health Sector Productivity Data, notes that:

“Hospital data (both inpatients and outpatients) tends to be the most readily available (and utilised for productivity studies) but [it] only provide[s] a partial view of the sector. To fully understand trends in the health system more system wide data are required” (Fraser and Nolan, 2017).

Partial service or agency-only measures can still be useful for assessing productivity. However, when using and interpreting these measures there needs to be good understanding of what they are showing and what they are not.

#### F7.1

All agencies have the starting capability to develop at least some simple productivity metrics - for at least parts of their systems (or some of their services).

#### F7.2

A better “whole-of-system” understanding of agencies’ business flows will enable more comprehensive measures of whole system productivity to be developed. Few agencies currently take a whole-of-system approach to productivity improvement.

## 7.4 Data analytic and interpretive skills are needed

Developing rigorous productivity measures, which can be used to support a range of decisions, requires data management skills and analytical skills to interpret the results of the measures and suggest ways to improve productivity.

### Data management skills

Productivity measures require good data. Some agencies have more advanced data systems than others, and most have multiple systems that must be integrated (or interfaces built) for them to operate together. Integrating data systems, or building new ones, requires specific capability. This includes the capability to design and construct processes for drawing together data from multiple systems, and institutions to govern the flow of input and output data within and between agencies.

Most state sector agencies already have some capability in this area. Designing and constructing data sets requires data architects and data engineers. Legal and policy professionals with skills in areas such as privacy law and data security are needed to develop effective institutions for the governance of data (Chapter 8). In addition to “technical” skills, data engineers, architects and others must be able to communicate effectively so they understand the context and purpose of the data infrastructure, and can support the business “owners”.

### Capability to analyse or interpret the productivity measures or metrics

For data to become knowledge it must be analysed. The capability to analyse and draw conclusions from productivity measures typically involves asking questions about what is driving observed changes in productivity and, if necessary, carry out further analysis to gain a deeper understanding.

The capability needed to analyse productivity figures will vary depending on the context. A more sophisticated analysis may require understanding of:

- The impact of the chosen approach on the results and alternative methodologies if necessary. For example, to understand how the method of adjusting for output quality or inflation impacts on results.
- How changes in financial processes and incentives drive concern for productivity and may impact the results. For example, changes to an agency’s cost-allocation or cost-recovery process can impact on productivity measures.
- The institutional environment within which the measured public services are being delivered.

“Data science” covers many skills useful for developing productivity measures. It has been an emerging field in the private sector for over a decade now, but its move into the public sector is more recent. This has led some agencies to start to identify data analytics as a skill set they need.

### Evidence about current levels of data and analytic capability is poor

The Commission has not been able to access detailed information about the numbers of people with the technical skill sets described above who are employed in the agencies covered by this inquiry. Job titles (such as Principal Advisor or Policy Analyst) do not provide information on either the skill-set or the educational background of particular staff; nor do they indicate whether they use specific skills in their current roles. For example, state sector staff may have degrees in economics or statistics (we know anecdotally that some do), but they may or may not use those skill sets in their daily work.

The State Services Commission (SSC) collects payroll information via the Human Resource Capability survey of public sector agencies. This includes roles coded to ANZSCO classifications by job title, but job titles do not accurately capture the capability in agencies, either in terms of staff skill-sets, staff education and training or the range of activities staff are engaged in.

It is likely that much of the existing capability in this area is currently being directed towards implementing “social investment” approaches, as these use much of the same capability that could be used to measure productivity.

## **Strategic documents have little workforce capability information, and tend to have an agency (rather than sector) focus**

Departmental Four Year Plans, Annual Reports, Statements of Intent, and Performance Improvement Framework (PIF) reviews also contain little information on skills and capability. The Four Year Plan Guide (SSC and The Treasury, 2016) states that Four Year Plans should cover workforce issues, including leadership capability and workforce capability and capacity, diversity and inclusion, and ICT people capability.

Four Year Plans reviewed by the Commission did not generally show how agencies will position their workforce to meet their future needs. They all acknowledged workforce and capability development as an issue (often related to the expanding role of technology), but responses typically involved high-level statements rather than specific strategies and actions. The Commission found very little detail or discussion of the skill-sets that agencies currently have, the skill sets they will need in the future, or the steps they will take to acquire those skill sets (eg, through training or progressively employing new staff).

For example, MSD's 2017 Four Year Plan (MSD, 2016) notes that:

Like other agencies the Ministry is already faced with workforce challenges and is bracing itself for change ahead including:

- the impact of disruptive technology on roles [...] this refers to new technologies that often disrupt the workforce as the impacts are unknown and untested
- the availability of staff for increasingly competitive<sup>16</sup> roles. (p. 24)

It also refers to "the adoption of new self-service digital channels" and notes the need to reduce processing staff, and invest in "technical skills in the back office, and digital competency at the frontline" (p. 38). This shift (from face-to-face services to online and digital channels) suggests an intent to improve efficiency. However, there is little discussion of the extent of change envisaged or the specific skills required.

Four Year Plans also tend to focus internally on individual agencies, rather than looking across the wider sector. For example, the Ministry of Health's 2014 Four Year Plan has a comprehensive analysis of departmental capabilities and gaps, but only for the Ministry itself:

An initial assessment of workforce capability gaps and pressures across whole-of-Ministry was presented in last year's Four-Year Plan and a stock take was recently commissioned ... to determine what had changed, remained the same or emerged. The results ... confirm current Ministry-wide workforce capability gaps in a number of the areas identified in the initial assessment. (p. 29)

The Ministry of Health's 2017 Four Year Plan notes that a revised New Zealand Health Strategy requires "a stronger data and analytical function at the centre" (p.17), and adds that the Ministry is establishing a Client Insights and Analytics team to support this. But it does not connect the Ministry workforce with those of other sectors or even with the rest of the health sector.

The New Zealand Police 2017 Four Year Plan proposes upskilling existing staff and bringing in new capability, but it is pitched at a very high level (for example, it is unclear whether it refers to sworn officers, non-sworn back office staff, or both):

Police will need to both increase overall staff numbers and develop new skills in a number of areas. This will be achieved by providing training to existing employees to develop these skills, and by hiring in specialists with capacity in these areas. Police will need to develop relevant training, management and career progression for these new capacities. (p. 58)

The Ministry of Justice's 2012 PIF notes:

The People Strategy includes a section on planning for future workforce needs and a section on attracting talented people. While the strategy is generally reported to be supported by staff, it is also seen as having stalled and lost focus. The Ministry's self-review noted an implementation on risk to the People Strategy due to the changing nature of what is required and the fact that forward planning happens on an ad hoc basis, often at a team or business unit level. (p. 51)

<sup>16</sup> Technical roles for which there is likely to be competition between employers for skilled staff

## The future supply of data and analytic capability is also unknown

In terms of the volume of people entering the workforce with data and analytic skills, the picture from the education system is mixed. Data on recent graduates from New Zealand universities (Table 7.1) does not show significant growth in economics, econometrics and statistics graduates - who are among those that would have the skills to interpret and analyse data (though there are other disciplines that can also conduct analysis). However, there has been a large increase in graduates in fields that could build infrastructure and datasets – such as information technology, computer science and electrical engineering.

**Table 7.1 Number of domestic students completing Bachelor's degrees by selected field of study**

Bachelor's degree by detailed field of study	2008	2009	2010	2011	2012	2013	2014	2015	% change
Econometrics	35	45	55	65	60	40	35	20	-
Economics	455	415	485	520	560	505	485	500	10%
Statistics	180	180	160	175	205	200	200	215	19%
Mathematics	140	165	140	180	205	225	185	175	25%
Information technology	385	390	370	450	565	565	630	645	67.5%
Physics	80	75	75	85	85	100	85	140	75%
Electrical and electronic engineering	120	110	130	170	225	200	215	225	87.5%
Computer science	325	315	330	335	405	465	585	645	98.5%

Source: Ministry of Education, 2016.

Notes:

- The increase in total numbers of graduates during this period was approximately 25%.

There are also new qualifications being offered by tertiary providers that recognise the growing importance of data in society. These qualifications tend to blend courses from computer science, mathematics, business studies, and domain knowledge. Information about the number of people attaining such qualifications is not well captured by existing government data collection.

### The state sector is not the only party seeking these skills

Quantitative skills are also in high demand in the private sector. This competition for talent puts upward pressure on pay rates. Some of the PIFs and Four Year Plans reviewed by the Commission raise this as a concern, but do not offer solutions or plans to address it. For example, the MSD Four Year plan outlines success factors (such as “making better decisions through using workforce data, and analytics” (pp. 24–25). This implies a shift towards staff with more scarce skill-sets such as data analytics where they will need to compete for staff with the private sector, but it is not explicitly articulated.

### There is potential for Statistics New Zealand to have a greater role

Statistics New Zealand has an interest in, and concern for, data and analytic capability in the wider state sector as well as in their internal capability. Its 2014 PIF describes the challenge as including wider state sector capability:

With the growth of Big Data, there is likely to be heavy demand for the key skills that Statistics NZ will need. The agency will need to be more deliberate about career development options and professional engagement across the State sector to build capacity and capability in its areas of technical speciality to meet its own needs and lift the capability of professionals supporting the Official Statistics System. (p.10)

The 2017 follow-up PIF suggests that the capability challenge is still large:

Statistics New Zealand has invested in a number of areas to support its existing and future workforce and has a workforce development strategy in place. However, it is still exploring its role in addressing the system capability gap (as are other agencies) and what its own future needs will be: where will it require deep technical understanding and where will it require generalists. (p. 16)

A new functional lead role for data and analytics has been established, and given to the Chief Executive of Statistics New Zealand and Chief Government Statistician. Although still in its establishment phase, this role is likely to involve Statistics New Zealand working with agencies to support them to do more with data.

A work programme being developed inside Statistics New Zealand includes capability building as one activity. The Commission understands that this will include looking at the data analysis landscape, including the people, roles, and skill-sets in this space. Looking beyond data scientists, Statistics NZ is also considering the leadership capability required and barriers such as risk aversion.

F7.3

There is little information about the level of capability to measure productivity across the state sector (in terms of both current capability and expected future need).

F7.4

Some technical and analytic capability exists in most agencies, and much can be done to measure productivity with this current capability, as long as each agency has a thorough understanding of its operational business flow, processes, and goals.

F7.5

Productivity measurement requires a similar skill set to that needed for social investment approaches; therefore, better connections between these two should be built.

## 7.5 Leaders need to drive productivity measurement

To be champions for measurement, leaders should themselves have a broad understanding of the concept of productivity, the assumptions behind it, the range of ways it can be measured, and potential uses for the measures. The views of the “technical people,” who understand the detail of the data, should be well-integrated into the strategic management of the organisation.

Leaders also need a good understanding of the limitations of the different types of measures, the types of business questions that productivity measures may be able to answer (and the questions they cannot answer), and how this is relevant to the business of their agency; this should complement a deep understanding of their own business flow content, described in section 7.3.

However, most importantly, state sector leaders need a commitment to embed productivity measurement and a productivity mind-set in their agencies. They must want to measure productivity - to build up the capability to conduct different types of measurement over time (as required) and to act on that information.

### **There is some misunderstanding and resistance around the term “productivity”**

In twenty semi structured interviews with current and former State Sector leaders, Pickens (2017) explored how state sector leaders think about and conceptualise efficiency:

All interviewees understood the importance of basing their decisions on the concepts underlying economic efficiency, although some were less familiar with the technical terms. In particular, many attributed efficiency to mean *productive* [ie, technical] efficiency only.

Most interviewees felt the understanding of efficiency across the state sector was poor. Almost all agreed that the language of economic efficiency was neither well understood nor necessarily well accepted as an appropriate objective for the state sector (Pickens, 2017, p. 31).

This suggests that a key barrier to measuring productivity is the understanding of it, rather than the capability to develop and implement such measures. In other words, it is about leaders and staff



understanding what decisions different kinds of measures can be used for. This misunderstanding is perhaps one of the causes of the cultural resistance to using productivity measures discussed in Chapter 6.

## Building understanding and encouraging measurement

State sector officials largely mirror the attitude leaders have towards the use of efficiency measures. Therefore, state sector leaders must lay the groundwork for efficiency improvement by demonstrating a commitment to organisational learning in general, and the use of productivity measures in particular.

Encouraging a receptive culture is not solely the responsibility of chief executives and senior leadership teams. Ministers, lower tier managers and informal leaders all have an important role to play. Indeed, it is often an official's direct manager that sets the "tone" of a workplace – particularly in regional offices.

There are a number of ways that leaders can encourage the use of productivity measures. These need to be considered as a self-reinforcing dynamic that fits into an agency's business model, rather than a "checklist" for those who do not understand their business model as well as they should. However, some suggestions are described in Box 7.2 below:

### Box 7.2 Examples of specific actions leaders can take to seed a receptive culture

- Regularly (and consistently), pay attention to, and prioritise, organisational learning and the pursuit of efficiency and effectiveness.
- Allocate resources in a way that supports learning and efficiency.
- Provide a role model for officials and coach other leaders in how to encourage learning and efficiency throughout the organisation.
- Put in place organisational systems and procedures to encourage learning and the use of productivity measures in decision-making.
- Foster an expectation that staff share information. Sanction staff that withhold information and reward those that develop systems to make sharing information easier.
- Create multiple channels of communication that enable staff to connect with and learn from others. This is particularly important in cases where staff operate from multiple (regional) locations.
- Seek input from staff on barriers to learning and improving productivity. Empower staff to develop solutions to the barriers, and act on their suggestions. Encourage staff by publicly acknowledging and rewarding their efforts.
- Include learning and the pursuit of efficiency in statements of organisational beliefs and values.
- Link managers' performance measures to the steps they take to encourage staff learning and knowledge sharing.
- Reward staff that demonstrate a commitment to learning and the pursuit of efficiency.
- Create "space" for staff to experiment with new ways of operating. Treat unsuccessful experiments as learning opportunities rather than "failures". Reward staff for experimenting – even when experiments are unsuccessful. Publicly emphasise the importance of learning from failure.
- Personally (and publicly) encourage people at all levels to ask questions and share stories about what they have learnt from previous experiences.
- Seed a workforce that embraces learning and productivity improvement by hiring and promoting on the basis of staff capacity for learning and ability to identify improvements in working practices.

Source: Adapted from NZPC (2014b)

Training for leaders in the techniques of productivity measurement is also useful. The Data Futures Forum has noted:

We think one gap is in education for decision-makers and their advisors, to help increase the demand and uptake for data-driven innovation and research. When business and community leaders, senior officials, and ministers see the opportunities in data, they begin to expect data-enabled services, to require broad evidence-based policy thinking and make space for data-driven innovation. (2014, p. 53)

Pickens (2017) also finds that senior leaders, in particular, “need to understand efficiency well enough to be able to persuade others of its importance” so that they can “explain the connection between efficiency and achieving broader community outcomes”. He notes that a:

formal training programme for state sector leaders and aspiring leaders may be required. This should include training in the measurement of efficiency and the use of efficiency indicators in state sector contexts. (Pickens 2017, Finding 3)

This training should explain productivity using examples and language relevant to each sector (eg, health, social services, justice). Pickens (2017) suggested that discussions about productivity should be presented in the language of the social outcomes sought, rather than the language of economic efficiency.

#### F7.6

State sector agencies already have much of the capability to measure productivity, but leaders are not currently incentivised to understand, measure and act upon it. Framing productivity in the language of social outcomes rather than that of economic efficiency would help make these messages more accessible.

#### F7.7

There is a lack of connection between the people with technical skills (to measure and understand productivity) and the managerial and strategic activities of agencies.

## 7.6 Sound financial management systems

New Zealand’s public sector financial management system governs the way agencies plan, resource, and report on their activities. As outlined in Chapter 1, this system already has the key elements needed to measure productivity. These key elements are:

- Ministers and departments agree and specify *ex ante* the outputs they will deliver over the coming year. Outputs are specified in terms of cost, quantity, quality, and timeliness.
- Output expenses are fully costed, including “the full cost of producing and supplying outputs measured in accrual accounting terms; and...the full allocation of overhead costs” (Section 2, Public Finance Act 1989).
- Outputs are combined into groups or classes that form a common set of goods and services, which become the basis of appropriations in the government’s budget. Appropriations are authorisations from Parliament for spending, and limits are set on how the funds can be used.
- Once costed and agreed, a department’s “baseline” is fixed in nominal terms. There are no automatic adjustments to reflect changes in input costs.<sup>17</sup> However, departments can bid for additional funding at the annual budget round.
- At the end of the financial year, agencies report on their performance against outputs in their annual reports and are subject to review by a Parliamentary select committee.

<sup>17</sup> Some appropriations – such as benefit or school expenses – are, however, regularly adjusted to reflect changes in volumes (eg, more benefit recipients or school students).

## **New Zealand’s fiscal management system provides a good base for measurement**

New Zealand’s financial system is well positioned for productivity measurement – an inheritance from the public sector reforms of the 1980s and 1990s. As outputs are already specified, agencies have a good starting point to develop productivity measures and metrics. Two critical elements of the financial management system are full costing and fiscal transparency. Van Dooren et al (2007) point out that the use of accrual accounting helps make the true cost of government activities more transparent, which both enhances financial management practices and enables increased efficiency.

## **There is a sound core, but limited financial involvement in strategic management processes**

Despite this sound core, the skills of agency financial teams are underutilised in setting strategic directions and improving service performance. Most government finance teams look closely at where funds are going (tracking) to check they are being spent on specified deliverables (output compliance). An SSC summary of lessons learned from PIF reviews comments that the:

best agencies develop and use information and analysis to support decision-making to add value and manage risk. The others avoid risk rather than manage it. They also see corporate functions as overhead that exists largely to pay the bills, meet compliance requirements and, at worst, keep monitoring entities off their backs (SSC, 2013b, p.7)

SSC go on to note that there seem to be “weakly developed capabilities across the State Services, in part because of weak utilisation and leveraging of the finance, information management and asset management functions. These key support functions are typically seen as transactional and need to be seen instead as real business partners” (2013b, p.31). The functions:

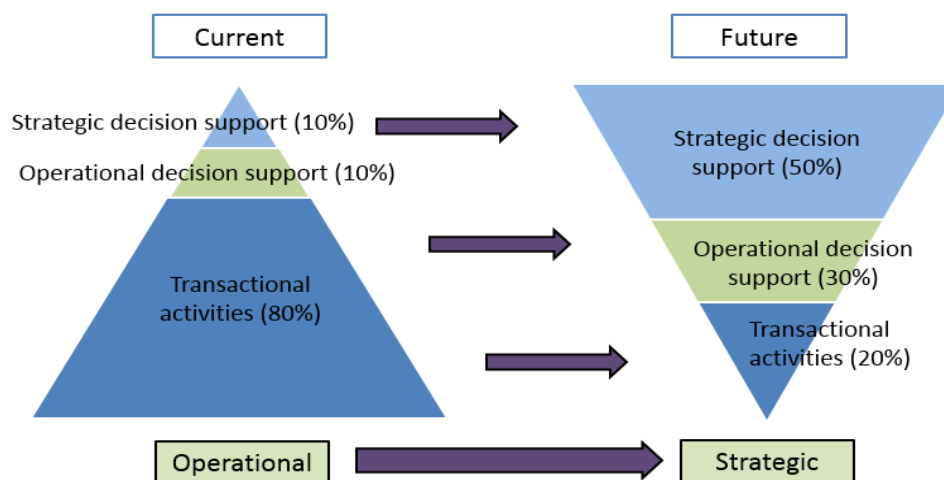
do not regularly inform executive decision-making. They tend to report to agency managements through a corporate services executive...and not to be strongly linked to strategy formulation or decision-making (p.33)

## **A more strategic focus is needed**

McKinsey (2017) is of the view that government finance functions, from finance ministries to national treasuries to department-level financial officers, need to expand their focus beyond budgeting and fiscal stewardship to actively drive outcomes, identify productivity-improvement opportunities, and champion change. They need to help state sector agencies map clear routes to better outcomes, develop finely tuned measures of progress, and provide early warnings of storms ahead (McKinsey, 2017).

A more strategic approach to financial management would connect finance teams better to the operational parts of an agency and look at what is being achieved as well as at what is spent. A proactive, rather than an audit, approach would help agencies’ finance teams make more meaningful use of efficiency measures (and also guard against misuse of them).

The Head of Finance Profession and Chief Government Accountant, based in the New Zealand Treasury, has identified reorienting government finance functions from operational, transactional activities, to a role that better supports strategic decisions, as an important goal. Figure 7.2 below is Treasury’s way of presenting this desired shift in emphasis.

**Figure 7.2 A shift in thinking from transactional to strategic financial management**

Source: The New Zealand Treasury (2017c)

To achieve this shift, the Treasury have set up a Financial Development Capability Programme; this includes establishing a group of “key position” CFOs (from the big agencies) who meet regularly, and a programme of activities offered to key finance staff from other agencies across the state sector. Activities include:

- training and professional development at three levels: senior management, mid-level, and early career;
- mentoring and secondments;
- identified agency “champions”; people who report to key position CFOs in all the main agencies (including the Ministries of Health, Education, Oranga Tamariki, and Police); and
- a shared workspace which has 265 members, and is used as a communication tool with the members of the government finance profession.

The Commission has not had any feedback on whether this work programme will be able to support agencies to build their productivity measurement capability. However, the Commission is encouraged that a shift from an audit/critique role to a more proactive, and supportive, role for finance teams would be a shift in the right direction.

**F7.8**

New Zealand has a well-developed government fiscal management system, which provides a good basis to measure efficiency.

**F7.9**

The Head of the Finance Profession is working to shift agencies from a transactional to a more strategic financial management capability. This would enable agency finance teams to better connect productivity measurement to operational decision making.

## 7.7 Robust and integrated data systems and tools

A core requirement for measuring productivity is sound and robust data collection, reporting, and storage systems. Quantitative information is needed for both inputs (including costs) and outputs (including volumes), of state sector services. These systems must capture what is produced when and by whom in terms of both quality and quantity. Accurate and useful productivity measurement is completely dependent on both the availability of good data, and the ability to use the data to draw sensible insights that have practical application.

Ideally, government IT systems should also be integrated (as much as possible) to enable productivity measurement. This would mean information on the same people (whether clients or staff) and the same

types of activities (whether whole services or outputs, or whether smaller collections of activities that contribute to an output), could be examined and analysed as parts of a whole. System integration could occur across service streams within one organisation, across different organisations within one sector, across the whole of government, or even to include services provided by non-government players. Across the sectors considered in this inquiry, system integration is beginning, but it is still at early stages.

## **There are large volumes of data across state agencies, which could support productivity measurement**

The ability to integrate data would allow more sophisticated (therefore, more useful) productivity measures and metrics to be developed. Government holds a large quantity of data, including in all the sectors in this inquiry. In fact, a significant number of state sector leaders interviewed by Pickens (2017) consider that too much data is collected, and that it is often unnecessary.

Core administrative datasets form the basis of all state agencies' record keeping, and business operations. These usually include legacy systems that have large volumes of (often dated) government data, which have been built on, and newer systems that have been added – sometimes a new data system every time a new service initiative is developed. However, integration is still at early stages and this limits the value that is able to be generated from New Zealand's data ecosystem (Data Futures Partnership, N.D. b). All the agencies covered by this inquiry have IT strategies and plans, (Department of Internal Affairs, ITDS survey, 2016) although most had been assessed as "needing development" for information management in their most recent PIF Reviews.

The Data Futures Partnership is an independent ministerial advisory group set up to provide a "collective voice on data issues and strengthen New Zealand's data ecosystem". It has been looking at issues around data sharing in New Zealand, particularly government data. It has stated that "the volume of data held by public and private sectors is growing exponentially and there is potential to create more value by effectively sharing and using many types of data" (Data Futures Partnership, N.D. a).

Although some government data systems are connected with each other, many are not. Those that are connected are mainly within a sector or across sectors by bi-lateral agreements for specific purposes (for example, IRD and MSD), rather than across multiple sectors and available for a range of purposes. Statistics New Zealand's Integrated Data Infrastructure (the IDI) is the notable exception, discussed in Box 7.3.

### **Box 7.3 Statistics New Zealand and the Integrated Data Infrastructure (IDI)**

The IDI is a large, secure, research database containing a wide range of de-identified data about people and households. It began with Statistics NZ survey data, involving the 2013 census and a range of other survey data. It now includes data from a range of other government agencies in the health, education, justice, students, social services and tax sectors, and some non-government organisations. These include Ministry of Education data on primary and secondary schools, tertiary education and some training programmes, IRD data on tax, income and working for families, MSD benefit data, Auckland City Mission data, ACC claims data and a number of data sets from the health sector.

The data in the IDI is "de-identified" at the point of use. It is identified when it enters the IDI (by the name and date of birth of the individual concerned). It is then "matched" with other data held in the IDI about the same person, then the identifying detail is stripped out and it becomes "de-identified".

Researchers can apply to use the data for specific projects at secure "data lab" locations. When the data is used by researchers it is always de-identified. This means that although it could in theory be used for operational purposes, it would only be for purposes where identification is not required (for example, to target services to cohorts of people with specified characteristics, but not to identified individuals).

Source: Statistics New Zealand, 2017b.

The Commission did not find any agencies that were using IDI linked data to measure their productivity, or to develop productivity measures. However, it can clearly be used to track long-term outcomes for groups of people, which suggests that there is potential to use IDI data as an input to productivity measurement (eg, quality adjustment of outputs).

Below is a scan of the sectors that are the focus of this inquiry, describing some activity they are engaging in to integrate their data and build better data systems.

### **The health sector: A semi-devolved and distributed sector is reflected in its IT supports**

Health sector data is kept in many different IT systems which are owned, governed and managed by different organisations across the sector. The largest collections are the National Minimum Dataset (NMDS) and the National Non-Admitted Patient Collection (NNPAC), which record hospital activity. These are held by DHBs, and shared with the Ministry of Health. Primary health information is kept in the Primary Health Care Organisation (PHO) enrolment data set. It is owned and controlled by PHOs and GP practices, and the data is shared with the DHBs and the Ministry of Health. Separate data sets also exist for disability services, mental health, well child services, immunisations, and for various specific conditions.

Although there is a lot of valuable data in the health system, it is not well connected, and there is no single organisation or group that owns, controls, governs or can access all the health sector information. Downs, (2017) comments that “To an outsider, the lack of primary care data that is collected and synthesised in New Zealand is surprising” (Downs, 2017, p. 25). Fraser and Nolan (2017) comment that greater use of existing data should be the focus rather than collecting new data, and that this [should] include “thinking about data access, standards and linking” (p. 5).

Some of the data from some of the national collections held by the Ministry of Health is entered into the IDI (including some NMDS data, and the PHO enrolments data). However, much of this data is limited in its content (such as the data from GP practices) and some of it is in the IDI on condition that its use is restricted.

### **The social support sector (MSD): Centralised and internally connected IT systems**

The Ministry of Social Development (MSD) also has a range of different data systems, which are all centrally owned, governed, and controlled. Like other organisations of its size and type, MSD has an ongoing work programme to consolidate and connect its systems as well as modernise and add new functionalities.

MSD has an Information Analysis Platform (IAP) which combines data from over 40 of MSD’s business systems since 1996, including the main benefit payment system, the case management systems, and a number of others. It essentially operates like an internal IDI, and is used by researchers inside MSD and the new Ministry for Vulnerable Children, Oranga Tamariki, for a number of purposes. These purposes include to contribute to their knowledge about programme effectiveness and to develop their outcomes models. MSD is using IAP data as an input to its Individualised Cost Allocation Model (iCAM) which has potential to measure productivity, discussed in Chapter 4.

### **The education sector**

As with the PHOs in the health sector, schools each own their own School Management Systems and provide data from them to the Ministry of Education. Meanwhile, the Ministry of Education runs a series of other data systems to manage all its own business (such as HR and Finance) and specific programmes (eg, specialist education). The school’s payroll service is managed and run by Education Payroll Ltd (EPL), a Crown Entity set up specifically for that purpose.

Education sector data that has been included in the IDI includes data for primary and secondary schools, tertiary education, industry training, and targeted training. This includes details of students enrolled in schools and tertiary institutions, and qualifications and standards achieved.

### **Justice sector data availability varies by agency**

Justice sector data is collected and maintained in a range of places. The Courts system, NZ Police and the Department of Corrections all provide some data to the Ministry of Justice. Courts system data is controlled

under the auspices of the judiciary, which meant that the Productivity Commission was unable to access it. However, police maintain comprehensive and rich data systems, which have potential to be useful for developing productivity measures. Administrative data collected via the police despatch systems, clearly has potential as an input to productivity measures, as can be seen in the Commission's Police productivity case study (Genet & Hayward, 2017).

The IDI includes justice sector data from the NZ Police on recorded crime (offenders and victims), Court charges data from the Ministry of Justice, and sentencing and remand data from the Department for Corrections.

### **Better connected systems would allow better information use**

There is clearly scope to make better use of existing data, including developing productivity measures; this will be demonstrated (for some agency data sets) in the case studies that the Commission has developed as part of this inquiry.

One cross agency initiative to try to address the ad-hoc nature of Government IT development is the appointment of the Chief Executive of the Department of Internal Affairs (DIA) as government's ICT functional lead, and Government Chief Information Officer (GCIO). A Government ICT Strategy was developed (in 2013, and revised in 2015). A work programme established in 2016 is reviewed and reported regularly, most recently in June 2017.

Like other Heads of Profession, the GCIO has set up several cross-agency stakeholder groups to gain agency support for the work programme. Goals include improving data analysis to inform decision-making, through the "development of infrastructure and capability to produce new insights to improve government services and to inform government policy and investment" and to develop the "Information, Technology and Digital Leadership, Capability and Workforce" (DIA, 2016).

A coordinated cross sector effort in IT development would be a positive move for productivity measurement. But this should be linked with business capability and include collaboration between the GCIO in the Department of Internal Affairs and the Functional Lead for Data and Analytics (the Government Statistician).

#### **F7.10**

Government agencies hold a wealth of data, some of which could be used to develop productivity measures, but it could be improved through better integration across agencies and with business re-engineering processes.

## **7.8 Data use rules, protocols and protections**

### **The system needs to build trust in how government will use data**

Data taken from administrative systems is often the best source for productivity measurement, as it is accurate and requires the lowest additional resource effort to produce. However, its use comes with some risk. In particular, there is a risk that the re-use of data collected as part of daily business is inconsistent with Principle 10 in the Privacy Act, which states: "An agency that holds personal information that was obtained in connection with one purpose shall not use the information for any other purpose".

Therefore, in addition to the hardware and software systems, there needs to be a body of rules, protocols and processes around how data is used. These frameworks and standards for the collection, storage, reporting and sharing of data are needed to protect privacy and ensure compliance with legal, ethical and human rights standards. In particular, public trust in government's use of data is important.

The Data Futures Partnership specifically addresses the challenge of ensuring trust. They noted that:

...existing models for enabling data integration and reuse fail because they do not address this central challenge. The dominant approach tends to build technically focussed point solutions that are highly specific to the particular context they are operating in. Moreover, data reuse interests tend to address only their own needs – frequently overlooking the interests of the data contributor. At best there is lip service to consent, minimal personal control for the contributor, or at worst coercive harvesting of data. Because these attempts fail at trust they become costly and hard to scale. (Mansell et al, N.D., p.7)

State sector leaders interviewed by Pickens (2017) also raised that collecting information from other organisations, including government agencies and third sector service providers, was sometimes difficult due to privacy concerns and suspicion as to how the information would be used.

## Protections for data use are still under development

The Data Commons Blueprint states that practices for safe management of personal and commercial creative and scientific knowledge, which balance personal and public interest and commercial and public goods, have “built up during the last hundred years” (Mansell, et al. N.D. p.15). However, they also say that these regimens are “woefully unfit” for what is happening in today’s world.

Various organisations in New Zealand are providing guidance to help agencies to navigate this new world. These include:

- *The Data Futures Partnership: A Path to Social Licence: Guidelines for Trusted Data Use*: Aim to enable organisations to maximise the value of data through building the trust of the clients and developing wider community acceptance. New Zealand organisations using the Guidelines will be able to follow advice about how to use, and communicate about data in a trusted way.
- *The Privacy Commissioner*: Provides comprehensive advice and guidance on his website, including an interactive “data safety toolkit” with tips on how to manage a privacy breach,
- *Government Chief Privacy Officer (GCPO)*: Based in the Department of Internal Affairs, the GCPO has issued core expectations for good practice for privacy management and governance in the State services. They have also developed guidance on privacy management, and a Privacy Maturity Assessment Framework to help agencies assess (and build) their existing capability.
- *The Data and Analytics functional lead*: This role is filled by the Government Statistician and Chief Executive of Statistics New Zealand. In addition to their capability building activities (described in section 7.4), other activities they are considering include:
  - **Good practice**: Setting data standards, such as data management practices.
  - **Knowledge curation**: Centralising knowledge and making it available to agencies.
  - **Data consultancy**: Services to support agencies with data challenges.

The goal is to help agencies to do “better things” with data – with Statistics New Zealand facilitating, connecting and brokering. They have established a group of agencies, and are working to consolidate standards and data management practices and to develop a roadmap for their work programme.

- *Social Investment Agency (SIA)*: Describe part of their role as “supporting agencies to make consistent and greater use of data and analysis”. The SIA is leading work to develop a data protection and use policy for the social sector; this is expected to include workshops and public engagement in early 2018.

A number of agencies are also taking proactive steps to address this issue (though often in response to negative publicity or data breaches). For example, the Ministry of Social Development had developed a Privacy, Human Rights and Ethics (PHRaE) Framework to ensure that they “act responsibly” when using personal information, and “remain focussed on achieving better outcomes for clients.

However, most of this work is still at very early stages, so it is not clear whether they will be successful – or whether enough is being done. In addition, much of the activity that the Commission has reviewed is being done inside government, and it is not clear how the public has been involved – or should be involved.

### F7.11

Better data integration, and the use of existing data that is routinely collected, would enable agencies to measure productivity. But work is needed to build public trust in government’s ability to integrate and use routinely collected data in ways that are safe.



## 7.9 Using the Performance Improvement Framework

There is scope to use the Performance Improvement Framework (PIF) to better support productivity measurement and improvement in agencies.

Departments receive periodic PIF reviews, which assess their current health and capability to meet future goals. The stated goal is to “support continuous performance improvement” (SSC 2015). These reviews have been successful, when they have been completed. However, no PIFs or follow up reviews have been published for social service agencies since 2014.<sup>18</sup>

Table 7.2 presents PIF ratings against selected organisational management measures for departments that are within scope of the inquiry, and for Statistics New Zealand, Treasury, and the State Services Commission (as they can contribute to government’s capability to measure the efficiency of public services).

**Table 7.2 Departments' PIF organisational management ratings (selected measures)**

Agency PIF	Leadership & workforce development rating	Information management rating	Efficiency rating	Financial management rating
Ministry of Education 2011	Needing development	Well placed	Needing development	Well placed
Ministry of Social Development 2011	Well placed	Needing development	Well placed	Well placed
Ministry of Health 2012	Needing development	Well placed	Needing development	Needing development
Department of Corrections 2012	Well placed	Needing development	Well placed	Well placed
New Zealand Police 2012	Needing development	Needing development	Needing development	Needing development
Ministry of Justice 2012	Needing development	Weak	Needing development	Well placed
State Services Commission 2013	Needing development	Needing development	Needing development	Needing development
Treasury 2014	Needing development	Needing development	Needing development	Well placed
Statistics New Zealand 2014	Needing development	Needing development	Weak	Needing development

Notes:

1. Agencies are rated on a four-point scale: strong, well placed, needing development, or weak. No departments were rated as “strong” on the measures reported in this table.
2. From 2013, the “Efficiency” rating changed to “Improving Efficiency and Effectiveness”. From 2016, it was dropped altogether.
3. Statistics New Zealand was also rated as needing development against its results measures “Deliver statistical products and services to meet customer needs” and “Manage data and information as a strategic asset”.

Although these are taken from the most recent PIFs for which there are ratings, these reviews were all undertaken more than five-years ago, and most agencies will have experienced some change since then. Follow up reviews, where they have been done, do not have ratings, and do not directly address all the issues raised in the PIF itself. For example, the Ministry of Justice PIF in 2012 had a “needs development” rating for efficiency, and noted that “to move to a well-placed rating, the Ministry will need to take a systematic and linked up approach to identifying the efficiency and effectiveness of its operating models”.

<sup>18</sup> A PIF has been conducted for the Ministry of Health, but it has not yet been released as at publication date

However, the follow up review in 2014, did not explicitly address focus on efficiency. Furthermore, as noted in Chapter 6, changes to the PIF model in 2016 saw the abandonment of explicit “efficiency” ratings.

The recent independent review of the PIF conducted by the School of Government found that there was significant value in and support for the PIF process. However, it also noted that while the PIF process was “seen as helping to improve clarity of an organisation’s purpose, [and] providing a clear and detailed understanding of what to focus on...actual improvements to management operations ... [were] less clear” (VUW, 2017, p. 6).

### F7.12

The Performance Improvement Framework could be used to better support agencies to measure and improve productivity.

## 7.10 Conclusion

Overall, the Commission has found that the state sector has many of the core business systems, processes and infrastructure, and some of the technical and data management skills and capability needed to consolidate input and output data into productivity measures and metrics. But the Commission has also found very few examples of productivity measurement techniques actually being applied or productivity data being analysed, across the state sector.

In summary:

- State sector agencies’ understanding of their business systems and flows is good in many places, patchy in others and “whole of system” understanding is poor.
- There are some gaps in technical data analytics capability, but it is possible to construct simple productivity measures and metrics without high levels of these skills. In many cases, simple measures should be enough to prompt the sorts of performance discussions outlined in Chapter 6. The Commission has conducted a number of case studies to explore measurement and data issues, as an aide to departmental staff developing productivity metrics (these are available in the online annexes to this report).
- Leadership capability and commitment to measure productivity need development.
- The core fiscal management system provides a solid basis from which to develop productivity measures and metrics. However, a greater focus on strategic financial management is needed, and better use of finance teams across the state sector to inform decisions and value identification.
- There is a lot of valuable data held by agencies across the state sector, much of it routinely-collected administrative data, which could potentially be used to develop productivity measures. This would be an even greater asset if the various data-sets were better integrated. However, more public engagement and better protections are required on the issue of government use and integration of routinely collected data, to build public trust and support.
- PIF reviews could be used to better support agencies to measure and improve their productivity.

## 8 Encouraging state sector productivity improvement

### Key points

- Productivity improvement emerges from changes that improve the way things are done. Agencies cannot improve their efficiency without a willingness to innovate or experiment. If government wishes to encourage innovation it will need to signal a tolerance for risk. Public sector leaders also need to champion innovation and a drive for more efficient and effective services.
- There is a general lack of good empirical evidence about what government agencies should focus on to improve their efficiency. But the wider literature suggests innovation in many areas can help drive productivity improvement, including human resource management practices, performance incentives and scale of operations.
- Devolved delivery of public services often leads to innovation and “pockets of excellence”. This is true regardless of whether the services are delivered publicly (eg, schools) or privately (eg, primary health care). In a well-functioning system, such excellence would spread. Yet the excellence often remains persistently isolated. The Commission wants to better understand what measures would assist in increasing the spread of successful innovation.
- The State Services Commission employs “business coaches” to help public agencies undertake business process improvement reviews. These approaches take a customer perspective to seek improvements in quality and efficiency. But there is limited information about their cost-effectiveness.
- Budget rules and processes are key drivers of government performance. New Zealand was an early leader in incorporating performance information into the budget cycle. But there is limited information about the extent to which efficiency measures are used to inform budget decisions. The budget process provides significant scrutiny of new/marginal expenditure, but does not provide good information on how effectively the base of existing expenditure is being used.
- In the private sector, competition drives market share towards more efficient firms. Over time, high-productivity firms grow, and low-productivity firms shrink and go out of business. These factors may explain as much as half of private sector productivity growth. But these effects are weak or absent in the public sector. Shifting users of public services from less productive to more productive providers and channels is a key challenge, which better data on productivity can promote.
- ICT investment is consistently identified as a driver of public sector efficiency. But the benefits of ICT investment will be limited unless accompanied by significant changes in skills, business processes, or organisational structures. These changes can be difficult where there are tightly specified input controls, such as commitments to a particular level of frontline staff. There is some “low-hanging fruit” where the use of modern communication methods would improve the efficiency of public services.
- The public sector can do a much better job of measuring and understanding the efficiency of its activity. This information also needs to be systematically used alongside other sources of performance information to deliver better outcomes for New Zealanders.

The terms of reference for this inquiry ask the Commission to provide advice on how to develop the capability, culture and systems that can support agencies to better measure, understand and improve productivity. Previous chapters have focused on measurement and building capability and systems to

support measurement; this chapter investigates systematic ways to *improve* the productivity of public services.

## 8.1 Sources of productivity improvement in the public sector

Despite its importance, the literature on measuring public sector productivity is relatively sparse (with some exceptions, see Chapter 2). There is also limited evidence on how to take action to improve efficiency. One literature review begins by noting that:

The evidence on the institutional drivers of efficiency in the public sector is surprisingly limited. Available evidence provides a very limited assessment of the impact on efficiency of varying the mix of inputs used [...] or of changing structural and managerial arrangements. (Van Dooren et al., 2008, p. 3)

Reviewing the literature, Kattel et al (2014) say:

What becomes evident [...] is that although the concept of productivity seems to have occupied a highly prominent place in contemporary public policy and administration rhetoric, it has not, for various political, conceptual and analytical reasons, really applied into the study and practice of public administration. (p. 18)

Because the productivity of public services is not systematically measured in most countries (except for the purposes of national accounts) there is little quantitative evidence linking particular actions or reforms in public services with evidence of changes in productivity.

One exception is discussed in Box 8.1 which finds a strong relationship between ICT expenditure and productivity.

### Box 8.1 Relationship between productivity and ICT expenditure, PPPs, and use of consultants

Dunleavy and Carrera measured the productivity of three UK public services over a decade (customs regulation, tax administration and the administration of social security payments), and correlate the productivity of these agencies with three possible sources of productivity improvement (in each case lagged for one year).

- *Spending on ICT.* The authors find a positive and relatively strong correlation between productivity and lagged ICT spending, noting “Taken on its own, the trend line here suggests that increased ICT spending alone could account for over half of the observed variance in productivity increases” (p. 276).
- *Public-private partnerships.* Private finance initiatives (PFIs, but more recognisable in New Zealand as a public-private partnership or PPP) were included by the authors as indicative “primarily of the large-scale business process modernizations that typically occur when new offices are opened or other facilities are relocated” (p. 275). They find a positive and reasonably close fit between productivity and lagged PPP spending, sufficient to explain 38% of productivity improvements (noting that these bivariable comparisons are not additive).
- *Consultancy spending.* The authors note that consultancy spending may also indicate large business reorganisations, or that it may indicate concern over the performance of public services. They find a weak and basically flat relationship between productivity and lagged consultancy spending. The authors do not distinguish between consultancy spending of different types.

The authors say that “[t]his differing result suggests that, primitive though this plotting exercise must be with the paucity of available data, there clearly are differences between the influence of the three independent variables here” (p. 277).

Source: Dunleavy & Carrera, 2013

One Australian review surveyed government agencies about barriers to achieving efficiency (Box 8.2).

**Box 8.2 Barriers to efficiency in Australian federal agencies**

A 2011 Australian review found that government's financial and budget frameworks can provide disincentives for promoting efficiency in the Australian government.

Agencies told the review that an absence of upfront investment in re-engineering government systems was a barrier. In particular, agencies reported an inability to invest in IT systems, and a reluctance to put forward proposals to do so in case they were asked to absorb the costs.

Agencies said that the annual appropriation process, and the processes for seeking ministerial approval to budget for an operating loss, deterred them from making upfront investment in systems designed to improve efficiency. They also reported a strong focus on balancing each budget annually, based on the belief that underspends would result in a reduction in future budget allocations. The annual budget process both discouraged agencies from seeking efficiencies, and made it hard for them to do so where it required investment that was not smooth over time.

Agencies said that the absence of a generally understood framework for risk management led to over-specification and inefficiency.

Agencies also noted barriers to working collaboratively with other agencies to reduce costs. Cultural barriers to working collaboratively were a part of this, but agencies also cited an incompatibility between a chief executive's legal obligations to be accountable for public expenditure, and a desire for more shared services.

Agencies disagreed on the extent to which the barriers reported were a result of financial and budgetary processes, or cultural practices that had built up around them.

Source: DoFD, 2011

Most evidence about sources of public sector productivity improvement comes from case studies or "grey literature". This chapter explores what eight reports<sup>19</sup> said about sources of productivity improvement, and discusses these sources in turn (Figure 8.1). This section of the draft report seeks evidence and comment on sources of productivity improvement in the public sector; the most important drivers of efficiency will be discussed in more detail in the final report.

Some areas are reported as sources of productivity improvement many times, including ICT investments, improving the capability of agencies to learn or innovate, human resource management practices, and the use of markets or competition. All of the areas identified from the reports in Figure 8.1 are discussed below.

<sup>19</sup> - Van Dooren et al. (2008) is an OECD-published literature review of the available empirical evidence on the institutional drivers of government efficiency, prepared as a contribution to the OECD's *Government at a Glance* publication.

- Simpson (2009), published in the *Journal of Economic Surveys*, mostly discusses issues arising in the measurement of productivity in the public services, but it also discusses studies that examine factors underlying productivity differences and productivity growth in public and private sector organisations.

- Dunleavy & Carrera's (2012) book *Growing the Productivity of Government Services* is a thorough guide to issues in public sector productivity, with a focus on practical issues in measurement.

- PWC (2013) discusses drivers of public sector productivity, and some specific examples of productivity initiatives, from an Australian perspective.

- Public Policy Forum (2014) is a report by a Canadian thinktank that outlined "some of the current best thinking on how we can advance an understanding of productivity in government in Canada" (p. i).

- DCLG (2015) is a report from a UK cross-agency Public Sector Efficiency Group that presented evidence on efficiency trends and drivers in a UK context, and insights for achieving further efficiency improvements.

- OECD (2017a) "provides a roadmap on how the OECD Public Governance Directorate can contribute to strengthening public sector productivity" (p. 1).

- McKinsey's (2017) Center for Government report argues there are opportunities to dramatically improve government productivity around the world, and that "governments need to adopt an ambitious, structured approach to transform the productivity of the state."

**Figure 8.1** Reported sources of productivity improvement in public services, selected literature

	Van Dooren et al, 2008	Simpson, 2009	Dunleavy & Carrera, 2012	PWC, 2013	Public Policy Forum, 2014	DCLG, 2015	OECD, 2017a	McKinsey, 2017
Internal drivers	Improved organisational learning, and innovation in service design		✓		✓	✓	✓	
	Human resource management practices	✓			✓	✓	✓	✓
External drivers	Performance budgeting			✓			✓	
	Automatic budget cuts, and hard budget constraints					✓	✓	
	Spending reviews						✓	✓
	Performance incentives		✓					✓
Organisational choices	Better regulation inside government						✓	
	Decentralisation to sub-national governments	✓						
	Scale of operations	✓						
	Markets and competition		✓		✓	✓		
Data and technology	ICT investment		✓		✓	✓	✓	✓
	Better procurement, and improved management of infrastructure and IT projects				✓			✓
	Data and targeting					✓		✓

## 8.2 Internal drivers

Internal drivers of efficiency relate to the way public services are organised and delivered by public agencies.

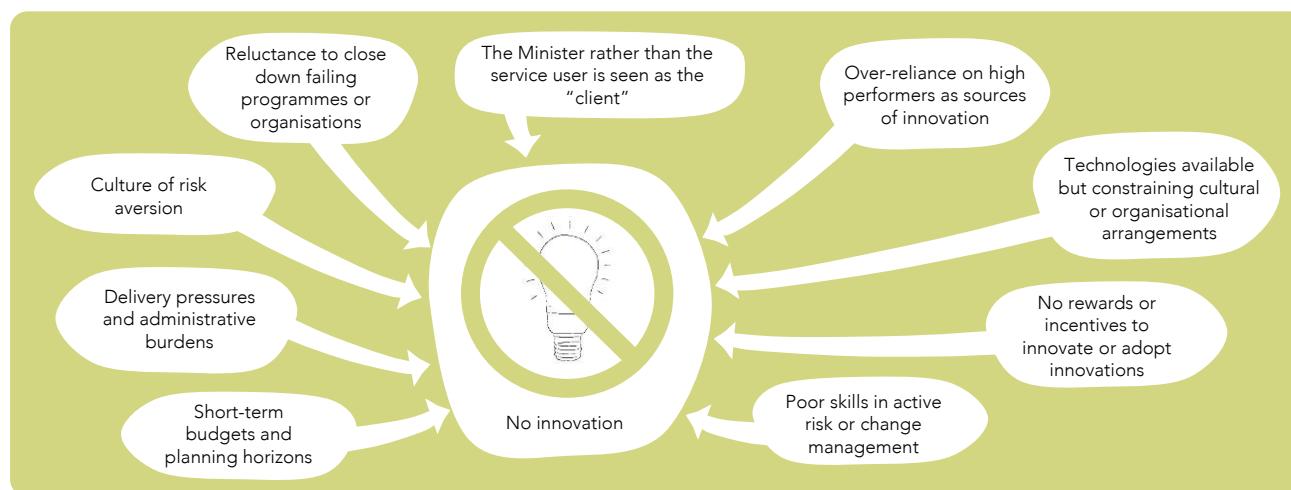
### Improved organisational learning/innovation

Even with the right incentives, agencies still need to learn from their own experiences to develop new, better ways of delivering public services. Agencies that are unable to learn or innovate will not respond to external incentives to improve efficiency, except through a reduction in the quality or quantity of public services delivered.

Political considerations can create risk aversion in the management of public services. From the perspective of public servants and politicians, the risk of innovation failing can outweigh the prospects that innovation will increase the efficiency or effectiveness of public services. In its inquiry into *More effective social services* the Commission found that:

The need for accountability and political risk management favours the use of prescriptive contracts, short contract periods and onerous reporting requirements. These factors work against the development and spread of innovation, and discourage productive and trusting relationships between government agencies and non-government providers. (NZPC, 2015b, p. 7)

**Figure 8.2 Barriers to innovation in the public sector**



Source: Mulgan & Albury, 2003; NZPC, 2015b.

In that report, the Commission discussed barriers to public sector innovation (Box 8.2). It found that:

A system that learns needs to have:

- clear goals to achieve better outcomes cost-effectively from social services;
- strong incentives to find, and the flexibility to try, new ways of doing things;
- information flows that provide ongoing feedback to clients, providers and commissioning organisations and citizens about what is working;
- a willingness to tolerate trials that fail (while dealing with failure quickly);
- the ability to structure trials and experiments in a way that can be scaled up if successful; and
- the flexibility to take up and spread successful innovations. (NZPC, 2015b, p. 13)

The Commission made a number of recommendations designed to promote a social services system that learns and is able to innovate. These recommendations were made in the context of social services, but have more general application to the other public services that are the focus of this inquiry. It recommended that:

- commissioning agencies should actively look for opportunities to engage providers to design and try out innovative service designs;
- commissioning agencies should improve the quality and transparency of information about service performance, and reward providers that innovate to improve their performance;
- commissioning agencies should promote and monitor the spread of innovation in devolved systems;
- to promptly respond to trends, commissioning agencies should use a wider range of data sources to monitor and evaluate service performance in real time;
- evaluations of service innovation should adhere to a set of principles for good evaluation developed by Superu; and

- services should be monitored and evaluated in a way that is commensurate to their scale and design, and government should not fund programmes where performance cannot be evaluated.

The Commission also noted the importance of government respecting the intellectual property of providers that successfully innovate, while respecting the right of government to own innovations that it funds. The Methodist Mission in its submission to this inquiry reiterated the need for this (sub. 1).

Since 2014, the OECD has had a project to identify key issues for innovation in government. It published an update report in 2017 (Box 8.3).

### Box 8.3 Global trends in government innovation

The OECD identified four cross-cutting factors for unlocking innovation in government:

- overcoming bureaucratic barriers;
- harnessing the power of citizens' ideas and the people behind them;
- building open, transparent and trust-based relationships with citizens; and
- enabling a culture that supports innovation.

The report makes four overarching recommendations for governments to maximise the potential for innovation:

- **Signal innovation as a priority.** [...] Innovation is unlikely to take root in government unless senior leaders communicate that it is an important priority. This includes acknowledgement that sometimes failure is an acceptable option. [...]
- **Enable connections across and beyond government.** This includes fostering communication and information sharing across and beyond government, by harmonising rules and developing connection points and networks. [...]
- **Promote trust through transparency and responsiveness.** [...] Governments can help build this trust by being open about activities and decisions that affect people. However, transparency is not enough. Citizen input must be considered and acted on, as appropriate, in visible ways.
- **Forge partnerships with all relevant players.** [...] Strategic and ongoing partnerships must be forged with civil society organisations, business, experts and the public. [...] Civil servants must therefore have the ability to balance and interpret the sometimes competing priorities of these different groups, and be empowered to make decisions on how to proceed with what they learn. (p. 15)

Source: OECD, 2017b

In the New Zealand context, the most pressing of these challenges are to signal that innovation in the delivery of public services is a priority, facilitating innovation in service design so that high-quality evaluation is built-in, and developing effective mechanisms for ceasing unsuccessful innovations and scaling successful ones.

### Business process improvement methodologies

The State Services Commission employs "business coaches" to work with agencies to redesign business processes (Box 8.4). They use techniques that have their roots in "lean" philosophy, which is based on the idea "that organisations can improve product or service quality for the same or less cost by continuously reviewing their processes from a customer perspective to remove waste by reducing duplication and inconsistency, and by identifying and resolving the root causes of operational problems" (NAO, 2011, p. 5).

### Box 8.4 Continuous Improvement



The SSC employs Continuous Improvement business coaches who work with public sector agencies to redesign business processes using “Systems Thinking” approaches that are similar to Lean, Vanguard, or Six Sigma methodologies.

SSC’s systems improvement coaches work alongside staff to understand customer interactions, how work flows through their system and where waste occurs. This information is used to challenge and change the way things are done to make improvements. (SSC, 2017a)

Coaches have worked with a range of public sector agencies including Land Information New Zealand and Education Payroll. Their work to redesign the Aviation Security, Immigration and Customs processes at Auckland International Airport won a 2015 IPANZ award for Excellence in Achieving Collective Impact.

SSC business coaches told the Commission that they typically find considerable scope for improvement when they work with public services:

#### **What we see**

All agencies’ people are committed and have the best of intentions, and many people are succeeding to deliver outstanding value despite the system

Typical themes we see include:

- Organisations traditionally arranged to suit themselves not the people they exist to serve
- Organisational purpose is rarely sharp or customer focussed
- Well meaning but detrimental improvement is everywhere – (replicating a paper process online)
- Poor measures or KPIs drive default purposes
  - can lead to worse performance and lower productivity
- Waste in all systems
- Limits of conventional management theory – e.g. functionalisation
- Cross agency ways of working are hard – developing the operating models to do this requires investing time, effort and everyone changing. (SSC, pers. comm.)

Education Payroll Ltd used the SSC coaches to review their business processes. Education Payroll reported increased customer satisfaction, decreased processing time, and fewer errors in processing. A \$60 000 investment in the SSC coaches had reduced costs by \$200 000 in one year (Education Payroll, pers. comm.).

Improving the customer experience and reducing waste are key objectives of the SSC’s Continuous Improvement process. Some public agencies that have used the SSC business coaches have subsequently employed their own coaches to continue this work.

A National Audit Office review of a similar business improvement approach in the UK customs system, found that:

- its success depends on how effectively leaders and managers apply the technique;
- the process did lead to productivity improvements, but their size could not be quantified because of other change programmes;
- staff engagement had improved only slightly, and was still well below the average for departments;
- the department did not understand the costs it had incurred in implementing the programme; and
- taking all these factors into account, “the programme is not yet value for money” (NAO, 2011, pp. 7-9).

Sari et al. (2017) review the application of lean in the Saskatchewan healthcare system. They say the cost of business process improvement programs in health systems are underreported and often underestimated, and conclude that outcome evaluation of such initiatives is required and should begin prior to implementation. But Radnor (2010) argues that any failure of business improvement processes like lean to be effective in public services stem from a lack of understanding or commitment on the part of the agency, rather than a failure of the lean philosophy.

Business improvement processes like this are attractive, especially because of the focus on improving the customer experience. Inefficiency in the public sector can often be characterised by poor user experience or long waits, and improving those experiences is desirable to improve quality and reduce costs.

But the Commission has not found good evidence about whether or not such approaches are cost-effective in public services. It is a problem if significant investment in business improvement processes relies on anecdote, rather than a more formal evaluation, to demonstrate its worth. As noted in Box 8.1, Dunleavy & Carrera find no relationship between lagged consultancy spending and productivity improvement. The Commission would like more evidence about “lean” and related business process improvement approaches.

### Q8.1

What evidence is there about the results of “lean” and related business improvement processes in public services? Are there evaluations about their use in the New Zealand public sector?

## Flexibility to manage

Giving managers and staff at different levels the authority to make changes or to innovate in response to new information appears to be as important as the agility and flexibility of the routines themselves.

Several studies have found that managers in the private sector not only have strong incentives to use performance information, but also have less restrictive internal processes. This means private sector managers have a greater capacity to respond to performance information than those in the public sector (Boyne & Chen, 2008, Hvidman & Anderson, 2014, Johansen & Zhu, 2014, Moyihan, 2008). Both the ability to change and the authority to change are needed.

Johansen, Taehee and Zhu (2016) studied differences in the use of performance information across private, public and not-for-profit hospitals in the US. They found managers in private hospitals were significantly more likely to use performance information to improve operational efficiency than managers in public hospitals. They concluded that private sector managers have both a stronger incentive to focus on service efficiency, and are less constrained by internal systems that weaken autonomy over subordinates and resource allocation.

Although the flexibility of managers to respond to performance information is critical to improving performance (Moynihan & Landuyt 2009; Moynihan & Pandey 2010), governments often focus on “creating increasingly advanced performance information systems while neglecting to increase managerial authority accordingly” (Nielsen, 2014, p.432).

Some organisations try to encourage better use of performance information through formal rules that require performance information to be collected and reported. Kroll (2013) argues that by increasing the regularity of performance information, such rules will make it more likely that managers will use the information when making decisions.

However, Moynihan and Lavertu (2012), reviewed the US Government Performance and Results Act (GPRA) and the Bush administration Program Assessment Rating Tool (PART). They found that rules compelling managers to collect performance information can result in managers acting “passively and not purposefully” – that is, they used data to create performance measures and set goals – but they did not use data for programme improvement, changing processes, reallocating resources, or managing employees.

Other research suggests that internal processes may have more relevance to the implementation of performance data. Van Dooren (2005) explored the use of performance information across government

departments in Belgium. He found that organisations that used performance information to guide decision making had frequent internal meetings to discuss performance results, which were attended by top managers, middle managers and operational staff.

### Devolved services

In some public services, the devolved delivery of public services can lead to innovation in service models. The Commission has heard that in devolved areas of the state sector (such as schools and hospitals), there is “no shortage” of organisations willing to innovate, and pockets of excellence – the problem is innovation often fails to spread. Particularly where these systems are subject to common ownership, or common governance, the lack of diffusion represents a system failure. For example, Downs (2017) reports how a range of models have emerged in primary health care in New Zealand:

Innovative changes to primary care delivery appear not to be driven by government policies per se. Rather, most initiatives are driven by local leaders who are inspired to change the way care is delivered. Some of these leaders have observed primary care systems in other countries and were motivated to implement delivery system reforms in their local New Zealand communities. The challenge for New Zealand policymakers is to create policy settings that encourage innovation whereby innovation is scaled nationally and not dependent on local leaders. (p. 43)

Despite the variety of primary care models that have emerged, little is known about their performance:

While there are exceptions, PHOs will readily admit that few of their programmes have been rigorously evaluated. Small PHOs note that they don’t invest in external evaluations whose cost would be disproportionately large relative to the cost of the programmes. Some PHOs have completed high-level internal examinations at how their programmes are performing. A small number of PHOs, such as Midlands Health Network, have hired external evaluators to assess programme performance. (p. 42)

Downs recommended these models be evaluated:

New Zealand is in an enviable position whereby a number of pilots have occurred throughout the country to address access to care. However, analysis has been limited. An updated and independent evaluation of what has been working in primary care should occur. That evaluation should hone in on the extent to which different initiatives have addressed the objectives of access, equity and integration. (p. 52)

The Commission agrees that such an evaluation is timely, and that it should also consider which models are the most efficient at delivering primary care services to their enrolled populations. But there are longstanding concerns about the lack of evaluation in the wider New Zealand public sector, and in particular, the link between agency activity and outcomes. In its inquiry into *More effective social services*, the Commission:

observed a large “stock” of existing social services that continue to be funded and run in much the same way over decades, with little evaluation of their impact or cost-effectiveness. At the same time, a flow of new initiatives attracts much attention but has little effect on the existing stock or on the performance of the system as a whole. This is consistent with an important inquiry finding that the current system is not good at evaluating programmes, or at expanding programmes that are effective and amending or phasing out programmes that are not. (2015b, p. 6)

The Commission recommended that “Superu should develop and adopt a set of principles for good evaluation and provide guidance to support those principles” (2015b, p. 200). Superu has done considerable work in this area, and published an evaluation guide (Superu, 2017).

But there are both demand and supply side reasons why evaluation is not done often or well in the New Zealand public sector (just as there are demand and supply side reasons for the absence of information about state sector productivity). In 1999, SSC said the reasons included:

- *low Ministerial demand for evaluation* – there is little in the way of incentives for Ministers to demand outcome evaluation; indeed there seem to be stronger incentives (in purchasing, and in budget arrangements, for example) for a short-term focus
- *historically poor outcome specification*
- *the lack of any central-agency champion for evaluation*

- *the lack of demand* for outcome information in current accountability frameworks;
- *funding difficulties*
- *methodological hurdles*
- *poor evaluation capability* in departments (mirrored in the private sector). (1999b, p. 11)

Many of these explanations remain plausible, although the state sector is more concerned with outcome information that it was in 1999. The Commission wants to understand more about the current state of evaluation in the state sector.

## Q8.2

What is the current state of evaluation in New Zealand public services? Is there demand for evaluation? Do departments have the capability and funding to undertake or purchase evaluation?

The slow spread of innovation may be a symptom of institutional problems such as:

- funding arrangements that limit an organisation's ability to retain cost savings achieved through innovation;
- few incentives on chief executives, principals and boards to share innovations with similar organisations;
- prescriptive central government service requirements that offer little scope for experimentation; and
- few avenues for communicating and demonstrating innovations.

Organisational factors can also hamper the spread of innovation, including:

- leaders without the vision and drive to seek out new ways of operating;
- organisational cultures that protect the status quo, or that are highly risk averse;
- a lack of resources and capability to experiment; and
- "staff politics" within an organisation (eg, some staff may fear innovation would bring a loss of status or power).

The characteristics of an innovation will also impact the rate at which it spreads. Innovations can be difficult to evaluate, and without "evidence" that the innovation works, other organisations may be reluctant to adopt it. For instance, a school (having limited resources) may evaluate an innovation using "soft" measures such as conversations with parents, teachers and pupils. Such qualitative evaluations may be enough for the school in question to judge the innovation's success or failure.

Yet soft measures may not be enough to convince other schools to trial the same innovation. For instance, the other schools may believe soft measures do not take account of differences between the student populations. In this instance, it is the lack of rigorous evaluation that acts as the barrier to the spread of innovation.

Furthermore, innovation is not costless or riskless. The more costly and risky an innovation is, the more important rigorous evaluation becomes.

The Commission is interested in gaining further insights into the level and spread of innovation within the state sector – particularly where decision-making power is devolved to local or regional operating units (such as schools and hospitals).

## Q8.3

In what areas of state sector is local/regional innovation occurring? Does innovation spread to other local/regional organisations? What factors explain the pace of innovation spread?

Some agencies have developed specific initiatives aimed at facilitating the identification and spread of innovation – for example, the Health Innovation Hub (HIH) aims to “accelerate smart ideas, products and services to improve health outcomes.” (HIH, n.d.). The HIH is a partnership between three large district health boards (Auckland, Counties Manukau and Canterbury), the Ministry of Business, Innovation and Employment, and the Ministry of Health.

It is difficult to know just how successful such initiatives have been. The Commission is interested in stakeholders’ views on the success of current initiatives and on other measures the government (or organisations) could take to promote the spread of innovation.

#### Q8.4

What measures exist to facilitate the spread of innovation? How successful have these measures been? What other measures or initiatives would assist in increasing the pace with which innovations spread?

Current and former New Zealand state sector leaders interviewed by Pickens (2017) reported that public sector pilot programmes did not work as intended. Criticism included a lack of proper assessment, a failure to end pilots when it became clear they did not meet their objectives or did so inefficiently, or an inability to scale pilots that were successful or showed promise.

#### F8.1

There are successful innovations in New Zealand’s public services, but they do not spread as far or fast as is desirable. This is a system failure.

### The current state: Management authority, flexibility and risk appetite

A cornerstone of the new public management frameworks brought in in the late 1980s was a shift in the role of ministers and public service departments to allow departmental chief executives greater decision-making authority and greater flexibility. They were freed up to make decisions over a whole range of aspects of their departments that were previously constrained, such as appointments and promotions, and were released from a range of detailed cash controls. Particular features were:

- a shift from control by ministers and control agencies (especially the Treasury and State Services Commission) of departmental inputs to the purchase of outputs;
- devolution of decision rights from the centre to departmental chief executives (with associated accountability and reporting requirements); and
- the move from cash to accrual accounting.

However, in several sectors (and it may be true more widely across the state sector) there remain some quite “heavy handed” and overly prescriptive settings. For example, the Commission’s inquiry into *New models of tertiary education* found that although government’s role in the tertiary education system significantly subsidises provision, it also limits flexibility and responsiveness in the system through prescriptive rules that dictate the nature, volume and location of delivery.

In addition, risk aversion consistently emerges as a barrier to innovation. This is particularly the case in terms of political risk, and there are some indications that this risk aversion may be increasing. Interviewing current and recent state sector leaders in New Zealand, Pickens (2017) reported:

Overall, interviewees felt government was too risk averse. Politicians and the public both tended to overreact to failure, even though failures were a small part of total activity and was necessary for innovation. This had meant state agencies were too reluctant to try new things in case those things failed, and to share information on failure even though it might be useful to others. (p. 27)

The *New models of tertiary education* inquiry also identified risk aversion as a barrier to innovation, with one submitter describing “the tremendous pressure to avoid public failure – whether real or perceived ... [as creating] an environment that is extremely risk averse and really only supports incremental change” (Ed. Collective, quoted in NZPC, 2017, p. 319).

Cultural norms in public sector organisations can be conservative and resistant to change. As the district health boards submitted:

DHB culture is highly patient-focused; to provide a more person-centred, effective, timely and safer care to their patients. There are however cultural barriers such as “we’ve always done it this way” mind set and underfunding of “soft skills” development. Sustained and aligned leadership is the key to changing this culture and DHBs do influence and instigate change over time. (sub. 17, p. 15)

**Q8.5**

How can government signal that it values innovation in the delivery of public services? Given some innovation will fail, how can it signal a greater tolerance for such risk?

**Q8.6**

How can Treasury and the State Services Commission support public sector agencies to be innovative and constantly seek improvements to the efficiency and effectiveness of service delivery?

## Human resource management practices

The production of public services is labour intensive, so it is not surprising that improvements to human resource management practices were consistently identified in reports as a key driver of public sector efficiency.

- Van Dooren et al (2008) found that employee satisfaction and morale were the most important drivers of performance. Soft aspects of HR (eg, skilled upper-level management, strong cultures that value employees and emphasise the importance of the agency’s work, and policies that empower employees) were more important than hard aspects (such as performance-related pay and job security).
- Canada’s Public Policy Forum (2014) argued that, more so than in the private sector, human capital is the key input to government productivity. It emphasised the need to focus on culture and talent, and having structures that allow the flexible deployment of human capital resources.
- The UK’s public sector efficiency group (DCLG, 2015) identified a number of drivers of efficiency, including holding down pay and incentive-based pay systems, though it also noted risks around recruitment and retention, quality, and perverse incentives from these approaches. It also noted shared services (back office functions) as a driver of efficiencies.
- The OECD (2017a) cautions that some countries that implemented public sector wage cuts or freezes in response to the global financial crisis were left with a lower skilled or demoralised workforce. It points to human resource management practices that focus on matching individuals to the right jobs, talent management, and flexible working environments as drivers of workplace productivity.
- McKinsey (2017) argues that to achieve a step-change in productivity will require government to “find or develop a range of functional skills that are currently underrepresented in the public sector – such as technologists, data analysts, and commercial project managers” (p. 27). It says that a cross-government view of talent recruitment, development and deployment is required to achieve this.

SSC reports that annual wage growth in the public sector has been below that of the private sector since September 2010, and that wage growth in the public service (core government departments) has been lower still. In the six-year period to June 2016, public sector wages increased by 8.8%, and private sector wages by 12.2% (SSC, 2016b).

The Council of Trade Unions submitted that employee engagement was an important driver of productivity improvement:

Constructive engagement between employers and unions contributes significantly to improved productivity. This approach has been adopted in several large firms. This includes the ‘High Performance High Engagement’ initiatives in Air New Zealand and KiwiRail, introduced with the support of unions. This involves redesigning work systems as part of a non-hierarchical and collaborative

approach to employee engagement. This includes design of appropriate and responsive output measurement to inform continuous improvement processes at the level of the work team and across the organisation. (sub. 9, pp. 9-10)

Similarly, the PSA submitted that “By engaging with workers at all levels, public sector agencies can better identify what outputs to measure as well as ways to improve productivity” (sub. 11, p. 2). Like the CTU, it recommends the adoption of High Performance – High Engagement workplace systems to build productivity (Box 8.5).

#### Box 8.5 High Performance Engagement

High Performance Engagement (or High Performance – High Engagement) means staff at all levels are involved in decisions. Air New Zealand says:

Over the past year we have worked with employees to increase their engagement and lift productivity. This includes creation of High Performance Engagement, a collaborative way of working where the people closest to the issues we face are actively involved in developing solutions. A High Performance Engagement charter was developed to set out clear objectives for Air New Zealand, our employees and their representatives. (Air New Zealand, 2015, p. 13)

KiwiRail reported in its 2016 Annual Report:

Central to our transformation programme is the leadership commitment to engage meaningfully with our frontline people, creating a culture which supports fairness, diversity, inclusion, safety and pride.

Some of our programmes which are changing the way we operate include a High Performance High Engagement approach. With the close involvement of union leaders and members, we are enabling direct and meaningful, involvement in problem solving and improving the way work is done, which raises productivity, improves safety and lifts our customer care. This is a significant commitment with our people and union partners to improve the company’s performance for our customers and shareholders. (KiwiRail, 2016, p. 14)

In 2017, the Ministry of Health made funding available to support the development of High Performance High Engagement systems between the Ministry, individual DHBs and unions (although “all of the workforce will be involved, regardless of whether individuals are union members or not” (MoE, 2017)). The DHB projects are at a very early stage.

The PSA submitted that current approaches to human resource management in the public sector were undermining the productivity and quality of public services in a number of ways:

The PSA and the State Services Commission have done extensive work on creating workplace cultures of integrity and productivity. These conditions are undermined when staffing numbers are capped at unsustainable levels and State servants work extensive unpaid hours to compensate for this.

The cap on public service employees has outlived what usefulness it had as a policy measure and continues to create unintended consequences, such as: professionals in client-facing roles spending a disproportionate amount of their time on administration and; the use of external contractors to make up the capacity and capability gap created by the cap. This additional capability is expensive and difficult to integrate into departmental workforce and system capability planning.

We do not see the increased use of fixed term agreements and other forms of insecure work, currently promoted in agency workforce plans, as an inevitable feature of the future of work but rather a choice by those agencies. Our understanding of integrity matters, both from the literature and from our practical experience, is that the increased use of insecure employment arrangements undermines integrity cultures and behaviours. (sub. 11, p. 12)

Gill’s survey of public servants in New Zealand, found that managers are motivated by feedback and support from peers, doing a job that is valued by the community, and receiving recognition for their work by ministers and senior management. Respondents in tiers 4 and 5 were more strongly motivated by the

opportunity for increased pay or promotion than those in tiers 1 to 3 (which were more motivated by peer recognition) (Gill, Kengmana & Laking, 2011).

There are also structural factors related to staffing. For example, while capping the numbers of Full Time Equivalent (FTE) staff in the core state sector may constrain expenditure and the growth of the state, it may also lead to more senior staff being retained (at higher cost) and reduced investment in training junior staff. There is some evidence of this in capability work undertaken by the Policy Project supporting the Head of the Policy Profession, which found an increase in Principal Policy Analysts across the public sector, but recruitment difficulties among Senior Policy Analysts, and a lack of investment in junior Graduate Analysts.

As noted by the PSA, political prescription of frontline staffing levels (eg, to employ more police or nurses) can also undermine efficiency by forcing frontline staff to undertake administrative tasks that might be delivered more efficiently by back office staff. This is discussed further in the next section.

### Q8.7

How can human resource management practices support public sector efficiency? What evidence is there about the effect of practices like High Performance Engagement on public sector efficiency?

## 8.3 External drivers

External drivers of efficiency typically relate to rules, rewards or sanctions that are applied to public agencies, often through the annual budget process.

### Performance budgeting

Performance budgeting refers to the use of performance information to inform resource allocation decisions in government. But it is a vague term that encompasses:

- the automatic (formula-based) allocation of resources based on results;
- providing performance information to inform budget decisions; and
- the provision of information about what performance outcomes are sought from budget allocations (including New Zealand's output-purchase approach to budgeting).

New Zealand was a leader at incorporating performance information into its budget process. The Public Finance Act 1989 reformed the entire framework of government financial management, reframing budget decisions from the level of inputs that government wishes to spend, to the level of services (outputs) that government wishes to purchase. The budget process is described in Chapter 6, which notes that efficiency measures are not commonly used as performance measures.

Curristine et al (2007) report that the use of performance information in budget processes has a number of advantages:

- providing a mechanism for politicians to clarify objectives, and set priorities;
- improving the monitoring of performance, including signalling which programmes are effective or ineffective;
- encouraging a greater emphasis on planning in management and budgeting, especially long-term planning;
- improving management, by encouraging managers to ask important questions about how programmes are performing;
- improving transparency about government's performance; and
- informing citizen's choices.



However, there is also scepticism about how much performance budgeting actually influences allocation decisions:

Performance budgeting has often been presented as a means of filling the “information gap” by providing a reporting mechanism between ministries of finance and line ministries/agencies covering inputs in the form of budget resources, institutional information, outputs and outcomes. The results, however, have been mixed in terms of providing the clear, useable and timely information for ministries of finance to make informed budget decisions, in particular with regard to prioritisation across spending areas. Ministries of finance have oftentimes found themselves overwhelmed with copious amounts of performance information of limited use to the budget process. (OECD, 2012, p. 21)

For example, Moynihan (2013) reports that the US Program Assistance Rating Tool (PART) mandated by President George W Bush was not widely used by Congress in making budget decisions, did not systematically influence budget decisions, and that it appeared to lose influence over time.

As a result, “most claims of performance measurement’s value in influencing decisions and improving services tend to be broad and disappointingly vague” (Ammons & Rivenbark, 2008, p. 305). Budget allocation decisions are deeply political. In 2003, Schick argued that:

Performance measurement has a problem ... The big problem is not in measurement, but in application. Much attention has been paid to the former, but not enough to the latter. With some notable exceptions, governments that invest in measuring performance rarely use the results in managing programmes. They do not base civil service salaries on performance, nor do they hold managers accountable for performance or allocate resources on this basis. Efforts to budget on the basis of performance almost always fail, as do reforms that aim to link pay and performance. It is common these days for governments to present performance information in budgets, annual reports and other official documents. But publication is a poor substitute for genuine use. (2003, p. 9)

Despite these disappointing reports, Moynihan & Beazley (2016) see a future for a more targeted form of performance budgeting, with information used less by budget officials as part of a budget cycle, and more by program managers in line ministries. This emphasises the vagueness of the term performance budgeting:

[I]n practice the line between performance budgeting and performance management is blurring in a way that reflects the limits of how performance data can be used in the budget process but also how they might be used for other management purposes. (2016, p. 13)

The Commission is interested in learning more about how information regarding the efficiency of public services is used to inform budget decisions.

### Q8.8

How is information about the efficiency of public services used to inform budget decisions in New Zealand? How does performance information affect Treasury advice about budget decisions?

## Automatic cuts

The UK’s public sector efficiency group (DCLG, 2015) identified the use of hard budget constraints, combined with spending flexibility, as a key driver of public sector efficiency, particularly over the short term. Hard limits provide public sector agencies with the incentive to prioritise, and spending flexibility gives them the freedom to reallocate spending or recombine inputs to improve their efficiency.

Under a policy of automatic cuts (also known as an “efficiency dividend”), the funding allocated to government agencies for existing programmes decreases by a set percentage each year. The idea is to withhold an amount roughly equivalent to productivity improvements observed in the market sector of the economy, on the basis that similar levels of productivity improvement have or can be achieved, even if they cannot be measured. Australia’s efficiency dividend provides an example of automatic budget cuts (Box 8.6).

### Box 8.6 Australia's efficiency dividend

The Australian government has applied an "efficiency dividend" – an annual reduction in funding to Commonwealth government agencies – since 1987. Most Australian state and territory governments also apply an efficiency dividend (Horne, 2012). The aim of the dividend is to:

- give managers the incentive to continually seek new or more efficient means of carrying out continuing government business
- allow government to redirect a portion of efficiency gains to higher priority activities
- clearly demonstrate Public Service efficiencies resulting from improvements in management and administrative practices. (DOFD, 2011, p. vii)

The dividend is based on the idea that as the public service becomes more productive, there is room for reducing inputs by an equivalent or lesser amount without reducing outputs. The dividend mimics the pressure that competition and profit motives provide in the private sector to motivate efficiencies.

The dividend generally applies to departmental expenditure (and in most cases not, in the New Zealand parlance, to non-departmental expenditure), and not to all agencies. In 2012, approximately 6.5% of Commonwealth expenditure was subject to the dividend (Horne, 2012). The rate of the dividend has varied between 1% and 1.5%, although there were additional one-off imposts of 2% in 2008-09 and 2.5% in 2012-13.

In 2011 the Australian Public Service Commission (APSC) reported a range of strategies employed by agencies in implementing the efficiency dividend including 'reducing travel, increasing use of video conferencing, closer monitoring of vacancies to determine if they need to be filled, delaying filling of vacancies, conducting accommodation reviews and reducing non-ongoing employment'. (Horne, 2012)

Arguments made in favour of the dividend include that:

- it is effective in returning efficiency savings because of its simple, predictable, and near-universal application;
- it assists in budget planning because of its predictability;
- it is efficient to administer;
- it provides an incentive for agencies to seek efficiencies;
- it does not impinge on the role of the agency CEO; and
- it is preferable to alternative approaches whereby the Department of Finance attempts to assess whether agencies are operating efficiently, without a good basis to form these judgements.

Arguments made against the dividend include that:

- it is a blunt instrument, disproportionately affecting agencies with a high degree of departmental expenditure;
- it disproportionately affects small agencies, which are less able to make trade-offs;
- it allows governments to avoid making funding decisions about specific programmes and activities;
- it can lead to agencies cutting staffing, with negative impacts on workload and outputs;
- the rate of the dividend is too high, and is unrelated to an assessment of efficiencies that can be realised;
- there is a lack of clear connection between the dividend and actual efficiency measures; and
- it can lead to gaming, whereby agencies seek funding for new proposals that is then used to offset the effect of the dividend. (Horne, 2012)

In 2011, the Australian government conducted a review of how best to promote efficiency in public sector on a continuing basis, and specifically consider the role of the efficiency dividend and potential alternatives to it (DoFD, 2011).

The report recommended that the efficiency dividend be retained in the short-term. It also recommended that over the longer term, more substantial changes to the Australian government's budget framework should be explored. A new framework should specifically target improvements in agencies' resource management, and provide better incentives including through gain-sharing arrangements. The report also found significant opportunities to gain efficiencies through making greater use of shared services, and by setting a long-term goal to reduce the number of federal agencies.

Automatic budget cuts have been implemented in other countries, particularly in Scandinavia. However, the OECD notes that where sensitive areas of spending are shielded from automatic cuts, this can undermine the overall rationale for such measures (OECD, 2017a). Efficiency adjusters, of varying levels, have also been a component of district health board funding models in the past.

Since the early 1990s, New Zealand public agencies have operated within fixed nominal baselines. Appropriations are fixed, and in general only change where a specific decision is taken to change appropriations. Agencies are expected to absorb price increases – they are, effectively, subject to an efficiency dividend equivalent to the rate of inflation, except where there are automatic adjusters (eg, district health board funding) or a decision is taken to spend more money.

This approach means that the government applies a high level of scrutiny to new spending. In order to receive funding for cost pressures, the burden of proof rests on agencies to demonstrate that it will not be able to deliver services effectively within existing funding levels. The government must then trade off increasing spending on existing programmes against any new policies. To help make these trade-offs within tight fiscal constraints, recent Budgets have moved towards having groupings of Ministers making trade-offs between themselves (e.g. the "Social Sector Ministers" group). (Bose et al, 2016, pp. 22-23)

In addition, there was a one-off reduction in most departments' baselines announced in Budget 2011. Following the GFC, the incoming government signalled that increases to agency budgets would be rare:

The minority National government imposed a cap on new spending of \$1.1 billion per annum (increasing by 2 per cent each year) from 2009, and signalled that most agencies would not receive additional funding for several years (New Zealand Treasury 2010). Chief executives were advised to innovate, prioritise and reorganise back office functions to deliver higher-quality services more efficiently, with a cap on non-frontline staff numbers and a diversion of resources to frontline services. Agencies were encouraged to amalgamate, use online platforms and achieve efficiencies in administration. (Brenton, 2016, p.138)

While this approach provides ministers with good information about the likely value of new investments – either new initiatives or increasing spending on existing programmes – it does not provide for much scrutiny of the bulk of public expenditure which is rolled over each year. The lack of information about the efficiency of these large baselines should be a significant concern for ministers.

The Four Year Plan process is intended to help fill this gap. Four year plans are designed to provide an integrated view of an agency or sector's medium-term strategy. They should outline the agency's strategic objectives and how it will meet them. SSC guidance says that:

The provision of Four Year Plans helps inform government resource allocation and decision-making (including Budget decision-making) by demonstrating the value created with existing baseline expenditure and resources; and by identifying the strategic choices, trade-offs and decisions, including around scaling and phasing of interventions, facing departments. They also enable the identification of system-wide opportunities, tensions and risks.

The main question that a department's medium-term planning should answer is "how will the department create increasing value for its customers and New Zealanders over the medium-term with the funding and balance sheet it has available?" (SSC, 2016c, p. 6)

The Commission is interested in views on how effectively four-year plans achieve the objective of providing assurance to ministers about how well departments are using their baseline funding to "achieve their strategy and the Government's priorities" (SSC, 2016c, p. 7).

**Q8.9**

What are the lessons from New Zealand's experience of fixed nominal baselines?

**Q8.10**

How well does the Four Year Plan process provide information about the value created with existing baseline expenditure and resources? How well does the Four Year Plan process identify how departments will increase the value they create over the medium-term?

## Spending reviews

In more recent documents, spending reviews are discussed as tools to improve public sector efficiency, reflecting the role that such reviews played in helping governments manage during and after the GFC. Spending reviews typically address issues of effectiveness and cost-effectiveness, as well as allocative and technical efficiency. They are a way to scrutinise the base of expenditure, and so complement the scrutiny that annual budget processes provide over marginal expenditure.

The failure of conventional budget preparation processes to fully address expenditure prioritization is not accidental. It is no easy matter for central decision-makers, whether in MOF [ministries of finance] or at the level of political leadership, to reallocate resources. To do so requires considerable information about the efficiency and effectiveness of baseline spending, and involves overcoming resistance from spending ministries and ministers. (Robinson, 2013, p. 7)

Spending reviews take many forms. Robinson identifies three main types:

- programme reviews, that examine specific government programmes; these can address issues of allocative and technical efficiency;
- process reviews, that scrutinise specific business processes used in the production of government services, such as procurement processes, IT systems, or HR practices; these are typically focused on technical efficiency; and
- agency reviews, which in principle cover all of an agency's programs and processes.

Comprehensive spending reviews are more successful during times of significant change or crisis, such as that experienced by many governments during and following the GFC. At other times, selective reviews that allow for more thorough scrutiny and a gradual reallocation of resources are likely to be more effective (Robinson, 2013). The OECD has identified four pre-conditions for effective spending review processes:

- political will is necessary to establish the reviews and take decisions about savings;
- review teams require the political and technical capacity to produce useful recommendations;
- performance information generated through the budget system should provide information able to support the review process; and
- because reviews designed to produce efficiency improvements may require time to deliver savings, links to a medium-term framework are important (OECD, 2012).

Seven baseline or output price reviews of New Zealand government departments were conducted between 1996 and 2004. However, these had varying results – most led to increases in funding levels, but few agencies actually resolved their pricing and resourcing problems – and none of these exercises seem to have

taken place in the thirteen years since 2004 (although Value for Money and line by line reviews have continued to be done).

**Table 8.1 Baseline and output price reviews, 1996-2004**

Department	Motivation	Result
Police (1996)	Cost and demand pressures.	No increase. Review found scope to achieve savings and deficiencies in police management information systems.
Statistics New Zealand (1996/7)	Statistics New Zealand had been unable to deliver third-party revenue targets, impacting on working capital.	No increase, but reduction in third party revenue targets.
Ministry of Foreign Affairs and Trade (1997)	Rising input costs and an inability to further reprioritise resources, resulting in a projected operating deficit.	An increase in some appropriations, but a decline in others after transfer of some Ministry property assets to Crown balance sheet. Net increase over time.
Education Review Office (1998)	Concern that significant remuneration changes for teachers (who formed ERO's main labour market) would put pressure on output prices.	Increase to baselines.
Department of Corrections (2002/3)	Significant cost increases and concern about impacts on output delivery.	Increase to baselines.
Department of Child Youth and Family Services (2002/3)	Funding sought outside budget cycle. Rising workloads and high-profile child deaths.	Additional funding provided for a plan to stabilise the department. Later it was merged with the Ministry of Social Development, and it has now been re-established as a separate Ministry.
Ministry of Justice (2004)	Newly-merged agency (the Department for Courts and Ministry of Justice) and increased demand from new legislation, rising population and increased crime resolution rates.	Increase to baselines.

Source: Norman, 1998; Petrie & Webber, 2001; Brough, 2005; Whitcombe, 2008

As part of the 2009 budget process, New Zealand departments were instructed to undertake "line by line reviews" of expenditure. Reviews were expected to identify:

- opportunities for savings from programmes that were ineffective, offered poor value-for-money, or were inconsistent with government priorities;
- programmes where further investigation was needed because of poor information about performance;
- ways to save on back office functions, including human resources, IT and communications; and
- areas where scope creep had resulted in agencies undertaking work that was not their core business, or which duplicated the work of other departments.

This was an internal process between departments and Treasury. As a result of the process some \$500 million in annual savings was identified.

**Q8.11**

What evidence is there on the effect on productivity of “line by line” reviews undertaken as part of Budget 2009? Should there be more regular scrutiny of base expenditure to identify efficiency savings? If so, what form could this process take?

## Performance incentives

There are a range of ways in which government might use incentives to motivate public sector performance

**Table 8.2 Potential mechanisms to motivate performance**

Mechanism	Reward	Sanction
Funding	Increased funding to the agency	Reduce or restrict agency funding
	Maintain funding levels	Eliminate agency funding
	Increase the staff budget	Cut the staff budget
	Provide management and employee bonuses	
Flexibility	Allow the agency to retain and carry over efficiency gains	Return all funding to the centre
	Allow flexibility to transfer funds between different programmes and/or operating expenditures	Restrict the ability to transfer funds
	Exempt the agency from certain reporting requirements	Increased reporting requirements Order a management audit of the agency
Public recognition	Publicly recognise the agency’s achievements	Publicly criticise the agency’s performance

Source: Curristine et al, 2007

Curristine et al. (2007) note, however, that few countries actually use performance results to financially reward or punish agencies. They note several reasons:

- a lack of political support for such responses;
- rewarding well-performing agencies with additional funding may not take into account government priorities;
- uncertainty in attributing performance outcomes to agency actions; and
- to avoid encouraging gaming or the manipulation of data.

The latter risk is illustrated by the period of “targets and terror” in the English National Health Service (NHS) (Box 8.7).

### Box 8.7 “Targets and terror” in the NHS

In the late 1990s, the Blair Labour government introduced a National Performance Assessment System, which measured the performance of health care providers (eg, hospitals and primary care organisations) against a range of indicators covering issues such as access, health improvement, health outcomes, patient/carer experiences, efficiency and effective delivery. Health care providers reported against indicators, and were awarded “star” ratings based on their performance. Although the performance measures were not strongly linked to funding, poor star ratings had consequences for managers – “health managers whose organizations lost their stars or never gained any could expected to be fired” (Hood, 2006, p.515).

Reported performance data showed impressive improvements against a number of metrics and English health providers (which were subject to the targets) improved at a much faster rate than those in neighbouring Wales (where targets were not applied) (Hauck and Street, 2007). For example, the proportion of patients waiting more than four hours in accident and emergency departments fell from 23% in 2002 to just over 5% in mid-2004. Le Grand (2007) attributed the poorer performance in Wales in part to the authorities there misdiagnosing the motivations of health providers and inappropriately applying a high-trust model.

However, further investigation revealed that while there were genuine performance improvements, there were also significant levels of gaming and misrepresentation:

- A study by the Commission for Health Improvement found that in a third of ambulance organisations, information on ambulance response times to emergency calls had been “corrected” to be reported to be less than eight minutes” (the official measure). (Bevan & Hood, 2006, p.531). The same study found that some patients in urgent need had been classified as having a lower priority for ambulance response (and hence falling outside the monitored “category A” group).
- Patients were required to wait in queues of ambulances outside A&E departments “until the hospital in question was confident that that patient could be seen within four hours.” (Bevan & Hood, 2006, p.531)
- In response to the target that a patient had to be admitted to a hospital bed within 12 hours of emergency admission, many hospitals turned “trolleys” into “beds” by putting them in hallways.
- The National Audit Office found that some health organisations had inappropriately adjusted their waiting list to meet their waiting time targets.
- Pressure to meet emergency department targets was also identified as a contributing factor to serious care failures at Stafford hospital (Francis, 2010).

## Performance-based pay

Performance-based pay is often used in the private sector with the aim of motivating employees to perform at higher levels, and its use is controversial in the public sector. Proponents claim that performance-based pay increases motivation to improve performance, and may also encourage higher quality staff into public employment.

Opponents argue that public sector staff are already fully or sufficiently motivated by intrinsic motivations and that performance pay could weaken intrinsic motivation; that performance pay tends to undermine unions and collective bargaining; that the desired performance is difficult or impossible to specify, measure, or attribute wholly to the work of the employee; or that the rewards may be discretionarily awarded on an unfair basis.

There is a lot of literature investigating the effects of performance-related pay in the public sector, and many literature reviews. For example, Hasnain et al (2012) find that:

- a majority of studies find a positive effect of performance-related pay, with higher quality studies generally more positive in their findings;
- the more rigorous studies relate to jobs such as teaching, health care and revenue collection, where outputs or outcomes are more readily observable; they find there is insufficient evidence of the effects of performance related pay in areas where outputs or outcomes are more difficult to measure;
- several studies identify problems with unintended consequences in gaming, though it is not clear if the gaming results in a decline in productivity compared to the counterfactual; and



- there are few studies that follow up performance-related pay over a significant period of time.

The Work Foundation (2014) finds widespread evidence of positive effects from performance-related pay in education, though much depends on the design of the scheme, teachers' support, and wider accountability and performance measures.

The evidence in health is also positive, but more so in primary care and the management of chronic diseases than in acute care.

The evidence in the civil service is slightly positive but more mixed. It finds evidence of gaming or outright cheating, some evidence of effort displacement on unrewarded tasks, but also some evidence of positive spillovers. The review also looks at effects on staff attitudes and motivation.

In the civil service, there is some evidence that financial incentives can displace intrinsic motivation; in education several studies find a negative effect on the job satisfaction and motivation of teachers; and the evidence in health is mixed.

### Q8.12

In what circumstances might performance related pay be effective? Where is it likely to be ineffective or harmful, and why?

## Incentives on agencies

Some theories of bureaucracy suggest (or assume) that public servants have an incentive to maximise their budgets (eg, Tullock, 1976; Niskanen, 1973). Larger budgets can bring tangible and intangible rewards to public servants, including higher status. Pickens (2017), reports that current and former senior public servants in New Zealand reported that increasing budgets and organisational size was associated with professional status. Writing of the UK, Barber says:

[I]ncentives in the public sector have traditionally been, to put it mildly, ambiguous. Ministers' performance tends to be judged on the size of the budget they negotiate with the Treasury, rather than how much they deliver or innovate. Civil servants tend to get bigger rewards and more status for managing more people or bigger budgets. Departments don't call time on budgets that aren't delivering. They rarely, if ever, go to the Treasury with ideas for spending less, even if, through innovation, doing so could improve outcomes. Some years ago, when I volunteered to return, in mid-year, part of a budget I had realised I wouldn't need, the relevant Treasury officials almost fell off their chairs in surprise – no-one ever did that, they commented. (2017, p. 3)

Similarly, bureaucracies may face incentives not to seek savings or efficiencies for a range of reasons, including well-meaning stewardship of their agency. Unspent resources might disappear from future budgets, inhibiting the ability of the agency to handle unexpected pressures. Because of intrinsic motivation, public servants might be reluctant to shift resources from programmes that are doing some good, even if the resources could be used more efficiently in other parts of government. Box 8.2 notes these effects as a barrier to efficiencies in the Australian public sector. Individual ministers may also be subject to the same incentives to retain expenditure within their Vote.

The Justice Sector Fund (Box 8.8) is one way of trying to overcome these problems.

### Box 8.8 The Justice Sector Fund

The Justice Sector Fund is a funding pool established in 2012 to allow the transfer of funding between justice sector agencies, and across years, so that savings in one part of the justice sector can be redirected to fund priorities within the sector.

By May 2017, a total of \$273 million of savings in the justice sector had been made available to fund 66 initiatives. These initiatives include a review of family violence law, expanding the use of restorative justice, reintegration programmes for people released from prison, and installing audio-visual links between courts and prisons.



Applications to the Justice Sector Fund have to show that an initiative will contribute towards reducing crime and reoffending, or assist the sector to modernise and become more effective. The fund provides a mechanism to trial new initiatives that may be funded in future Budgets.

Source: Ministry of Justice, 2017.

The Justice Sector Fund should increase the incentive on agencies to seek efficiencies and make savings available for reinvestment in other justice sector initiatives. The Commission is interested in understanding more about the incentives on public sector agencies to seek efficiencies.

### Q8.13

What evidence is there about incentives on public sector agencies to seek efficiencies? What evidence is there that initiatives like the Justice Sector Fund increase agencies' incentives to seek efficiencies? What has been its effect on the overall efficiency of the justice system?

## 8.4 Organisational choices

Organisational choices refer to the decisions that are made about the way public services are structured and governed.

### Better regulation inside government

Government regulation of the private sector plays an important role in managing risk and protecting the public. Over-regulation can constrain innovation and harm productivity. Government also regulates its own activities, through a range of laws, rules, norms and myths<sup>20</sup>. The OECD (2017a) discusses the possibility that administrative regulation inside government may harm the productivity of government services.

Public service delivery, e.g. in education, health, and transport, can be improved by addressing constraints and burdens imposed by superfluous, outdated or badly targeted regulation, and burdensome information obligations. Thus, public sector productivity can be significantly affected by the quality of RIG [regulation inside government] that will lead to improved targeting and efficiency.

This is an underresearched area, where evidence is even more scanty than for the private sector in general. (p. 29)

The New Zealand Initiative submitted that:

Small changes in the quality of proposed regulation can have enormous consequence for overall national productivity: simply imagine if the Health & Safety regime had been better designed to avoid the kind of excessive risk aversion it created, or if the Anti Money Laundering / Know Your Customers regulations had taken more care to scale costs imposed to degree of real risk involved in private sector activity. Relatively small amounts of public sector expenditure in better consultation and regulatory drafting could have had enormous consequence for public sector productivity. (sub. 8, p. 7)

Professor Zoe Radnor of the University of Leicester provided the Commission an anecdote from a British hospital. Staff at the hospital who cycled to work were required to sign a "bicycle book" at the front desk; Professor Radnor wanted to find out why. She discovered hundreds of these books, boxed up and stored for decades, dating back to World War II, when because of fuel rationing, staff who cycled were entitled to extra food rations.

The Commission is interested in whether regulation inside government constrains public sector productivity in New Zealand.

<sup>20</sup> In 2015 a government taskforce on "Loopy rules" (DIA, 2015) reported to its surprise that a large number of loopy rules complained about by the public were in fact myths.

**Q8.14**

Is regulation inside government a barrier to improving the productivity of public services? What examples or evidence is there?

## Decentralisation to sub-national governments

Van Dooren et al (2008) report evidence that decentralisation (the devolution of functional responsibilities to sub-national governments) is associated with efficiency gains. However, the supporting evidence cited appears to describe a better matching of services and taxation levels to local preferences, rather than improvements in technical efficiency. The Commission provided a framework for allocating functions between central and local government in its 2013 inquiry *Towards better local regulation*. As a small, unitary state, there are limited opportunities to decentralise state services to sub-national governments in New Zealand; therefore, this issue is not considered further. However, the next section asks whether there are issues of insufficient scale in the organisation of some New Zealand public services.

## Scale of operations

Van Dooren et al (2008) report efficiency gains from increasing the scale of operations, based on evidence collected mainly in the health and education sectors. For example, McCallion et al (1999) used DEA analysis to investigate efficiency differences between small and large hospitals in Northern Ireland, and found support for the then policy to rationalise hospitals in Northern Ireland. However, the OECD noted that other public service values, such as equity, access and quality also need to be taken into account when considering options to leverage scale.

The organisation of New Zealand's health and education system invites questions of whether there are opportunities to improve the efficiency of these services through increased scale. There are 20 DHBs, 32 PHOs, about 2500 schools and more than 5000 playcentres or licenced ECE services. A recent report on primary health care in New Zealand recommended consolidation of PHOs and DHBs:

Risk can be better managed among large organisations that include a large number of enrolees. Some of New Zealand's 31 PHOs do not have enough scale to take on expected and unexpected risks that are more easily spread across large populations. Also, the large number of PHOs combined with 20 DHBs has resulted in a vast array of complex and overlapping relationships. Consolidation of the sector to four to six PHOs and a similar number of DHBs and the establishment of similar geographic boundaries could greatly simplify the current system. Amalgamation could streamline data sharing, support collaboration between the sectors, reduce administration, increase accountability and simplify shared risk. (Downs, 2017, p. 6)

In the case of pharmacy services, government regulations on the ownership of pharmacies have limited scale, with a consequential lack of competition, innovation, and efficiency (Box 8.9).

### Box 8.9 Pharmacy ownership restrictions

In order to receive a pharmacy licence, a pharmacy must be majority-owned by a pharmacist, and a pharmacist may own a majority stake in no more than five pharmacies. Prescribers are prohibited from having an interest in a pharmacy.

The Ministry of Health considers that these restrictions:

- add unnecessary compliance costs for license applicants and the regulator;
- reduce competition;
- reduce investment in pharmacy;
- restrict innovative ownership models (eg, community trusts);
- act as a barrier to the integration of health services to develop better patient-centred care.

The Ministry considers that managing the risks associated with the supply of medicines to the public is best done through licencing conditions, rather than ownership restrictions. Past attempts to liberalise pharmacy ownership requirements have been successfully opposed by the Pharmaceutical Guild and Pharmaceutical Society.

Source: Ministry of Health, 2016

### Q8.15

Are there issues of insufficient scale in the organisation of health or education services in New Zealand? What evidence is there of this?

## Markets and competition

Competitive markets are a key driver of productivity growth in the private sector.<sup>21</sup> First, competition drives market share towards more efficient firms. Therefore, in a competitive environment, high-productivity firms are more likely to grow and expand while low-productivity firms shrink and exit (opening up room for more efficient producers). Second, competition encourages the uptake of costly productivity-raising actions by firms that may not otherwise take place. For example, because adopting new technology is costly, producers facing less competition have less incentive to do so. Put differently, if a firm is making good profits in a low-competition environment, there is a higher opportunity cost to changing production practices via new technology.

These effects are weak or absent in the public sector. In general, public services do not face competitive pressures to increase efficiency, though a public service ethos may encourage efficiency. Most public services continue to be delivered by government agencies that are not at risk of losing customers to competitors, or going out of business because they are outcompeted. In rare circumstances, government might respond to perceived poor performance by transferring functions between agencies, or establishing new vehicles to deliver services. In some public services where the public can exercise choice of provider, government may regulate to reduce or stop reallocation for other policy reasons (eg, through school zones). In its inquiry into *New models of tertiary education*, the Commission (2017) found that the only realistic way for private tertiary establishments (PTEs) to increase their market share was to buy another PTE and hope to retain its TEC-allocated student places.

DCLG (2015) describes four ways in which markets and competition can aim to bring “private sector-type disciplines and incentives to public services”:

- **New entry competition/market creation:** Creating competition between public sector delivery units, either by allowing the entry of private sector providers, or giving consumers greater choice in where they receive their services; however, it notes that a new unit may be less effective in the short term, and this driver depends on units being able to grow or reduce capacity in response to changes in demand.
- **Strengthened incentives:** Giving service providers the motivation or incentives to deliver higher quality and/or lower cost services, including through financial and non-financial incentives and deterrents; though careful design is necessary to prevent unintended consequences.
- **Cost benchmarking:** bringing the worst performing units up to a minimum standard, by holding them to account (including through naming and shaming the worst performing units).
- **Intelligent outsourcing:** Commissioning private firms or not-for-profits can bring external experience and perspectives, and realise the benefits of competition, if commissioning is well-designed. (DCLG, 2017, p. 89)

<sup>21</sup> However, these effects are not as strong in the New Zealand private sector as they are in other countries, contributing to low productivity growth rates (Conway, 2016).

The selective application of markets and competition can also improve quality of service and increase satisfaction from the users of public services. In its inquiry into *More effective social services*, the Commission (2015b) recommended that government should trial providing service users with greater control via client-directed budgets, and consider a wider suite of commissioning models in contracting for services. The New Zealand Initiative submitted that government's investment approach could replicate some of the benefits of market competition in driving innovation:

NGOs operating on outcome (or output)-based contracts allow a wide variety of approaches for delivering those outcomes. More efficient providers can expand to serve more of the market; less efficient ones win fewer contracts and provide fewer services. The Investment Approach, broadly speaking, then facilitates an entrepreneurial market discovery process similar to that which operates in other competitive markets. Even better, the government may be able to learn from the process to improve its delivery of similar outputs or outcomes by Ministries. (sub. 8, p. 2)

New Zealand Kindergartens *Te Pūtahi Kura Puhou o Aotearoa* submitted that the existence of for-profit provision of early childhood education represented the diversion of public funds for personal gain, and away from children's learning (sub. 10). However, that judgement should depend on the quality of outcomes delivered in terms of *Te Whāriki* (the early childhood curriculum), as well as other goals such as supporting the ability of parents to participate in the workforce.

One recent example of competition is the introduction of hospital choice in the English National Health Service (Box 8.10).

#### Box 8.10 **Hospital competition in the English NHS**

In 2006 the British government introduced market-based reforms into the English NHS. Every patient in England could choose their hospital for secondary care, and hospitals had to compete for revenue by attracting patients. Prices were fixed, and hospitals competed to attract patients based on quality of care. To facilitate this, the government introduced a new information and booking system that allowed patients or their GP to book hospital appointments online or over the phone (the "choose and book system").

Cooper et al. (2010) found that hospitals in more competitive geographic markets saw mortality decrease more quickly than in hospitals in less competitive markets; prior to the policy there had been no difference in mortality trends. The authors conclude that hospital competition in the English NHS saved lives. They also showed (2012) that competition between public providers caused public hospitals to improve their productivity by reducing length of stay; by contrast, competition from private hospitals did not improve public hospital performance, and instead left them with a costlier case mix of patients and an increased length of stay.

Using a similar methodology, Gaynor et al (2010) finds that hospitals in areas of greater competition saw lower death rates, and shorter lengths of stay, without increasing operating costs or cutting staff.

Cookson et al. (2013) examined whether this competition had undermined equity of access to care, by looking at access rates from low socioeconomic communities. They find no evidence that increased competition had a harmful effect on equity of access. "If anything, we find that competition may have very slightly improved socio-economic equity, by helping to facilitate the slightly more rapid growth of elective inpatient admissions over time in deprived areas" (p. 419).

Dunleavy and Carrera (2013) recommend that public sector efforts should concentrate on "strengthening provider succession" to shift users of public services from less productive to more productive providers:

What is needed is a genuine analogue to the competition and succession processes (i.e., the transfers of output and activity from less efficient to more efficient firms) that account for half of all productivity advances in the private sector. A wide range of different NPM [new public management] solutions have only transferred across superficial business processes in rather crude ways, without transferring risks.

And constantly fiddling with the location of the interface between the public and private sectors, or promoting quasi-markets, has been a waste of time in terms of generating more genuine provider succession. This service architecture tinkering mainly just displaces the nature of productivity growth problems. In most public services the key to securing more provider succession is to refocus on substantive changes of services, fostering innovations that can bring in new providers plus solutions. For instance, don't try to run the existing local public libraries system more cheaply. Set up a wholly e-book national lending library in competition, and then see how citizens really want to read books, and what books they want to read. (pp. 299-300)

Provider succession can occur in a market where there are multiple providers, customer choice, and government is prepared to let providers grow, shrink, and exit. However, where there are monopoly providers of public services, providers succession can be particularly difficult to achieve. The reasons include:

- *Legal considerations.* Unlike the private sector, government's procurement decisions can be challenged via judicial review. In 2013 the Ministry of Health issued a Request for Proposals for regional and national problem gambling services. The incumbent provider tendered and was unsuccessful, but successfully challenged the tender decisions in 2015.
- *Political and public lobbying.* The public has a greater interest in the awarding of public sector contracts, and their views can influence outcomes. In 2006 the Ministry of Health ended public funding for a well-child phone service operated by Plunket (PlunketLine) in favour of an American-owned firm, McKesson, which already operated Healthline and successfully tendered for an integrated service. PlunketLine had not been meeting the terms of its contract, with as many as 87% of calls unanswered in some months (Prime Minister, 2006). Plunket was (and is) an iconic and popular organisation. A public petition to reinstate funding for PlunketLine attracted more than 50,000 signatures, and support from opposition politicians. Funding was reinstated in 2009.
- *Professional resistance.* Like the public, professional groups can also exert influence over contracting outcomes, and perhaps influence whether providers can successfully deliver. In 2006, Auckland's DHBs issued a tender for their monopoly community laboratory services, and awarded it to a new provider. The new service was more efficient, in part through a proposed reduction in collection locations. GPs, who had established relationships with the incumbent providers, campaigned against the new providers. The decision by the DHBs to award the contract to a new supplier was challenged by way of judicial review, but eventually upheld by the Court of Appeal in 2009. Nevertheless, health practitioners complained of lower service quality, and in late 2009 the DHBs returned 10% of the contract to the original provider (Deuchrass, 2009; Du Chateau, 2007).

#### Q8.16

Are there viable ways to strengthen "provider succession" to shift users of public services from less productive to more productive providers?

## 8.5 Data and ICT

Data and ICT drivers of efficiency seek to utilise increased availability of data and new technology to improve the way public services are targeted or delivered.

### ICT investment

ICT investment is consistently identified as a source of productivity improvement in public services. Dunleavy and Carrera find a strong positive relationship between levels of investment in ICT and productivity growth in UK central departments, accounting for as much as half of the observed variance in productivity rates.

There is an extensive literature on the relationship between ICT investment and productivity. In the 1970s and 1980s, some commentators observed rapid computerisation but slowing productivity growth; in 1987 Robert Solow wrote, "You can see the computer age everywhere but in the productivity statistics" (Solow, 1987, p. 36). Milgrom and Roberts theorised in 1990 that successful firms in the modern economy exploited complementarities, rather than improvements in only one aspect of their operations; success is "not a matter

of small adjustments made independently at each of several margins, but rather have involved substantial and closely coordinated changes in a whole range of the firm's activities" (1990, p. 513).

Reviewing the literature in 2000, Brynjolfsson and Hitt conclude:

Concerns about an information technology "productivity paradox" were raised in the late 1980s. Over a decade of research since then has substantially improved our understanding of the relationship between information technology and economic performance. The firm-level studies in particular suggest that, rather than being paradoxically unproductive, computers have had an impact on economic growth that is disproportionately large compared to their share of capital stock or investment, and this impact is likely to grow further in coming years.

In particular, both case studies and econometric work point to organizational complements such as new business processes, new skills and new organizational and industry structures as a major driver of the contribution of information technology. These complementary investments, and the resulting assets, may be as much as an order of magnitude larger than the investments in computing technology itself. (p. 45)

There is some evidence for similar effects in public services. For example, Garciano and Heaton (2010) examined the relationship between IT utilisation and productivity in US police departments from 1987-2003. At a high level, they found no relationship, except where IT use was also associated with changes in policing practice:

However, while there is no statistically significant relationship between general IT and crime fighting and deterrence (in spite of our large samples), productivity improvements become relatively large when IT adoption is undertaken as part of a whole package of organizational changes. These results endorse the complementarity hypothesis. Police agencies, like firms, appear likely to enjoy the benefits of computerization only when they identify the specific ways in which the new information and data availability interact with existing organisational practices and make adjustments accordingly. (p. 196)

## E-government

Over the last decade, there has been a lot of discussion about e-government. There is a commonly held view that the usual activity of government can be expanded, consolidated, simplified - or at any rate changed - through the increasing use of new information technologies.

The World Bank (2015) defines e-government as "the use of information and communications technology (ICT) to improve the efficiency, effectiveness, transparency and accountability of government". The OECDs recommendation of the Council on Government's digital strategies takes it a step further with the expectation that technology will act not only as a strategic driver for improving public sector efficiency, but also to "to create open, innovative, participatory and trustworthy public sectors" (2015, p. 4). The OECD (2016b) views digital technologies as having potential far beyond efficiency and effectiveness to improving transparency and participation in government, social inclusiveness and government accountability.

Van Dooren et al (2008) cite five different stages in the development of e-government:

- **One-way communication:** Information provision.
- **Two-way communication:** Request and response.
- **Transaction:** Service delivery and financial transaction.
- **Vertical and horizontal integration:** Inter-governmental and intra-governmental integration.
- **Political participation:** Online voting is the example cited by Van Dooren et al, but multiple forms of citizens input to government could be considered here (eg, online surveys, submissions).

All of these are relevant to efficiency, but the second and third of these stages probably have the most scope for improving efficiency in the short-term, and the best evidence about their ability to improve efficiency – as the communication and transactional impacts are most visible.

State sector leaders interviewed by Pickens (2017) identified technology as a key enabler of productivity improvements. This is consistent with a survey cited by Van Dooren et al (2008) where 83% of chief

information officers, of 50 states and 38 major federal agencies, believed that e-government had made government more efficient and 63% claimed that it had reduced costs.

However, Van Dooren et al (2008) found little evidence to test the claims about the benefits of e-government, and found the evidence that existed was anecdotal rather than based on empirical research. They concluded that e-government practices seem to “reduce time demands but increase task demands on staff” and quote an OECD e-government project as stating:

The initial impressive visible results of e-government (government websites, a number of sophisticated transactional services, development of portals) contrast with the next stage of e-government which requires the development of hidden infrastructure, connected back office arrangements and more complex services. (p. 14)

The New Zealand State sector, including the sectors covered by this inquiry, are at varying stages in terms of the five stages of e-government described above. Most agencies sit somewhere around the third of these five stages, having good one-way electronic communication and adequate two-way communication, and there are increasing numbers of agencies that are putting a lot of emphasis on shifting service delivery and transactional services online. Government has also made a number of attempts to coordinate and shape its ICT investment more effectively. This includes the launch of the first Government ICT Strategy in 2013, a review in 2015, the establishment of the Government Chief Information Officer as functional lead on ICT across government and establishing a six-monthly ICT Functional Leadership reporting cycle.

The Ministry of Social Development submitted that its Simplification project was in large part based on redesigning services to leverage technology, with the aim of improving efficiency (Box 8.11).

#### Box 8.11 **MSD's Simplification project**

Since March 2015, MSD has been implementing a large work programme of integrated technical and business changes to its transactional benefit system activities, as part of its “Simplification” work programme. This has involved an effort to change the way clients access services, and in particular to increase the uptake of online channels. It has also focused on improving the efficiency of MSD systems and processes, to reduce transactional work for staff and unnecessary compliance for clients. Specific initiatives include:

- implementing MyMSD, an easy-to-use, mobile service that clients can access from phones, tablets or computers; 430 000 clients have registered to use it as at October 2017;
- rolling out the MyStudyLink smartphone app for students;
- enhancing voice-enable technology used in its contact centre so clients can do more for themselves over the phone;
- eliminating dozens of unnecessary letters and made others available on-line;
- enabling medical practices to lodge clients’ work capacity certificates electronically, which has significantly reduced clients’ need to visit service centres;
- enabling an on-line application form to be pre-populated with data that MSD already holds on clients; and
- improving ease of use of the Work and Income website.

MSD submitted to the inquiry:

Simplification is client-centred, and the work programme is based on insights from clients and staff. It involved consultation and co-design with clients and staff before implementation, and testing and feedback during and after implementation. Simplification has also learnt how to use a scaled agile approach to introduce incremental technology improvements in many small steps rather than in a few 'big bang' releases, especially for client-facing improvements.



The programme started implementing integrated technical and business changes from March 2015. Since then, MSD has developed, delivered, and monitored more than 20 initiatives with attributable work-effort reductions.

Simplification has taken an active benefits-management approach throughout implementation. Transactional work-effort reductions are the combined effect of initiatives to increase online uptake and to streamline and automate processes. Changes in transactional work are evidence-based and are tracked against monthly and end-state targets and are reported on in detail monthly. Changes to digital uptake, processing times, and volumes are measured directly from the Ministry's technology systems in combination with direct observations and time and motion studies.

...

The above changes, aimed at making our services to be "simple at the front", have been complemented by implementing enhancements to business processes making our system "smarter at the back". This work has involved streamlining and automation of business processes, elimination and reduction of manual processing activities, and greater and smarter use of analytics.

MSD has succeeded in shifting clients to online channels, with more than half of "declarations of income" now occurring using the online MyMSD channel, with reductions in people accessing the phone contact centre or visiting a service centre.

The 2017 Budget Estimates for Vote Social Development says that "[t]he impact of efficiency savings and the realisation of simplification benefits sees the overall trend [in departmental expenditure] decreasing from 2018/19 onwards" (p. 61). MSD's annual net savings from simplification are estimated to be \$49.1 million by 2020/21. There are also commensurate reductions in MSD staffing levels, with MSD managing this largely through attrition (rather than restructuring).

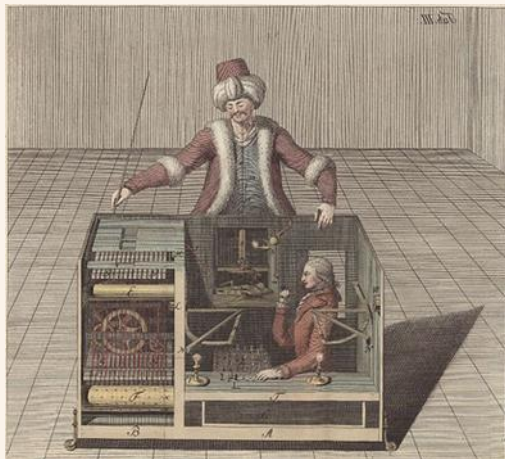
Source: Sub. 12, pp. 2-3

One of the big drivers of government activity to increase the volume of service delivery and transactional activity that can be completed online is likely to be BPS 10, which aims to ensure that "people have easy access to public services". The target of BPS 10 is that 80% of the twenty most common government transactions will be completed digitally by 2021.

A push to move services online too quickly poses a risk that services may shift onto an online format before the back-office processes are ready. This can result in front-end services being conducted online while behind the scenes the back-office processes continue to be done manually. This has occurred, for example, in some of MSD's Simplification activity – where the rapid expansion of online applications was not able to be matched by automated back-end processes. This means that some of MSD's benefit applications that are filed online are still entered into the main data systems manually by a central processing unit (CPU). This suggests that some transactions are going online before the interface between the online system and the main database had been developed; in other words, the back-end processes are gradually being automated, but at a slightly slower rate than the front-end automation.

This illusion of automation can be described as a "Mechanical Turk" after a chess-playing machine that toured Europe and America in the eighteenth century – which pretended to be a robot (or automaton), but had a human hidden inside (see Box 8.12).



**Box 8.12 The Mechanical Turk**

This illusion of automation known as a “Mechanical Turk” is named after the famous chess playing “automaton” created in Austria in 1770 by Wolfgang Von Kempelen. This device was designed to appear to be a chess playing robot, and people who challenged it to a game believed they were competing against a machine. However, the Turk was fake. Although the cabinet at which the “automaton” was seated, appeared to be operated entirely by clockwork, it was actually operated by a human hidden inside.

Mechanical Turk is also the name of an internet crowdsourcing “marketplace for work” website developed and run by Amazon. It claims to enable “individuals or

businesses to use human intelligence to perform tasks that computers are currently unable to do”. It provides an online place to exchange small and simple “Human Intelligence Tasks” (or HITs) that take only a few minutes for very low pay. In the words of Amazon chief executive Jeff Bezos:

Normally, a human makes a request of a computer, and the computer does the computation of the task. But artificial artificial intelligences like Mechanical Turk invert all that. The computer has a task that is easy for a human but extraordinarily hard for the computer. So instead of calling a computer service to perform the function, it calls a human. (Pontin, 2017)

Source: Hitlin, 2016

Rather than being either a fraud, as the original Turk was, or a pragmatic response to the relative costs of programming computers to do small jobs as per Jeff Bezos’s “artificial artificial intelligence”, the “Mechanical Turks” in our current systems are more likely to represent a transitional phase between a fully manual and fully online service. Public agencies are effectively testing the boundaries of how much day-to-day activity can be put online (and how much ought to be done by humans), as well as socialising clients and the public into the greater use of online services, at the same time as we are rapidly building the back-office services that connect the systems together in useful ways to enable them to be more fully automated.

The Commission was told there was no shortage of “low hanging fruit” in terms of opportunities to modernise public services. For example, some medical providers continue to communicate with each other or the public by fax or post. Victims of crime are required to attend police stations in person and complete handwritten forms (where email addresses are recorded separately because there is no field for them on the form). Summoned jurors may need to call an answerphone to find out whether or not they are required at court. In many cases, the use of modern communication channels would be more efficient, especially when considered from the service users’ point of view.

### **Getting benefits from IT investment requires changes to business processes**

The Commission has observed in at least two previous inquiries that the adoption of new technologies, without complementary changes in skills, business processes or organisational structures, limits the benefits from IT investment.

In its inquiry into the services sector of the economy, the Commission observed that firms that used technology to reshape their business models saw greater productivity gains than those that adopted technology to improve their existing business processes:

Extracting the highest possible returns from ICT requires complementary investments, including changes in business organisation. These complementary investments are typically more expensive than the direct cost of the ICT. (NZPC, 2014a, p. 162)

Similarly, in its inquiry into *New models of tertiary education*, the Commission (2017) found that tertiary education providers tended to graft new technology to improve their existing business models, rather than using it to fundamentally reshape their business models. Dunleavy and Carrera have also noted a tendency for governments to adopt sustaining innovations. Their analysis of productivity in UK departments “strongly suggests that one of the most general and dominant problems in growing the productivity of government services has been countering bureaucratic conservatism about digital era changes” (2013, p. 274). But they note that overcoming that conservatism is not easy:

[B]ureaucratic conservatism in adopting digital technologies is not just a one-time problem, which can be easily broken down by e-government or other one-off initiatives, and thereafter marginalized. Instead, long-lived government bureaucracies have a capacity to adopt early technological changes in ways that erect new forms of obstacles to future change. When government organizations incorporate previous waves of innovation, usually with strong time lags, they tend to concretize them in forms that resist further developments. (2013, p. 280)

The available evidence suggests that IT investment, when coupled with changes to organisational processes, is a potentially strong driver of productivity improvement in public services.

**F8.2**

ICT investment can be a strong driver of productivity improvement in public services, but only where it is accompanied by significant changes to organisational processes.

**Q8.17**

What is the experience of New Zealand government agencies in using information technology to improve the productivity of public services? What other changes have been necessary for the success of these initiatives?

The need for complementary investments, business changes or organisational changes to maximise the benefits of IT investment, suggests that such changes may be more difficult to achieve in public services where inputs are more tightly-specified (including through professional regulation, or for political messaging). For example, where politicians desire to fund a greater number of “front-line” police or medical professionals, this could undermine the implementation of technological solutions that would allow routine tasks like reporting to be undertaken more efficiently by back-office staff.

**Q8.18**

Are there input controls (for example, in legislation, or as a result of professional regulation), that prevent IT investment from realising productivity improvements in public services?

## **Better procurement / Improved management of large infrastructure and IT projects**

McKinsey considers that for many countries there is considerable scope to improve government procurement.

Of course, better procurement can help improve efficiency within governments by simplifying and speeding up purchasing procedures. But a push for smarter procurement also drives greater quality and innovation and reduces delivery time of the goods and services purchased, which in turn directly impacts citizen outcomes and satisfaction. (2017, p. 99)

It argues that many governments can realise savings from procurement not only through standardisation and coordinated procurement aimed at reducing the price paid for goods and services, but also through improvements to inventory, supply and order management.

In New Zealand there has been a significant effort to improve government procurement in recent years, with the designation of MBIE’s Chief Executive as government’s “functional leader” for procurement. MBIE’s 2016 annual report says it manages 18 All-of-Government contracts, involving more than 940 government agencies (including schools), and will achieve savings of \$740 million over the life of those contracts. The

Government Chief Information Officer's most recent update on the government's ICT Strategy and Action Plan reported that by March 2017 it had delivered \$107 million against a 2017 target of \$100 million annual savings. These savings include:

- The cost of ICT services negotiated with the market with government acting as a single client, as compared with the cost that would otherwise have been incurred if agencies had engaged with the market individually.
- Cost avoided by performing required activities once rather than by each agency, (e.g. primary procurement and security certification). This has also raised sector capability and reduced risk. (Office of the Minister of Internal Affairs, 2017, pp. 2-3)

However, other government efforts were less successful, including Health Benefits Limited which did not realise anticipated savings for district health boards, ceasing operation in 2015.

In its *More effective social services* inquiry, the Commission noted weaknesses in New Zealand's approach to the commissioning of government funded social services.

The commissioning of social services is a challenging task. It is not generally undertaken in New Zealand in a structured, consistent and effective way. Commissioning organisations should actively build the required skills, capability and knowledge base and use them to substantially lift the quality of commissioning. (2015b, p. 11)

In particular, it found that overly prescriptive and standardised procurement processes could make services less efficient by being too inflexible.

### **Management of large infrastructure and IT projects**

McKinsey (2017) points to improved management of major infrastructure and IT projects as a key source of productivity improvement within government. With respect to major infrastructure projects, they report benefits from working existing infrastructure harder, bringing greater rigour to the prioritisation, planning and monitoring of major projects, and improving the management of project delivery.

In IT, McKinsey emphasises the need to avoid risk-prone megaprojects, emphasising the need to build commercial capabilities in managing such projects, with firm limits for cost and timeliness. They report that public sector IT projects are six times as likely to experience cost overruns, and 20% more likely to be delayed, than comparable private sector projects. There are well-known examples of this in the New Zealand public sector, such as Novopay.

Dunleavy and Carrera (2013) also expressed concern that the private sector market for supplying IT services to government has not been competitive in some countries, including the UK and Japan. They argue uncompetitive markets lead to IT firms investing in lobbying for contracts and other rent-seeking behaviours:

Once the firms acquire very large blocks of work they do not have to be innovative, focusing instead on just curating old-fashioned IT systems over long periods, using proprietary solutions as much as feasible, implemented at huge scale and in very long contracts. ... this way they can simultaneously maintain a high cost base to generate profits from, and yet help insulate their market share, an oligopolist's dream set up.

As an industrial strategy, allowing a closed oligopoly to develop in government IT is lamentably short-sighted, for both government departments and the firms involved. The privileged firms cannot grow their markets via exports, because their expertise is solely in running outdated and expensive à la carte systems, tweaked to indulge the conservatism of large bureaucracies and to respond to the (often uncosted) 'value guidance' (sometimes just whims) of politicians. The firms have no interest in promoting technologies that would produce low-cost, modular solutions that could potentially go on to win business in much larger-scale world markets. (Dunleavy & Carrera, 2013, p. 279)

In 2014 the Commission noted that the New Zealand government was far from the leading edge in terms of ICT procurement or adoption:

On the contrary, the technologies that government agencies use are often older than those used in the private sector. For example, some agencies continued to use Windows XP long after it has been

superseded by more recent versions and right up till security support was withdrawn by Microsoft. ... The Government is not yet an effective and informed purchaser of ICT. (NZPC, 2014a, p. 195)

In 2014, the Commission recommended that:

The Treasury, the State Services Commission and the Department of Prime Minister and Cabinet, working together with the Government Chief Information Officer, should agree on and mandate a means to ensure that significant government ICT projects are evaluated and reviewed and the results disseminated both within government and more widely. (NZPC, 2014a, p. 199)

There has not been a formal government response to the *Boosting productivity in the services sector* inquiry indicating whether or not government intends to take up this recommendation. But in 2015 government introduced changes to the government's Investment Management System, and a revised Government ICT Strategy and Action Plan.

### The investment management system

The focus of changes in the investment management system include providing more ministerial oversight over the entirety of government's investment activity, and providing ministers with assurance about agencies capability to manage investments, and requiring "investment-intensive agencies" to prepare long-term investment plans.

Since 2008, major government investments have been subject to Gateway reviews, whereby projects are assessed by an independent review team at key decision points. Review reports are intended to support, and are confidential to, the agency leading the initiative and the Treasury. But Treasury also publishes regular reports on "lessons learned" from Gateway reviews. The latest such report finds that:

- Governance is a significant and growing concern, in part because the increasing trend towards large complex multi-agency and all-of-government projects requires a list in capability. There are significant concerns that expectations are exceeding capability in this area.
- RAID (Risks, Assumptions, Issues and Dependencies) management is a consistently weak area in the majority of projects reviewed.
- Although the number of recommendations around Business Cases has dropped since the introduction of Better Business Cases (BBC), it remains a key theme and there is still significant uncertainty about use of the BBC process. (New Zealand Treasury, 2017b, pp. 2-3)

### Government's ICT Strategy and Action Plan

The Government ICT Strategy and Action Plan emphasises the importance of agencies being more nimble in adopting technology:

The mission, structure and staffing of ICT units and service delivery workforces will need to change. Just as Government needs to manage the economic and social consequences of disruptive technological change, e.g., workforce automation, public sector leaders must carefully plan and manage transitions to new ways of integrating and interacting with technology.

The way agencies develop services will also need to change. This will require agencies adopt innovative approaches to developing services, such as focusing on minimum viable products that are iterated rapidly. (DIA, 2015)

To support this, the strategy outlines how government agencies will implement common policies and standards. McKinsey comments favourably on the New Zealand "building a government cloud program to make it easier for government to buy IT from small and midsize enterprises" (2017, p. 106). The strategy also notes opportunities to align IT investment with "agency modernisation initiatives".

The Commission is interested in submissions on how competitive the market for government IT services is, and whether this could be inhibiting different IT investments that could achieve productivity improvements in public services.

#### Q8.19

Has the government's management of large infrastructure and information technology projects improved? In what ways could it improve further?

## Data and targeting

The UK's public sector efficiency group concluded in its final report that data and targeting will be a significant driver of efficiency, particularly over the long-term. It reported that these approaches "offer new ways of delivering services and tailoring interventions to the needs of different client groups, with consequential efficiency gains" (DCLG, 2015, p. 85).

In its inquiry into *More effective social services*, the Commission (2015) found that the social services sector had been slow to use data and analytics to innovate. The Commission recommended the construction of a wide-access, client-centred, social services data network to enable the social services system to learn, innovate, and become more effective. It also found an absence of capability within government to use data more effectively.

However, the terms of reference for this inquiry direct the Commission to not duplicate work on issues, such as where to invest, "that are being developed as part of the social investment approach".

## 8.6 Conclusion

This inquiry is in part about how government can improve the measurement of efficiency in the delivery of public services to New Zealanders. Those services matter for all New Zealanders, but particularly the most vulnerable in society. The outcomes that society wants from these services – effective healthcare, quality education, safety and security, decent housing, and dignity – matter a great deal to the wellbeing of New Zealand.

The public and officials are rightly concerned with how well government achieves these outcomes – what activities government does to deliver better health, education, justice and social development outcomes. But they must also be concerned with how to productively use public resources to maximise the outputs of government.

A focus on how well outcomes are achieved, and how efficiently outputs are delivered, are not in conflict. Both are necessary. A single-minded focus on technical efficiency, without care about what services are being produced and why, is truly unproductive: "There is surely nothing quite so useless as doing with great efficiency what should not be done at all" (Drucker, 1963). Similarly, delivering the right services in an inefficient way is also wasteful – it means that public services will deliver fewer outputs, and the resulting outcomes in terms of better health, education, justice or social development will also be reduced.

Indeed, a careful and thorough focus on efficiency will require an agency to understand and take a broad assessment of its activity – is it doing the right things? Is it doing them the best way that it can? Is it looking for opportunities to improve?

This inquiry finds that the public sector can do a much better job of measuring and understanding the efficiency of its activity, and that this information needs to be used systematically alongside other sources of performance information to deliver better outcomes for New Zealanders.

# Summary of questions

## Chapter 8 – Encouraging productivity

**Q8.1**

What evidence is there about the results of “lean” and related business improvement processes in public services? Are there evaluations about their use in the New Zealand public sector?

**Q8.2**

What is the current state of evaluation in New Zealand public services? Is there demand for evaluation? Do departments have the capability and funding to undertake or purchase evaluation?

**Q8.3**

In what areas of state sector is local/regional innovation occurring? Does innovation spread to other local/regional organisations? What factors explain the pace of innovation spread?

**Q8.4**

What measures exist to facilitate the spread of innovation? How successful have these measures been? What other measures or initiatives would assist in increasing the pace with which innovations spread?

**Q8.5**

How can government signal that it values innovation in the delivery of public services? Given some innovation will fail, how can it signal a greater tolerance for such risk?

**Q8.6**

How can Treasury and the State Services Commission support public sector agencies to be innovative and constantly seek improvements to the efficiency and effectiveness of service delivery?

**Q8.7**

How can human resource management practices support public sector efficiency? What evidence is there about the effect of practices like High Performance Engagement on public sector efficiency?

**Q8.8**

How is information about the efficiency of public services used to inform budget decisions in New Zealand? How does performance information affect Treasury advice about budget decisions?

**Q8.9**

What are the lessons from New Zealand’s experience of fixed nominal baselines?

**Q8.10**

How well does the Four Year Plan process provide information about the value created with existing baseline expenditure and resources? How well does the Four Year Plan process identify how departments will increase the value they create over the medium-term?

**Q8.11**

What evidence is there on the effect on productivity of “line by line” reviews undertaken as part of Budget 2009? Should there be more regular scrutiny of base expenditure to identify efficiency savings? If so, what form could this process take?

**Q8.12**

In what circumstances might performance related pay be effective? Where is it likely to be ineffective or harmful, and why?

**Q8.13**

What evidence is there about incentives on public sector agencies to seek efficiencies? What evidence is there that initiatives like the Justice Sector Fund increase agencies' incentives to seek efficiencies? What has been its effect on the overall efficiency of the justice system?

**Q8.14**

Is regulation inside government a barrier to improving the productivity of public services? What examples or evidence is there?

**Q8.15**

Are there issues of insufficient scale in the organisation of health or education services in New Zealand? What evidence is there of this?

**Q8.16**

Are there viable ways to strengthen "provider succession" to shift users of public services from less productive to more productive providers?

**Q8.17**

What is the experience of New Zealand government agencies in using information technology to improve the productivity of public services? What other changes have been necessary for the success of these initiatives?

**Q8.18**

Are there input controls (for example, in legislation, or as a result of professional regulation), that prevent IT investment from realising productivity improvements in public services?

**Q8.19**

Has the government's management of large infrastructure and information technology projects improved? In what ways could it improve further?

# Findings

The full set of findings and recommendations from the report are below.

## Chapter 2 – Measuring state sector productivity

### Findings

**F2.1**

There is a significant volume of research that applies different approaches to the measurement of state sector productivity, from which a lot can be learned. However, in recent years, progress in the development of robust and regularly-reported public sector productivity measures has slowed, both in New Zealand and internationally.

**F2.2**

Productivity measurement can occur at a number of levels (eg sector, institution, service). Measures are likely to be most useful where they align with the decision rights or scope of control of decision makers.

**F2.3**

Different types of output and input measures will answer different questions. For example, gross output measures are useful for understanding the total output of a sector or organisation, while value added measures are useful for assessing the marginal additional value added (or removed).

## Chapter 3 – Measuring outputs

### Findings

**F3.1**

Productivity measures show how efficiently outputs (the final goods or services produced by an entity for individuals or organisations) are produced, and hence are not a complete measure of state sector performance. Other performance information including measurement of outcomes should be captured through separate processes.

**F3.2**

Data limitations should not necessarily be a barrier to productivity measurement. Any data limitations should be clearly articulated, and measurement efforts should be undertaken in parallel with work to improve data collection.

## Chapter 4 – Selecting and costing inputs

### Findings

**F4.1**

The capital charge (or equivalent Treasury discount rate) is a reasonable proxy for inclusion in the calculation of capital inputs.

**F4.2**

Challenges measuring the impact of technological improvement on inputs, or the contribution of client characteristics to service production, can often be managed through the quality adjustment of outputs.



## Chapter 5 – Adjusting for complexity, quality and price changes

### Findings

#### F5.1

Productivity measures should account for differences or changes in an organisation's operating environment. Key differences that it can be useful to account for include the characteristics of the clients of the services, the size and scope of the organisations, and market structure.

These differences can be reflected in measures by:

- measuring the outputs related to different population subgroups separately (segmenting the population) and treating them as distinct outputs;
- limiting the range of providers studied to those from similar environments; and
- adjusting the volumes of outputs for differences in the operating environment (eg severity of treatments).

#### F5.2

Properly measured, productivity captures changes in both the volume and quality of outputs. Quality changes can be reflected in productivity measures by:

- grouping outputs so that only products and services of the same specification are compared;
- adjusting outputs based on the resulting change in status directly attributable to the services received; or
- adjusting outputs based on final outcomes; for example, the population's education level or life expectancy, without necessarily relating change to any intervention.

#### F5.3

Removing the effect that prices may have on comparisons over time helps ensure productivity measures accurately reflect changes in volumes. However, there is scope for the introduction of significant error from using the wrong deflators. Price effects can be removed using market information (eg, the consumer or producer price indices, or subgroups of these indices) or direct volume indices.

## Chapter 6 – Putting productivity measures in place

### Findings

#### F6.1

For large sections of the state sector, there is little measurement or understanding of productivity. In some sectors covered by this inquiry, the Commission was unable to access the data needed to conduct productivity assessments.

#### F6.2

The lack of available data from parts of the health sector and from parts of the justice sector limits meaningful measurement of the efficiency of the New Zealand justice and health systems.

#### F6.3

Current public sector financial management systems provide weak incentives for agencies to seek efficiencies within existing programmes.

**F6.4**

Concepts such as “productivity” or “efficiency” have negative connotations for some parts of the public sector. These connotations include concerns that the pursuit of efficiency may compromise important social outcomes and values, or lead to reduced service quality.

**F6.5**

The absence of productivity or efficiency measures likely reflects low demand from ministers, and insufficient demand from senior public sector leaders for the necessary information and analysis.

**F6.6**

“Core” public service agencies do not appear to have the openness to change and innovation required to drive service and productivity improvements.

**F6.7**

The following principles are useful for expanding and integrating productivity measures into public service practice and performance systems:

- design productivity measures to complement outcomes;
- design measures with the involvement of staff who deliver services;
- as far as possible, collect productivity data as part of business-as-usual activity;
- use productivity information primarily as the basis for conversations and learning about service improvement;
- ensure agency leaders actively support the use of productivity measures;
- develop measures that enable comparisons to be made between similar organisations, business units or outputs, to identify and promote sharing of good practice; and
- treat productivity measures as one input into performance decisions, rather than the sole factor with high stakes impacts.

## Chapter 7 – Capability and systems needed to measure productivity

### Findings

**F7.1**

All agencies have the starting capability to develop at least some simple productivity metrics - for at least parts of their systems (or some of their services).

**F7.2**

A better “whole-of-system” understanding of agencies’ business flows will enable more comprehensive measures of whole system productivity to be developed. Few agencies currently take a whole-of-system approach to productivity improvement.

**F7.3**

There is little information about the level of capability to measure productivity across the state sector (in terms of both current capability and expected future need).

**F7.4**

Some technical and analytic capability exists in most agencies, and much can be done to measure productivity with this current capability, as long as each agency has a thorough understanding of its operational business flow, processes and goals.

**F7.5**

Productivity measurement requires a similar skill set to that needed for social investment approaches; therefore, better connections between these two should be built.

**F7.6**

State sector agencies already have much of the capability to measure productivity, but leaders are not currently incentivised to understand it, measure it and act upon it. Framing productivity in the language of social outcomes rather than that of economic efficiency would help make these messages more accessible.

**F7.7**

There is a lack of connection between the people with technical skills (to measure and understand productivity) and the managerial and strategic activities of agencies.

**F7.8**

New Zealand has a well-developed government fiscal management system, which provides a good basis to measure efficiency.

**F7.9**

The Head of the Finance Profession is working to shift agencies from a transactional to a more strategic financial management capability. This would enable agency finance teams to better connect productivity measurement to operational decision making.

**F7.10**

Government agencies hold a wealth of data, some of which could be used to develop productivity measures, but it could be improved through better integration across agencies and with business re-engineering processes.

**F7.11**

Better data integration, and the use of existing data that is routinely collected, would enable agencies to measure productivity. But work is needed to build public trust in government's ability to integrate and use routinely collected data in ways that are safe.

**F7.12**

The Performance Improvement Framework has potential as a tool to support agencies to measure and improve productivity.

## Chapter 8 – Encouraging productivity

### Findings

**F8.1**

There are successful innovations in New Zealand's public services, but they do not spread as far or fast as is desirable. This is a system failure.

**F8.2**

ICT investment can be a strong driver of productivity improvement in public services, but only where it is accompanied by significant changes to organisational processes.

# Appendix A Public consultation

## Submissions

<b>Individual or Organisation</b>	<b>Submission number</b>
Bruce D White Consulting	Sub 013
District Health Boards	Sub 017
Hermann Grobler	Sub 005
Inland Revenue	Sub 007
Ministry of Social Development	Sub 012
New Zealand College of Critical Care Nurses	Sub 015
New Zealand Council of Trade Unions	Sub 009
New Zealand Kindergartens	Sub 010
New Zealand Nurses Organisation	Sub 014
NextEra Global	Sub 016
Patricia M Harrison	Sub 003
PEPworldwide NZ Limited	Sub 006
Public Service Association	Sub 011
Te Rito Maioha Early Childhood New Zealand	Sub 002
The Methodist Mission	Sub 001
The New Zealand Initiative	Sub 008
Youth Horizons	Sub 004

## Engagement meetings

Accident Compensation Corporation  
Australian Department of Finance  
Bruce D White Consulting Ltd  
Canterbury District Health Board  
Central TAS  
Department of the Prime Minister and Cabinet – Policy Project  
Dr Chris Eichbaum – Victoria University  
Education New Zealand  
Education Payroll Limited  
HM Treasury  
James Mansell  
Jon R Blondal, Senior Budge Official of the OECD  
Justice Sector – Ministry of Justice, Department of Corrections, New Zealand Police  
Land Information New Zealand  
McKinsey Global Institute Council  
Methodist Mission Southern  
Ministry for Vulnerable Children, Oranga Tamariki  
Ministry of Health  
Ministry of Education  
Ministry of Justice

Ministry of Social Development  
Murray Horn  
New Zealand Council of Trade Unions  
New Zealand Customs Service  
New Zealand Police  
New Zealand Treasury  
New Zealand Treasury – Strategic Performance Improvement Group  
NZCER  
Professor Arthur Grimes – Victoria University  
Professor Donald Moynihan – Robert M la Follette School of Public Affairs, University of Wisconsin  
Professor Martin Connor – Executive Director, Centre for Health Innovation, Griffith University  
Professor Patrick Dunleavy  
Rural Women New Zealand  
Social Investment Agency  
State Services Commission  
Statistics NZ

## Roundtables

Official's workshop on Measuring State Sector Productivity, Wellington	27 October 2017
'The Health Roundtable - NZ Chapter', Wellington	23 Nov 2017

## Conferences and seminars

'Empowering Customers & Communities', Wellington	5 July 2017
Dr Richard Meade, 'The Economics of Social Services', The Treasury	17 August 2017
Professor Zoe Radnor, "'Lean" and public sector productivity: panacea or paradox?', Reserve Bank of New Zealand	23 August 2017
Professor Martin Connor, Executive Director, Centre for Health Innovation, Griffith University at the Reserve Bank of New Zealand	30 October 2017

# References

- Accent and RAND Europe. (2010). *Review of Stated Preference and Willingness to Pay Methods*. Retrieved 13 September 2017 from [http://webarchive.nationalarchives.gov.uk/+http://www.competition-commission.org.uk/our\\_role/analysis/summary\\_and\\_report\\_combined.pdf](http://webarchive.nationalarchives.gov.uk/+http://www.competition-commission.org.uk/our_role/analysis/summary_and_report_combined.pdf)
- Air New Zealand. (2015). *Sustainability Report / 2015*. Retrieved 20 November 2017 from [www.airnewzealand.co.nz/assets/PDFs/J004845-Sustainability-Report-2015\\_FINAL-ART-NEW\\_MEDRES.pdf](http://www.airnewzealand.co.nz/assets/PDFs/J004845-Sustainability-Report-2015_FINAL-ART-NEW_MEDRES.pdf)
- Ammons, D. N., & Rivenbark, W. C. (2008). Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina benchmarking project. *Public Administration Review*, 68(2), 304–318.
- Andersen, S. C., & Moynihan, D. P. (2016). Bureaucratic investments in expertise: Evidence from a randomized controlled field trial. *The Journal of Politics*, 78(4), 1032–1044.
- APC [Australian Productivity Commission]. (2015). *On productivity: concepts and measurement*. Productivity Commission Staff Research Note. Retrieved 7 July 2017 from [www.pc.gov.au/research/supporting/concepts-measurement/concepts-and-measurement.pdf](http://www.pc.gov.au/research/supporting/concepts-measurement/concepts-and-measurement.pdf)
- APC. (2017a). *Report on government services, courts*. Retrieved 10 October 2017 from [www.pc.gov.au/research/ongoing/report-on-government-services/2017/justice/courts/rogs-2017-volumec-chapter7.pdf](http://www.pc.gov.au/research/ongoing/report-on-government-services/2017/justice/courts/rogs-2017-volumec-chapter7.pdf)
- APC. (2017b). *Shifting the Dial: 5 Year Productivity Review. Supporting Paper 5 – Integrated Care*. Canberra: Australian Productivity Commission.
- APC and NZPC. (2012). *Strengthening trans-Tasman economic relations*. Retrieved 10 October 2017 from [www.productivity.govt.nz/inquiry-content/1774?stage=4](http://www.productivity.govt.nz/inquiry-content/1774?stage=4)
- Atkinson, T. (2005). *Atkinson review: final report. Measurement of government output and productivity for the national accounts*. Retrieved 7 July 2017 from [www.ons.gov.uk/ons/guide-method/method-quality/specific/public-sector-methodology/articles/atkinson-review-final-report.pdf](http://www.ons.gov.uk/ons/guide-method/method-quality/specific/public-sector-methodology/articles/atkinson-review-final-report.pdf)
- Baekgaard, M. (2017). *Prospect theory and public service outcomes: Examining risk preferences in relation to public sector reforms*. *Public Administration*.
- Barber, M. (2017). *Delivering better outcomes for citizens: practical steps for unlocking public value*. Retrieved 24 November 2017 from [www.gov.uk/government/publications/delivering-better-outcomes-for-citizens-practical-steps-for-unlocking-public-value](http://www.gov.uk/government/publications/delivering-better-outcomes-for-citizens-practical-steps-for-unlocking-public-value)
- Barslund, M., & O'Mahony, M. (2012). *Output and productivity in the education sector. Indicators for Evaluating International Performance in Service Sectors (INDICSER)*. Discussion Paper 39. Research Directorate General, European Commission, Brussels.
- Behn, R.D. (2003). The big questions of public management. *Public administration review*, 55(4), 313–324.
- Brenton S. (2016). *The politics of budgetary surplus*. London: Palgrave Macmillan.
- Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 287–302.
- Berman, E., & Wang, X. (2000). Performance measurement in US counties: Capacity for reform. *Public Administration Review*, 60(5), 409–420.
- Berwick, D., Loehrer, S., & Gunther-Murphy, C. (2017). Breaking the rules for better care. *Journal of the American Medical Association*, 317(21):2161-2162.

- Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *The American economic review*, 95(3), 616–636.
- Better Public Services Advisory Group. (2011). *Better Public Services Advisory Group report*. Wellington: State Services Commission.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public administration*, 84(3), 517–538.
- Black, S. (1998). Measuring the value of better schools. Federal Reserve Bank of New York. *Economic Policy Review*, 4(1), 87–94.
- Bose, D., Philip, R., & Sullivan, R. (2016, June/July). *Returning to surplus: New Zealand's recent fiscal consolidation experience*. Paper prepared for the NZAE annual conference, Auckland. Retrieved 20 November 2017 from [www.nzae.org.nz/wp-content/uploads/2016/10/Returning\\_to\\_surplus\\_NZAE.pdf](http://www.nzae.org.nz/wp-content/uploads/2016/10/Returning_to_surplus_NZAE.pdf)
- Boser, U. (2014). *Return on Educational Investment: 2014. A district-by-district evaluation of U.S. educational productivity*. Retrieved 8 June 2017 from <https://cdn.americanprogress.org/wp-content/uploads/2014/07/ROI-report.pdf>
- Bourdeaux, C., & Chikoto, G. (2008). Legislative influences on performance management reform. *Public Administration Review*, 68(2), 253–65.
- Boyle, R. (2006). *Measuring Public Sector Productivity: Lessons from International Experience*. CPMR Discussion Paper, no. 35. Retrieved 14 September 2017 from [www.ipa.ie/fileUpload/Documents/CPMR\\_DP\\_35\\_Measuring\\_Public\\_Sector\\_Productivity\\_Lessons\\_from\\_International\\_Experience.pdf](http://www.ipa.ie/fileUpload/Documents/CPMR_DP_35_Measuring_Public_Sector_Productivity_Lessons_from_International_Experience.pdf)
- Boyne, G., & Chen, A. (2007). Performance targets and public service improvement. *Journal of Public Administration Research and Theory*, 17, 445–477.
- Broatch, M. (2017, April). Cataract surgery: sight for sore eyes. *Listener*.
- Brough, J. N. (2005). *Baseline and output price reviews in New Zealand: a research report submitted to the Victoria University of Wellington in partial fulfilment of the requirements for the Master of Public Management (M.P.M.)*. Wellington, New Zealand: Victoria University of Wellington.
- Brumby, J., Edmonds, P., & Honeyfield, K. (1996, August). *Effects of public sector financial management reform in New Zealand*. Paper presented to the Australasian Evaluation Society Conference, Wellington, New Zealand.
- Brynjolfsson, E., & Hitt, L. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4), 23–48.
- Burgess, S., & Ratto, M. (2003). The role of incentives in the public sector: issues and evidence. *Oxford review of economic policy*, 19(2), 285–300.
- Byrne, D. M., Fernald, J. G., & Reinsdorf, M. B. (2016). Does the United States have a productivity slowdown or a measurement problem? *Brookings Papers on Economic Activity*, 2016(1), 109–182.
- Cookson, R. A., Laudicella, M., & Li Donni, P. (2013). Does hospital competition harm equity? Evidence from the English National Health Service. *Journal of Health Economics*, 32(2), 410–422.
- Cooley, W.W. (1983). Improving the performance of an educational system. *Educational researcher*, 12(6), 4–12.
- Cooper, Z., Gibbons, S., Jones, S., & McGuire, A. (2010). *Does hospital competition improve efficiency? An analysis of the recent market-based reforms to the English NHS*. CEP Discussion Paper No 988. Retrieved 20 November 2017 from <http://cep.lse.ac.uk/pubs/download/dp0988.pdf>

- Cooper, Z., Gibbons, S., Jones, S., & McGuire, A. (2012). *Does hospital competition improve efficiency? Evidence from a quasi-experiment in the English National Health Service*. CEP Discussion Paper No 1125. Retrieved 20 November 2017 from <http://cep.lse.ac.uk/pubs/download/dp1125.pdf>
- Cumming, J., Mays, N., & Daubé, J. (2010). How New Zealand has contained expenditure on drugs. *British Medical Journal*, 340, 1224–1227.
- Currstine, T., Lonti, Z., Joumard, I. (2007). Improving public sector efficiency: Challenges and opportunities. *OECD Journal on Budgeting*, 7(1), 1608–7143.
- Data Futures Forum. (2014). *Harnessing the economic and social power of data*, NZDFF, Paper three. Retrieved 6 September 2017 from <http://datafutures.co.nz/assets/Uploads/Data-Futures-FORUM-NZDFF-harness-the-power.pdf>
- Data Futures Partnership. (2017). *A path to social licence: Guidelines for trusted data use*. Data Futures Partnership, Wellington. Retrieved 6 September 2017 from <https://trusteddata.co.nz/wp-content/uploads/2017/08/Summary-Guidelines.pdf>
- Data Futures Partnership. (N.D. a). *About the partnership*. Retrieved 22 September 2017 from <http://datafutures.co.nz/about-the-partnership/>
- Data Futures Partnership. (N.D. b) Terms of Reference, <http://datafutures.co.nz/assets/Uploads/Approved-DFP-WG-Terms-of-Reference-14-March-2016.pdf>
- Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A., & Weale, M. (2004). *Developing new approaches to measuring NHS outputs and productivity*. Retrieved 6 September 2017 from [www.york.ac.uk/che/pdf/tp31.pdf](http://www.york.ac.uk/che/pdf/tp31.pdf)
- DCLG (UK Department for Communities and Local Government). (2015). *Public sector efficiency: Final report from the cross-Whitehall analysts' group*. Government Economic Service, London.
- Demaj, L. (2017). What can performance information do to legislators? A budget-decision experiment with legislators. *Public Administration Review*, 77(3), 366–379.
- Demaj, L., & Schedler, K. (2014). What drives legislators' information behaviour? An eye-tracking experiment with legislators. Retrieved 22 September 2017 from [www.alexandria.unisg.ch/234449/1/Demaj%20%26%20Schedler\\_2014\\_What%20Drives%20Legislators%20Information%20Behavior.pdf](http://www.alexandria.unisg.ch/234449/1/Demaj%20%26%20Schedler_2014_What%20Drives%20Legislators%20Information%20Behavior.pdf)
- Department of Internal Affairs. (2017). *Guidance and resources: privacy*. Retrieved 1 December from [www.ict.govt.nz/guidance-and-resources/privacy](http://www.ict.govt.nz/guidance-and-resources/privacy)
- Department of Internal Affairs. (2015). *Government ICT strategy and action plan: Technology*. Retrieved 20 November 2017 from [www.ict.govt.nz/strategy-and-action-plan/strategy/technology/](http://www.ict.govt.nz/strategy-and-action-plan/strategy/technology/)
- Department of Internal Affairs. (2016). *Information Technology and Digital Skills Survey*. Wellington: Author.
- Department of Internal Affairs. (2017). *Privacy maturity assessment framework and guidance on privacy management*. Department of Internal Affairs. Retrieved 9 December 2017 from [www.ict.govt.nz/assets/Guidance-and-Resources/User-guide-for-the-Privacy-Maturity-Assessment-Framework-version-2-0.pdf](http://www.ict.govt.nz/assets/Guidance-and-Resources/User-guide-for-the-Privacy-Maturity-Assessment-Framework-version-2-0.pdf)
- Deuchrass, A. (2009). *Australian owner appoints new boss for troubled Labtests*. Retrieved 20 November 2017 from [www.nbr.co.nz/article/australian-owner-appoints-new-boss-troubled-labtests-110746](http://www.nbr.co.nz/article/australian-owner-appoints-new-boss-troubled-labtests-110746)
- Diewert, E. (2011). Measuring productivity in the public sector: some conceptual problems. *Journal of Productivity Analysis*, 36: 177–191.



- Diewert, W. (2017). *Productivity Measurement in the Public Sector: Theory and Practice*. Retrieved 7 September 2017 from [http://econ.sites.olt.ubc.ca/files/2017/02/pdf\\_paper\\_erwin-diewert-17-01-TheoryandPractice.pdf](http://econ.sites.olt.ubc.ca/files/2017/02/pdf_paper_erwin-diewert-17-01-TheoryandPractice.pdf)
- Dixit, A. (2002). Incentives and organizations in the public sector: an interpretative review. *The journal of human resources*, 37(4), 696–727.
- Djellal, F., & Gallouj, F. (2008). *Measuring and improving productivity in services: issues, strategies and challenges*. Cheltenham, United Kingdom: Edward Elgar Publishing.
- DoFD (Department of Finance and Deregulation) (2011). *Report of the review of the measures of agency efficiency*. Retrieved 10 November 2017 from [www.finance.gov.au/publications/measures\\_of\\_agency\\_efficiency/docs/measures\\_of\\_agency\\_efficiency.pdf](http://www.finance.gov.au/publications/measures_of_agency_efficiency/docs/measures_of_agency_efficiency.pdf)
- Douglas, J. (2006). *Measurement of public sector output and productivity*. Retrieved 7 September 2017 from [www.treasury.govt.nz/publications/research-policy/ppp/2006/06-09/tpp06-09.pdf](http://www.treasury.govt.nz/publications/research-policy/ppp/2006/06-09/tpp06-09.pdf)
- Downs, A. (2017). *From theory to practice: The promise of primary care in New Zealand*. Retrieved 20 November 2017 from [www.fulbright.org.nz/wp-content/uploads/2017/09/DOWNS-From-Theory-to-Practice-The-Promise-of-Primary-Care-in-New-Zealand-.pdf](http://www.fulbright.org.nz/wp-content/uploads/2017/09/DOWNS-From-Theory-to-Practice-The-Promise-of-Primary-Care-in-New-Zealand-.pdf)
- Drucker, P. (1963). Managing for business effectiveness. *Harvard Business Review*. Retrieved 20 November 2017 from <https://hbr.org/1963/05/managing-for-business-effectiveness>
- Du Chateau, C. (2007, January). Blood on the lab floors. *New Zealand Herald*. Retrieved 20 November 2017 from [www.nzherald.co.nz/nz/news/article.cfm?c\\_id=1&objectid=10419977](http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10419977)
- Duckett, S. (2013). *Australia's bad drug deal: high pharmaceutical prices*. Melbourne, Australia: Grattan Institute.
- Dull, M. (2009). Results-model reform leadership: Questions of credible commitment. *Journal of Public Administration Research and Theory*, 19, 255–84.
- Dunleavy, P. (2015). Public sector productivity: puzzles, conundrums, dilemmas and their solutions. In J. Wanna, H-A Lee, & S. Yates (Eds.) *Managing under austerity, delivering under pressure*. Canberra, Australia: ANU Press, 25–42.
- Dunleavy, P. (2016, November). *Public Sector Productivity – Measurement Challenges, Performance Information and Prospects for Improvement*. Paper to the Annual Meeting of the OECD Senior Budget Officers' Network OECD, Paris, France.
- Dunleavy, P., & Carrera, P. (2013). *Growing the productivity of government services*. Cheltenham, United Kingdom: Edward Elgar.
- East, P. (1997). Opening address, in State Services Commission (1997), *Future issues in public management*. Wellington: SSC.
- Elias, S. (2017, August). *Managing criminal justice*. Address given at the Criminal Bar Association Conference. Retrieved 16 October 2017 from [www.courtsfnz.govt.nz/speechpapers/CBA.pdf](http://www.courtsfnz.govt.nz/speechpapers/CBA.pdf)
- Engler, R. (2010). *Academic performance of first-year bachelors students at university*. Ministry of Education. Retrieved 8 September 2017 from [www.educationcounts.govt.nz/data/assets/pdf\\_file/0006/76488/Performance-firstYear-18052010b.pdf](http://www.educationcounts.govt.nz/data/assets/pdf_file/0006/76488/Performance-firstYear-18052010b.pdf)
- Eurostat. (2001). *Handbook on price and volume measures in national accounts*. Luxembourg: Office for Official Publications of the European Communities.

- Francis, R. (2010). *Robert Francis Inquiry report into Mid-Staffordshire NHS Foundation Trust*. Retrieved 20 November 2017 from [http://webarchive.nationalarchives.gov.uk/20130104234315/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_113018](http://webarchive.nationalarchives.gov.uk/20130104234315/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_113018)
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., & Neely, A. (2007). Towards a definition of a business performance measurement system. *International journal of operations & production management*, 27(8), 784–801.
- Frank, R. G., Berndt, E. R., & Busch, S. H. (1999). Price indexes for the treatment of depression In Triplett, J. E. (Ed.), *Measuring the prices of medical treatments*, pp. 72–102, Washington, DC, USA: Brookings Institution Press.
- Fraser, H., & Nolan, P. (2017). *Understanding health sector productivity*. Wellington: New Zealand Productivity Commission.
- Gabbittas, O., & Jeffs, C. (2008). *Assessing productivity in the delivery of public hospital services in Australia: Some experimental estimates*. Paper presented at the 30<sup>th</sup> Australian Health Economics Conference, Adelaide, Australia.
- Garciano, L., & Heaton, P. (2010). Information technology, organization and productivity in the public sector: Evidence from police departments. *Journal of Labor Economics*, 28(1), 167-210.
- Gaynor, M., Moreno-Serra, R., & Propper, C. (2010). *Death by market power: Reform, competition and patient outcomes in the National Health Services*. CMPO Working Papers. Bristol, United Kingdom: Bristol University.
- Gemmell, N., Nolan, P., & Scobie, G. (2017a). *Estimating quality-adjusted productivity in tertiary education: Methods and evidence for New Zealand*. Wellington: Victoria University of Wellington Chair in Public Finance and New Zealand Productivity Commission.
- Gemmell, N., Nolan, P., & Scobie, G. (2017b). *Public sector productivity: Quality adjusting sector-level data on New Zealand schools*. Wellington: New Zealand Productivity Commission.
- Genet, T. (2017). *Productivity measurement case study: Courts*. New Zealand Productivity Commission Research Note 2017/04.
- Genet, T., & Hayward, M. (2017). *Productivity measurement case study: Police*. New Zealand Productivity Commission Research Note 2017/09.
- Gibson, J., & Boe-Gibson, G. (2014). *Capitalizing performance of 'free' schools and the difficulty of reforming school attendance boundaries*. Working Paper 8/14, Department of Economics, University of Waikato.
- Gibson, J., Sabel, C., Boe-Gibson, G., & Kim, B. (2007, July). *House prices and school zones: does geography matter?* Paper presented to a conference of the Spatial Econometrics Association, Cambridge.
- Gill, D., & Schmidt, T. (2011). Organisational performance management: concepts and themes. In Gill, D. (Ed.), *The iron cage recreated: the performance management of state organisations in New Zealand*. Wellington: Institute of Policy Studies, pp.9–36.
- Gill, D. (2011). Regulatory management in New Zealand: what, how and why? In Frankel, S. (Ed.), *Learning from the past, adapting to the future: regulatory reform in New Zealand*. Wellington: Lexis Nexis, pp.173–214.
- Gill, D., Kengema, L., & Laking, R. (2011). Information that managers use: results from the managing for organisational performance survey. In Gill, D. (Ed.), *The iron cage recreated: the performance management of state organisations in New Zealand*. Wellington: Institute of Policy Studies, pp.375–402.

- Giordano, R., Lanau, S., Tommasino, P., & Topalova, P. B. (2015). *Does public sector inefficiency constrain firm productivity: evidence from Italian provinces*. Washington, DC, United States of America: International Monetary Fund.
- Goodridge, P. (2007). Index numbers. *Economic & Labour Market Review*, 1(3), 54–57.
- Gordon, J., Zhao, S., & Gretton, P. (2015). *On productivity: concepts and measurement*. Canberra, Australia: Australian Productivity Commission.
- Green, N. (2017). *Productivity measurement case study: early childhood education*. New Zealand Productivity Commission Research Note 2017/05.
- Greenberg, R., & Nunamaker, T. (2005). Dysfunctionality in performance measurement when outputs are difficult to measure: A research note. *Advances in Management Accounting*, 14, 277–289.
- Hanushek, E., & Ettema, E. (2015). *Defining Productivity in Education: Issues and Illustrations*. Hoover Institution Stanford, CA: Stanford University.
- Hanushek, E. (2010). *The economic value of higher teacher quality*. NBER Working Paper No. 16606.
- Hasnain, Z., Manning, N., & Pierskalla, J. (2012). *Performance-related Pay in the Public Sector: A Review of Theory and Evidence*. World Bank Policy Research Working Paper 6043. Retrieved 20 November 2018 from <https://openknowledge.worldbank.org/bitstream/handle/10986/6046/WPS6043.pdf?sequence=1&isAllowed=y>
- Hatry, H. (1975). Wrestling with police crime control productivity measurement. In Wolfe, J., & Heaphy, J. (Eds.), *Readings on productivity in policing*. Washington DC, United States: Police Foundation.
- Hauck, K., & Street, A. (2007). Do targets matter? A comparison of English and Welsh national health priorities. *Health Economics*, 16, 275–290.
- HIH [Health Innovation Hub]. (n.d.). *Our role*. Retrieved 20 November 2017 from <http://innovation.health.nz/about/our-role>
- Henri, J-F. (2006). Organizational culture and performance measurement systems. *Accounting, organizations & society*, 31(1), 77–103.
- Hitchener, S., & Gill, D. (2011a). The formal system as it evolved. In Gill, D. (Ed.), *The iron case recreated: the performance management of state organisations in New Zealand*. Wellington: Institute of Policy Studies, pp.81–118.
- Hitchener, S., & Gill, D. (2011b) The formal system: themes and conclusions In Gill, D. (Ed.), *The iron cage recreated: the performance management of state organisations in New Zealand*. Wellington: Institute of Policy Studies, pp.119–140.
- Hitlin, P. (2016). *What is mechanical turk?*. Pew Research Centre. Retrieved 20 November 2017 from [www.pewinternet.org/2016/07/11/what-is-mechanical-turk/](http://www.pewinternet.org/2016/07/11/what-is-mechanical-turk/)
- Ho, A. T. (2006). Accounting for the value of performance measurement from the perspective of city mayors. *Journal of Public Administration Research and Theory* 16, 217–37.
- Hood, C. (2002). Control, bargains, and cheating: The politics of public-service reform. *Journal of Public Administration Research and Theory*, 12(3), 309–332.
- Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–521.
- Hood, C. (2007). Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles? *Public Money and Management*, 27(2), 95–102.

- Horne, N. (2012). *The Commonwealth efficiency dividend: an overview*. Canberra, Australia: Parliament of Australia.
- Howell, M. (2016). *Measuring productivity in the education sector*. Unpublished research note. Productivity Commission, Wellington.
- Hvidman, U., & Andersen, S. (2014). The impact of performance management in public and private organizations. *Journal of Public Administration Research and Theory*, 24(1), 35–58.
- Insights MSD (iMSD) (2017). *Service Delivery Cost Allocation Model for Individual Outputs: 2017 version*. Wellington: Ministry of Social Development.
- Johansen, M., Taehee, K., & Zhu, L. (2016). Managing for results differently: Examining managers' purposive performance information use in public, nonprofit, and private organizations. *American Review of Public Administration*, 29(1), 1–20.
- Johansen, M., & Zhu, L. (2014). Market competition, political constraints, and managerial practice in public, nonprofit and private American hospitals. *Journal of Public Administration Research and Theory*, 24(1), 159–184.
- Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *The Phi Delta Kappan*, 81(3), 199–203.
- Kattel, R., Cepilovs, A., Drechsler, W., Kalvet, T., Lember, V., & Tönurist, P. (2014). *Can we measure public sector innovation? A literature review*. LIPSE Working papers (no. 2). Rotterdam, the Netherlands: Erasmus University Rotterdam.
- Knopf, E. (2017). *History of efficiency measurement by the New Zealand Public Health Sector: post 2000*. Wellington: New Zealand Productivity Commission.
- Kiwirail. (2016). *Annual Integrated Report 2016*. Retrieved 20 November 2017 from [www.kiwirail.co.nz/uploads/Publications/Annual%20Integrated%20Report%202016.pdf](http://www.kiwirail.co.nz/uploads/Publications/Annual%20Integrated%20Report%202016.pdf)
- Kroll, A. (2013). The other type of performance information: Nonroutine feedback, its relevance and use. *Public Administration Review*, 73(2), 265–276.
- Kroll, A. (2015). Drivers of performance information use: Systematic literature review and directions for future research. *Public Performance & Management Review*, 38(3), 459–486.
- Langford, P. H. (2009). Measuring organisational climate and employee engagement: Evidence for a 7 Ps model of work practices and outcomes. *Australian Journal of Psychology*, 61(4), 185–198.
- Lau, E., Lonti, S., & Schultz, R. (2017). Challenges in the measurement of public sector productivity in OECD countries. *International Productivity Monitor*, 32, 180–195.
- Le Grand, J. (2007). *The other invisible hand: Delivering public services through choice and competition*. Princeton, NJ, United States: Princeton University Press.
- Lewis, S., & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education*, 45(2), 245–264.
- Linna, P., Pekkola, S., Ukko, J., & Melkas, H. (2010). Defining and measuring productivity in the public sector: managerial perceptions. *International Journal of Public Sector Management*, 23(3), 300–320.
- Lonti, Z., & Gregory, R. (2007). Accountability or countability? Performance measurement in the New Zealand public service, 1992–2002. *Australian Journal of Public Administration*, 66(4), 468–484.
- Lovell, C., & Baker (2005). Experimental Estimates of Productivity Change in the Non-Market Sector: Preliminary Evidence from Government Services. Australian Productivity Commission: Internal Research Memorandum.

- Mansell, J., Laking, R., Matheson, B., & Light, R. (N.D.). *Data Commons Blueprint: A high trust, lower cost alternative to enable data integration and reuse*. Retrieved 9 December 2017 from <http://datacommons.org.nz>
- Marshall, M. (2009). Applying quality improvement approaches to healthcare. *BMJ: British Medical Journal*, 339(7725), 819-820.
- McCallion G., McKillop, D., Glass, J., & Kerr, C. (1999). Rationalizing Northern Ireland hospital services towards larger providers: Best-practice efficiency studies and current policy. *Public Money and Management* 19(2): 27.
- McClay, S., & Harrison, R. (2003). *The impact of schooling zoning on residential house prices in Christchurch*. Paper presented to the annual conference of the New Zealand Association of Economists, Christchurch, New Zealand.
- McKillop, D., Glass, C., Kerr, C., & McCallion, G. (1999). Efficiency in Northern Ireland hospitals: A non-parametric analysis. *The Economic and Social Review*, 30 (2), 175–196.
- McKinsey & Company. (2017). *Government productivity: unlocking the \$3.5 trillion opportunity*. Retrieved 20 November 2017 from [www.mckinsey.com/industries/public-sector/our-insights/the-opportunity-in-government-productivity](http://www.mckinsey.com/industries/public-sector/our-insights/the-opportunity-in-government-productivity)
- McLaughlin, C., & Coffey, S. (1992). Measuring productivity in services. In Lovelock, C. (Ed.), *Managing Services: Marketing, Operations, and Human Resources*, 2<sup>nd</sup> Edition. Englewood Cliffs, New Jersey, United States: Prentice-Hall.
- Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy and organization. *American Economic Review*, 80(3), 511-28.
- Ministry of Education. (2016). *Field of Specialisation for Students Gaining Qualifications from Tertiary Education Providers*. Retrieved 5 December 2007 from [www.educationcounts.govt.nz/statistics/tertiary-education/retention\\_and\\_achievement](http://www.educationcounts.govt.nz/statistics/tertiary-education/retention_and_achievement)
- Ministry of Health. (2016). *Regulatory impact statement: Therapeutic products regulation – replacement of the Medicines Act 1981 and the Medicines Regulations 1984 with a new legislative scheme for therapeutic products*. Retrieved 20 November 2017 from [www.health.govt.nz/system/files/documents/pages/regulatory-impact-statement-therapeutic-products-regulation.doc.docx](http://www.health.govt.nz/system/files/documents/pages/regulatory-impact-statement-therapeutic-products-regulation.doc.docx)
- Ministry of Health. (2017). *Annual report 2016/17*. Wellington: Author
- Ministry of Justice. (2017). *About the justice sector*. Retrieved 20 November 2017 from [www.justice.govt.nz/justice-sector-policy/about-the-justice-sector](http://www.justice.govt.nz/justice-sector-policy/about-the-justice-sector)
- Ministry of Social Development (2016). *Ministry of Social Development 2017 four year plan*. Retrieved 9 December 2017 from [www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/corporate/four-year-plan/2017-four-year-plan.pdf](http://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/corporate/four-year-plan/2017-four-year-plan.pdf)
- Ministry of Social Development. (2017). Privacy, human rights and ethics framework (unpublished internal document).
- Moore, S. & Hayward, M. (2017). *Productivity measurement case study: Ministry of Social Development*. New Zealand Productivity Commission Research Note 2017/07.
- Moynihan, D. P. (2008). *The dynamics of performance management: constructing information and reform*. Washington, D.C., United States: Georgetown University Press.
- Moynihan, D. P. (2013). Does public service motivation lead to budget maximization? Evidence from an experiment. *International Public Management Journal*, 16(2), 179–196.

- Moynihan, D. P. (2015). Uncovering the circumstances of performance information use: findings from an experiment. *Public Performance & Management Review*, 39(1), 33–57.
- Moynihan, D. P. (2016). Political use of performance data. *Public Money & Management*, 36(7), 479–481.
- Moynihan, D. P., & Beazley, I. (2016). *Toward next-generation performance budgeting: Lessons from the experiences of seven reforming countries*. Washington, DC., United States of America: World Bank.
- Moynihan, D. P., & Kroll, A. (2016). Performance management routines that work? An early assessment of the GPRA Modernization Act. *Public Administration Review*, 76(2), 314–323.
- Moynihan, D. P., & Landuyt, N. (2009). How do public organizations learn? Bridging structural and cultural divides. *Public Administration Review*, 69(6), 1097–1105.
- Moynihan, D. P., & Lavertu, S. (2012). “Does involvement in performance reforms encourage performance information use? Evaluating GPRA and PART.” *Public Administration Review* 72(4): 592–602.
- Moynihan, D. P., & Pandey, S. K. (2010). The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory*, 20(4), 849–866.
- Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2012). Prosocial values and performance management theory: Linking perceived social impact and performance information use. *Governance*, 25(3), 463–483.
- Moynihan, D. P., Pandey, S. K., & Wright, B. E. (2013). Transformational Leadership in the Public Sector. *Public administration reformation: Market demand from public organizations*, 18, 87.
- Mulgan, G, & Albury, D. (2003). *Innovation in the public sector*. London, United Kingdom: Cabinet Office Strategy Unit. Retrieved 20 March 2015 from [www.sba.oakland.edu/faculty/mathieson/mis524/resources/readings/innovation/innovation\\_in\\_the\\_public\\_sector.pdf](http://www.sba.oakland.edu/faculty/mathieson/mis524/resources/readings/innovation/innovation_in_the_public_sector.pdf)
- Murray, R. (1992). Measuring public-sector output: the Swedish report. In Griliches, Z. (Ed.), *Output measurement in the service sectors, Vol. 56*. Chicago, Illinois, United States: University of Chicago Press for the National Bureau of Economic Research.
- Murray, R. (2007). *Developing a quality adjusted output measure for the Scottish education system*. Paper presented at a Workshop on Measuring Education and Health Volume. National Accounts and Financial Statistics Division, Statistics Directorate, OECD, Paris.
- Murray, R. (2010). Measuring output from the public sector: A critical examination of the Atkinson Review. *Review of Income and Wealth*, 56(2), 413–423.
- National Audit Office. (2011). *PaceSetter: HMRC’s programme to improve business operations*. London, UK: The Stationery Office.
- New Zealand Police. (2016). *Four year plan 2015/16 – 2018/19*. Retrieved 10 November 2017 from [www.police.govt.nz/sites/default/files/publications/strategic-plan-2015-08-06-four-year-plan.pdf](http://www.police.govt.nz/sites/default/files/publications/strategic-plan-2015-08-06-four-year-plan.pdf)
- New Zealand Police. (2017). *New Zealand Police workplace survey results 2017: report of findings*. Retrieved 16 November 2017 from [www.police.govt.nz/about-us/publication/new-zealand-police-workplace-survey-results-2017](http://www.police.govt.nz/about-us/publication/new-zealand-police-workplace-survey-results-2017)
- New Zealand Treasury. (1994). *Improving output costing: guidelines and examples*. Retrieved 21 September 2017 from [www.treasury.govt.nz/publications/guidance/reporting/outputcosting/outputcosting.pdf](http://www.treasury.govt.nz/publications/guidance/reporting/outputcosting/outputcosting.pdf)
- New Zealand Treasury. (1996). *Putting it together: an explanatory guide to the New Zealand public sector financial management system*. Retrieved 21 September from [www.treasury.govt.nz/publications/guidance/publicfinance/pit/](http://www.treasury.govt.nz/publications/guidance/publicfinance/pit/)

- New Zealand Treasury. (2009). *New Zealand's long-term fiscal statement: Challenges and choices*. Retrieved 9 December 2017 from [www.treasury.govt.nz/publications/media-speeches/speeches/pdfs/sp-longterm-29oct09.pdf](http://www.treasury.govt.nz/publications/media-speeches/speeches/pdfs/sp-longterm-29oct09.pdf)
- New Zealand Treasury. (2011). *Administrative and support services benchmarking report for FY 2008/09 and FY 2009/10*. Retrieved 9 December 2017 from [www.treasury.govt.nz/statesector/performance/bass/benchmarking/2008-10/bnchmrk-rep.pdf](http://www.treasury.govt.nz/statesector/performance/bass/benchmarking/2008-10/bnchmrk-rep.pdf)
- New Zealand Treasury. (2015). *Guide to social cost benefit analysis*. Retrieved 9 December 2017 from [www.treasury.govt.nz/publications/guidance/planning/costbenefitanalysis/guide/cba-guide-jul15.pdf](http://www.treasury.govt.nz/publications/guidance/planning/costbenefitanalysis/guide/cba-guide-jul15.pdf)
- New Zealand Treasury. (2016). *He tirohanga mokopuna: 2016 statement on the long-term fiscal position*. Retrieved 9 December 2017 from [www.treasury.govt.nz/government/longterm/fiscalposition/2016/he-tirohanga-mokopuna/ltfs-16-htm.pdf](http://www.treasury.govt.nz/government/longterm/fiscalposition/2016/he-tirohanga-mokopuna/ltfs-16-htm.pdf)
- New Zealand Treasury. (2017). *The Estimates of Appropriations for the Government of New Zealand for the Year Ending 30 June 2018*. Retrieved 10 July 2017 from [www.treasury.govt.nz/budget/2017/estimates](http://www.treasury.govt.nz/budget/2017/estimates)
- New Zealand Treasury. (2017a). *District health board financial performance to 2016 and 2017 plans*. Retrieved 12 June 2017 from [www.treasury.govt.nz/publications/informationreleases/health/dhb-performance](http://www.treasury.govt.nz/publications/informationreleases/health/dhb-performance)
- New Zealand Treasury. (2017b). *New Zealand Gateway Reviews: Lessons learned report 2017*. Retrieved 20 November 2017 from [www.treasury.govt.nz/statesector/investmentmanagement/review/gateway/lessons/gateway-lessons-learned-report-jul17.pdf](http://www.treasury.govt.nz/statesector/investmentmanagement/review/gateway/lessons/gateway-lessons-learned-report-jul17.pdf)
- New Zealand Treasury. (2017c). *Head of the Government Finance Profession*, unpublished Powerpoint presentation, personal communication
- NHS Executive. (1999). *The NHS performance assessment framework*. Retrieved 20 November 2017 from <http://webarchive.nationalarchives.gov.uk/20031117050616/http://www.doh.gov.uk:80/pub/docs/doh/paf.pdf>
- Nielsen, P. A. (2014). Performance management, managerial authority, and public service performance. *Journal of Public Administration Research and Theory*, 24(2), 431–458.
- Nielsen, P. A., & Moynihan, D. P. (2016). Romanticizing bureaucratic leadership? The politics of how elected officials attribute responsibility for performance. *Governance*, 30(4), 541-559.
- Nielsen, P. A., & Moynihan, D. P. (2017). How do politicians attribute bureaucratic responsibility for performance? Negativity bias and interest group advocacy. *Journal of Public Administration Research and Theory*, 27(2), 269–283.
- Niskanen, W. (1973). *Bureaucracy: servant or master?* London, United Kingdom: Institute of Economics Affairs.
- Norman, R. (1998). *Re-invented government management innovations in New Zealand's public sector*. Wellington: Victoria University of Wellington.
- Norman, E. J. R. (2003). *Obedient servants? Management freedoms and accountabilities in the New Zealand public sector*. Wellington: Victoria University Press.
- Norman, R., & Gregory, R. (2003). Paradoxes and pendulum swings: performance management in New Zealand's public sector. *Australian Journal of Public Administration*, 62(4), 35-49.
- NZPC. (New Zealand Productivity Commission). (2014a). *Boosting productivity in the services sector*. Retrieved 1 July from [www.productivity.govt.nz/sites/default/files/services-inquiry-final-report.pdf](http://www.productivity.govt.nz/sites/default/files/services-inquiry-final-report.pdf)
- NZPC. (2014b). *Regulatory institutions and practices*. Retrieved 1 July from [www.productivity.govt.nz/sites/default/files/regulatory-institutions-and-practices-final-report.pdf](http://www.productivity.govt.nz/sites/default/files/regulatory-institutions-and-practices-final-report.pdf)

- NZPC. (2015a). *More effective social services, Appendix F: the economics of social services*. Retrieved 1 July from [www.productivity.govt.nz/sites/default/files/social-services-final-report-main.pdf](http://www.productivity.govt.nz/sites/default/files/social-services-final-report-main.pdf)
- NZPC (2015b). *More effective social services*. Retrieved 3 June from [www.productivity.govt.nz/sites/default/files/social-services-final-report-main.pdf](http://www.productivity.govt.nz/sites/default/files/social-services-final-report-main.pdf).
- NZPC. (2016). *Better urban planning*. Retrieved 24 June from [www.productivity.govt.nz/sites/default/files/MASTER%20COMPILED%20Better%20urban%20planning%20with%20corrections%20May%202017.pdf](http://www.productivity.govt.nz/sites/default/files/MASTER%20COMPILED%20Better%20urban%20planning%20with%20corrections%20May%202017.pdf).
- NZPC. (2017). *New models of tertiary education*. Retrieved 24 June from [www.productivity.govt.nz/sites/default/files/New%20models%20of%20tertiary%20education%20FINAL.pdf](http://www.productivity.govt.nz/sites/default/files/New%20models%20of%20tertiary%20education%20FINAL.pdf)
- O'Mahony, M., & Stevens, P. (2009). Output and productivity growth in the education sector: Comparisons for the US and UK. *Journal of Productivity Analysis*, 31:177–194.
- O'Mahony, M., & Stevens, P. (2006). International comparisons of output and productivity in public service provision: A review. In Boyne, G., Meier, K., O'Toole, L., & Walker, R. (Eds.) (2006), *Public sector performance: perspectives on measurement and management*. Cambridge: Cambridge University Press.
- OECD (2001a). *Measuring productivity (OECD Manual)*. Retrieved 20 November 2017 from [www.oecd.org/std/productivity-stats/2352458.pdf](http://www.oecd.org/std/productivity-stats/2352458.pdf)
- OECD. (2001b). *Specifying outputs in the public sector*. Paris, France: OECD Publishing, Paris, France.
- OECD (2009). *Measuring capital (OECD Manual)*. Retrieved 9 December 2017 from [www.oecd.org/publications/measuring-capital-oecd-manual-2009-9789264068476-en.htm](http://www.oecd.org/publications/measuring-capital-oecd-manual-2009-9789264068476-en.htm)
- OECD. (2010). *Improving health sector efficiency*. OECD Health Policy Studies. OECD Publishing, Paris, France.
- OECD. (2012). *Budgeting levers, strategic agility and the use of performance budgeting in 2011/12*. Retrieved 20 November 2017 from [www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/SBO\(2012\)10&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/SBO(2012)10&docLanguage=En)
- OECD. (2015). *OECD Recommendation of the Council on digital government strategies*. Retrieved 20 October 2017 from [www.oecd.org/gov/digital-government/recommendation-on-digital-government-strategies.htm](http://www.oecd.org/gov/digital-government/recommendation-on-digital-government-strategies.htm)
- OECD (2016a). *Engaging public employees for a high-performing civil service*. OECD publishing, Paris, France
- OECD (2016b) *Digital government strategies for transforming public services in the welfare areas*. (OECD Comparative Study) OECD publishing, Paris, France.
- OECD. (2017a, April). *Understanding public sector productivity*. Report to the 55<sup>th</sup> session of the Public Governance Committee. OECD publishing, Paris, France.
- OECD. (2017b). *Embracing innovation in government: global trends*. Retrieved 20 November 2017 from [www.oecd.org/gov/innovative-government/embracing-innovation-in-government.pdf](http://www.oecd.org/gov/innovative-government/embracing-innovation-in-government.pdf)
- Office of the Auditor-General. (2011). *Central government: cost effectiveness and improving annual reports*. Retrieved 9 December from <https://oag.govt.nz/2011/improving-annual-reports/docs/cost-effectiveness-and-improving-annual-reports.pdf/view>
- Office of National Statistics. (2012). *Public service productivity estimates: Education*. Retrieved 1 September 2017 from [www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity](http://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity)



- Office of the Deputy Prime Minister. (2005). *Measurement of output and productivity of the fire and rescue service. A conceptual framework*. Retrieved 7 September 2017 from [http://webarchive.nationalarchives.gov.uk/20070507024724/http://www.communities.gov.uk/pub/810/MeasurementofOutputandProductivityoftheFireandRescueServicePDF565Kb\\_id1123810.pdf](http://webarchive.nationalarchives.gov.uk/20070507024724/http://www.communities.gov.uk/pub/810/MeasurementofOutputandProductivityoftheFireandRescueServicePDF565Kb_id1123810.pdf)
- Office of the Minister of Internal Affairs. (2017). *ICT Functional leadership: Progress report October 2016 to March 2017*. Cabinet Paper. Retrieved 20 November 2017 from [www.ict.govt.nz/assets/Functional-Leadership/ICT-Functional-Leadership-Progress-Report-October-2016-to-March-2017-redctd.pdf](http://www.ict.govt.nz/assets/Functional-Leadership/ICT-Functional-Leadership-Progress-Report-October-2016-to-March-2017-redctd.pdf)
- Office of the Minister of Social Services. (2009). *Capping the size of core government administration and giving priority to front-line service delivery*. Retrieved 16 August 2017 from [www.ssc.govt.nz/sites/all/files/cabinet-paper-capping-core-govt-administration.pdf](http://www.ssc.govt.nz/sites/all/files/cabinet-paper-capping-core-govt-administration.pdf)
- Offices of the Deputy Prime Minister and the Minister of State Services. (2012). *Better public services paper 2: better system leadership*. Retrieved 16 August 2017 from [www.ssc.govt.nz/sites/all/files/bps-2306510.pdf](http://www.ssc.govt.nz/sites/all/files/bps-2306510.pdf)
- Ostrom, B. J., & Hanson, R. A. (1999). *Efficiency, timeliness, and quality: A new perspective from nine state criminal trial courts*. Williamsburg, VA: National Centre for State Courts.
- Palmer, G. (1988). Political perspectives. In Martin, J., & Harper, J. (Eds.) *Devolution and accountability*. Wellington: GP Books.
- Parker, D., Waller, K., & Xu, H. (2013). Private and public services: productivity and performance migration. *International Journal of Productivity and Performance Management*, 62(6), 652–664.
- Peacock, S., Chan, C., Mangolini, M., & Johansen, D. (2001). *Techniques for Measuring Efficiency in Health Services*. Australia Productivity Commission Working Paper. Retrieved 3 November 2017 from [www.pc.gov.au/research/supporting/measuring-health-services/tmeihs.pdf](http://www.pc.gov.au/research/supporting/measuring-health-services/tmeihs.pdf)
- Pekonen, P. (2011). *Lessons from the Finnish productivity programme*. Retrieved 12 October 2017 from [www.oecd.org/gov/budgeting/49041877.pdf](http://www.oecd.org/gov/budgeting/49041877.pdf)
- Petrie, M. (1998). *Organisational transformation: The Income Support experience*. Wellington: Department of Social Welfare.
- Petrie, M., & Webber, D. (2001). *Review of evidence on Broad outcome of public sector management regime*. Treasury working paper 01/06. Retrieved 20 November 2017 from [www.treasury.govt.nz/publications/research-policy/wp/2001/01-06/twp01-06.pdf](http://www.treasury.govt.nz/publications/research-policy/wp/2001/01-06/twp01-06.pdf)
- Pickens, D. (2017). *Efficiency and Performance in the New Zealand State Sector: Reflections of Senior State Sector Leaders*. Wellington, New Zealand: New Zealand Productivity Commission.
- Plimmer, G., Cantal, C., & Qumseya, T. (2017). *Staff perceptions of performance and effectiveness in the New Zealand State Sector: Further analysis of the 2016 Public Service Association survey on workplace dynamics*. Wellington, New Zealand: New Zealand Productivity Commission.
- Pontin, J. (2007). *Artificial intelligence, with help from the humans*. New York Times, 25 March 2007. Retrieved 20 November 2017 from [www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html](http://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html)
- Prime Minister (Helen Clark). (2006, May 10). *PlunketLine – Quality of Service* [New Zealand Parliamentary Debates]. 631 NZPD 2897.
- Pritchard, A. (2003). Understanding government output and productivity. *Economic Trends*, 596, 27–40.
- Public Policy Forum. (2014). *Changing the game: Public sector productivity in Canada*. Retrieved 10 November 2017 from [www.ppforum.ca/sites/default/files/Public\\_Sector\\_Productivity\\_01-07-15\\_1.pdf](http://www.ppforum.ca/sites/default/files/Public_Sector_Productivity_01-07-15_1.pdf)

- PWC. (2013). *Improving public sector productivity through prioritisation, measurement and alignment*. Retrieved 20 November 2017 from [www.pwc.com/gx/en/psrc/pdf/pwc-improving-public-sector-productivity-through-prioritisation-measurement-and-alignment.pdf](http://www.pwc.com/gx/en/psrc/pdf/pwc-improving-public-sector-productivity-through-prioritisation-measurement-and-alignment.pdf)
- Radnor, Z. (2010). Transferring Lean into government. *Journal of Manufacturing Technology Management*, 21 (3), 411-428.
- Rehm, M., & Filippova, O. (2008). The impact of geographically defined school zones on house prices in New Zealand. *International Journal of Housing Markets and Analysis*, 1(4):313–336.
- Robinson, M. (2013, May). *Spending reviews*. Paper prepared for 3<sup>rd</sup> annual meeting of OECD senior budget officials. Retrieved 20 November 2017 from [www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/SBO\(2013\)6&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PGC/SBO(2013)6&doclanguage=en)
- Rouse, P., & Swales, R. (2006). Pricing public health care services using DEA: Methodology versus politics. *Annals of Operation Research*, 145:265–280
- Ryall, T. (2009). *Government caps core government administration and pledges better performance*. Retrieved 16 August 2017 from [www.beehive.govt.nz/release/government-caps-core-government-administration-and-pledges-better-performance](http://www.beehive.govt.nz/release/government-caps-core-government-administration-and-pledges-better-performance)
- Sandiford, P., Consuelo, D., & Rouse, P. (2017). How efficient are New Zealand's district health boards at producing life expectancy gains for Māori and Europeans? *Australian and New Zealand Journal of Public Health*, 41 (2), 125–129.
- Sari, N., Rotter, T., Goodridge, D., Harrison, L., & Kinsman, L. (2017). *An economic analysis of a system wide Lean approach: cost estimates for the implementation of Lean in the Saskatchewan healthcare system for 2012-2014*. BMC Health Services Research Paper. Retrieved 28 November 2017 from <https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-2477-8>
- Schick, A. (1996). *The spirit of reform: managing the New Zealand state sector in a time of change*. Wellington: State Services Commission and New Zealand Treasury.
- Schick, A. (2003). *The performing state: reflection on an idea whose time has come but whose implementation has not*. Retrieved 20 November 2017 from [www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PUMA/SBO\(2003\)17&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/PUMA/SBO(2003)17&docLanguage=En)
- Schreyer, P. (2010). *Towards measuring the volume output of education and health services: A handbook*. OECD Statistics Working Papers 2010/02, Paris. Retrieved 20 November 2017 from <http://dx.doi.org/10.1787/5kmd34g1zk9x-en>
- Schreyer, P. (2012). Output, outcome and quality adjustment in measuring health and education services. *Review of Income and Wealth*, 58(2): 257–278.
- Scott, G. C. (2001). *Public management in New Zealand: lessons and challenges*. Wellington: New Zealand Business Roundtable.
- Scott, R., & Boyd, R. (2017). *Interagency performance targets: a case study of New Zealand's results programme*. Washington DC, United States of America: IBM Centre for the Business of Government
- Shapiro, I., Shapiro, M. D., & Wilcox, D. (2001). Measuring the value of cataract surgery. In Cutler D.M., & Berndt, E.R. (Eds.), *Medical care output and productivity*. Chicago, USA: University of Chicago Press, 411–438.
- Sharpe, A., Bradley, C., & Messinger, H. (2007). *The measurement of output and productivity in the health care sector in Canada: An overview*. CLS Research Report 2007–06. Ottawa: Centre for the Study of Living Standards.

- Simpson, H. (2009). Productivity in Public Services. *Journal of Economic Surveys*, 23(2), 250–276.
- Smith, A., & Norman, R. (1998) Change manager case study – George Hickton: salesman, in Norman, R., *Reinvented government – management innovations in New Zealand’s public sector*. Wellington: Victoria University of Wellington.
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International journal of public administration*, 18(2-3), 277–310.
- Smith, P. C., & Street, A. (2005). Measuring the efficiency of public services: the limits of analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(2), 401–417.
- Solow, R. (1987, July). We’d better watch out. *New York Times Book Review*, p.36.
- Spencer, K. (2014). *Accountability and performance monitoring of regulatory Crown entities*. Wellington: New Zealand Productivity Commission.
- SSC (State Services Commission). (1999a). *Looping the loop: evaluating outcomes and other risky feats*. Occasional paper No.7. Wellington: Author
- SSC. (1999b). *Essential Ingredients: Improving the Quality of Policy Advice*. Occasional paper No. 9. Wellington: author.
- SSC. (2013a). *Building capability within the State Services*. Retrieved 17 October 2017 from [www.ssc.govt.nz/building-capability-within-the-state-services](http://www.ssc.govt.nz/building-capability-within-the-state-services)
- SSC. (2013b). *Core guide 3: getting to great; lead reviewer insights from the Performance Improvement Framework*. Wellington: Author
- SSC. (2015). *Performance Improvement Framework and Lead Questions*. Retrieved 9 December 2017 from [www.ssc.govt.nz/pif-framework](http://www.ssc.govt.nz/pif-framework)
- SSC. (2016a). *Performance improvement framework: key changes for 2016*. Retrieved 17 November 2017 from <http://ssc.govt.nz/sites/all/files/PIF%20key%20changes%20for%202016.pdf>
- SSC. (2016b). *Remuneration*. Retrieved 20 November 2017 from [www.ssc.govt.nz/public-service-workforce-data/hrc-remuneration#wagemovements](http://www.ssc.govt.nz/public-service-workforce-data/hrc-remuneration#wagemovements)
- SSC. (2016c). *2017 four year plan guide*. Retrieved 20 November 2017 from [www.ssc.govt.nz/sites/all/files/2017-four-year-plan-guide-v2.pdf](http://www.ssc.govt.nz/sites/all/files/2017-four-year-plan-guide-v2.pdf)
- SSC. (2017a). *Systems thinking and continuous improvement*. Retrieved 20 November 2017 from [www.ssc.govt.nz/systems-thinking-and-continuous-improvement](http://www.ssc.govt.nz/systems-thinking-and-continuous-improvement)
- SSC and New Zealand Treasury. (2008). *Performance measurement. Advice and examples on how to develop effective frameworks*. Retrieved 7 July 2017 from [www.ssc.govt.nz/upload/downloadable\\_files/performance-measurement.pdf](http://www.ssc.govt.nz/upload/downloadable_files/performance-measurement.pdf)
- Statistics Finland. (2014). *Changes to the production of statistics and products in 2015*. Retrieved 10 October 2017 from [www.stat.fi/ajk/tiedotteet/2014/uutinen\\_040\\_2014-12-18\\_en.html](http://www.stat.fi/ajk/tiedotteet/2014/uutinen_040_2014-12-18_en.html)
- Statistics New Zealand. (2010). *Measuring government sector productivity in New Zealand: a feasibility study*. Retrieved 9 June 2017 from [www.stats.govt.nz/browse\\_for\\_stats/economic\\_indicators/productivity/measuring-govt-productivity.aspx](http://www.stats.govt.nz/browse_for_stats/economic_indicators/productivity/measuring-govt-productivity.aspx)
- Statistics New Zealand. (2013). *Education and health industry productivity, 1996–2011*. Retrieved 9 June 2017 from [http://archive.stats.govt.nz/browse\\_for\\_stats/economic\\_indicators/productivity/education-health-industry-productivity-1996-2011.aspx](http://archive.stats.govt.nz/browse_for_stats/economic_indicators/productivity/education-health-industry-productivity-1996-2011.aspx)

- Statistics New Zealand. (2014). *Productivity statistics: sources and methods (10<sup>th</sup> edition)*. Retrieved 5 December 2017 from [http://archive.stats.govt.nz/browse\\_for\\_stats/economic\\_indicators/productivity/productivity-stats-sources-methods-tenth-ed.aspx](http://archive.stats.govt.nz/browse_for_stats/economic_indicators/productivity/productivity-stats-sources-methods-tenth-ed.aspx)
- Statistics New Zealand. (2016). *Internet service provider survey: 2016*. Retrieved 10 November 2017 from [www.stats.govt.nz/browse\\_for\\_stats/industry\\_sectors/information\\_technology\\_and\\_communications/SPSurvey\\_HOTP2016.aspx](http://www.stats.govt.nz/browse_for_stats/industry_sectors/information_technology_and_communications/SPSurvey_HOTP2016.aspx)
- Statistics New Zealand. (2017a). *Productivity Statistics: 1978–2016*. Retrieved 9 June 2017 from [www.stats.govt.nz/browse\\_for\\_stats/economic\\_indicators/productivity/ProductivityStatistics\\_HOTP78-16.aspx](http://www.stats.govt.nz/browse_for_stats/economic_indicators/productivity/ProductivityStatistics_HOTP78-16.aspx)
- Statistics New Zealand. (2017b). *Integrated Data Infrastructure*. Retrieved 5 December 2017 from [www.stats.govt.nz/browse\\_for\\_stats/snapshots-of-nz/integrated-data-infrastructure.aspx](http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx)
- Statistics New Zealand. (n.d.). *Consumers price index visualisation*. Retrieved 10 November 2017 from [www.stats.govt.nz/datavisualisation/cpi.html](http://www.stats.govt.nz/datavisualisation/cpi.html)
- SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) (1997). *Data envelopment analysis: A technique for measuring the efficiency of government service delivery*. Retrieved 28 November 2017 from [www.pc.gov.au/research/supporting/data-envelopment-analysis/dea.pdf](http://www.pc.gov.au/research/supporting/data-envelopment-analysis/dea.pdf)
- Superu. (2017). *Making sense of evaluation: A handbook for the social sector everyone*. Retrieved 9 December 2017 from [www.superu.govt.nz/sites/default/files/Publications/Evaluation%20Handbook%20Dec%202017.pdf](http://www.superu.govt.nz/sites/default/files/Publications/Evaluation%20Handbook%20Dec%202017.pdf)
- Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, 10(2-3), 211–227.
- Syverson, C. (2016). *Challenges to mismeasurement explanations for the US productivity slowdown*. (No. w21974). Cambridge, Massachusetts, USA: National Bureau of Economic Research.
- Taylor, J. (2011). Factors influencing the use of performance information for decision making in Australian state agencies. *Public Administration*, 89(4), 1316–1334.
- Technical Advisory Services (TAS). (2017, October). *2017 CCDM Overview*. Presentation to Productivity Commission, Wellington, New Zealand.
- TAS. (n.d). *Care capacity demand management*. Retrieved 27 November 2017 from <https://tas.health.nz/strategic-workforce-services/safe-staffing-health-workplace/care-capacity-demand-management>.
- Ter Bogt, H. J. (2004). Politicians in search of performance information? Survey research on Dutch aldermen's use of performance information. *Financial Accountability & Management*, 20(3), 221–252.
- Tipper, A. (2013, July). *Output and productivity in the education and health industries*. Paper presented at the 54th Conference of the New Zealand Association of Economists, Wellington, New Zealand.
- Topp, V., Soames, L., Parham, D., & Bloch, H. (2008). *Productivity in the mining industry: measurement and interpretation*. Canberra, Australia: Australian Productivity Commission.
- Tullock, G. (1976). *The Vote Motive*. London, United Kingdom: Institute of Economic Affairs.
- United Nations, European Communities, Organisation for Economic Co-operation and Development, International Monetary Fund, World Bank (2009). *System of national accounts*. Retrieved 11 December 2017 from <https://unstats.un.org/unsd/nationalaccount/docs/sna2008.pdf>

- Van Dooren, W., Bouckaert, G., & Halligan, J. (2015). *Performance management in the public sector*. Oxford, UK: Routledge.
- Van Dooren, W., Manning, N., Malinska, J., Kraan, D. J., Sterck, M., & Bouckaert, G. (2006). *Issues in output measurement for "Government at a glance"*. Retrieved 28 August 2017 from [www.oecd.org/std/na/37417214.pdf](http://www.oecd.org/std/na/37417214.pdf)
- Van Dooren, W., Lonti, Z., Sterck, M., & Bouckaert, G. (2008). *Institutional Drivers of Efficiency in the Public Sector*, OECD.
- Verbeeten, F.H.M. (2008). Performance management practices in public sector organizations: impact on performance. *Accounting, auditing & accountability journal*, 21(3), 427–454.
- Victoria University of Wellington, School of Government. (2017). *Independent review of the performance improvement framework*, June 2017, online, ISBN: 978-0-478-43481-1.
- Whitcombe, J. (2008). Contributions and challenges of 'new public management' New Zealand since 1984. *Policy Quarterly*, 4(3), 7–13.
- Whitehead, J. (2006, September). *The imperative for performance in the public sector*. Speech to public sector governance seminar, APEC Vietnam, Da Nang. Retrieved 19 September 2017 from [www.treasury.govt.nz/publications/media-speeches/speeches/pdfs/spch-10sep06.pdf](http://www.treasury.govt.nz/publications/media-speeches/speeches/pdfs/spch-10sep06.pdf)
- Work Foundation. (2014). *A review of the evidence on the impact, effectiveness and value for money of performance-related pay in the public sector*. Retrieved 20 November 2017 from [www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/381600/PRP\\_final\\_report\\_TW\\_F\\_Nov\\_2014.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/381600/PRP_final_report_TW_F_Nov_2014.pdf)
- World Bank. (2015). *E-Gov research and resources*. Retrieved 20 October 2017 from [www.worldbank.org/en/topic/ict/brief/e-gov-resources#egov](http://www.worldbank.org/en/topic/ict/brief/e-gov-resources#egov)
- Wu, T., Chen, M., & Yeh, J. (2010). Measuring the performance of police forces in Taiwan using data envelopment analysis. *Evaluation and Program Planning*, 33 (3), 246–254.
- Yang, K., & Hsieh, J. Y. (2007). Managerial effectiveness of government performance measurement: testing a middle-range model. *Public Administration Review*, 67(5), 861–879.
- Zhang, H., Van de Walle, S., & Zhuo, Y. (2016). Does trust in the performance measurement organization influence how public managers use performance information? *Public Performance & Management Review*, 40(2), 409–430.

NEW ZEALAND  
PRODUCTIVITY COMMISSION  
Te Kōmihana Whai Hua o Aotearoa

