

Costs and benefits of public sector data provision

John Houghton

**Centre for Strategic Economic Studies
Victoria University, Australia**

John.Houghton@pobox.com



Presented by Greg Laughlin, Principal Policy Advisor, ANDS

Costs and benefits of data provision

Public Sector Information (PSI)

- Public Sector Information (PSI) is any kind of information that is produced or collected by a public agency as a part of its mandated role.
- PSI may be the basis for industries that use or reuse the raw data to produce knowledge-intensive products and services.
- Case studies exploring the costs and benefits that PSI **agencies** and their **users** experience in making information freely available, and preliminary estimates of the **wider economic impacts** of open access to PSI.

- The aim of this study was to explore costs and benefits via three case studies:

The Australian Bureau of Statistics (national statistics);
GeoScience Australia (fundamental spatial data); and
The National Water Commission (hydrological data).

Costs and benefits of data provision

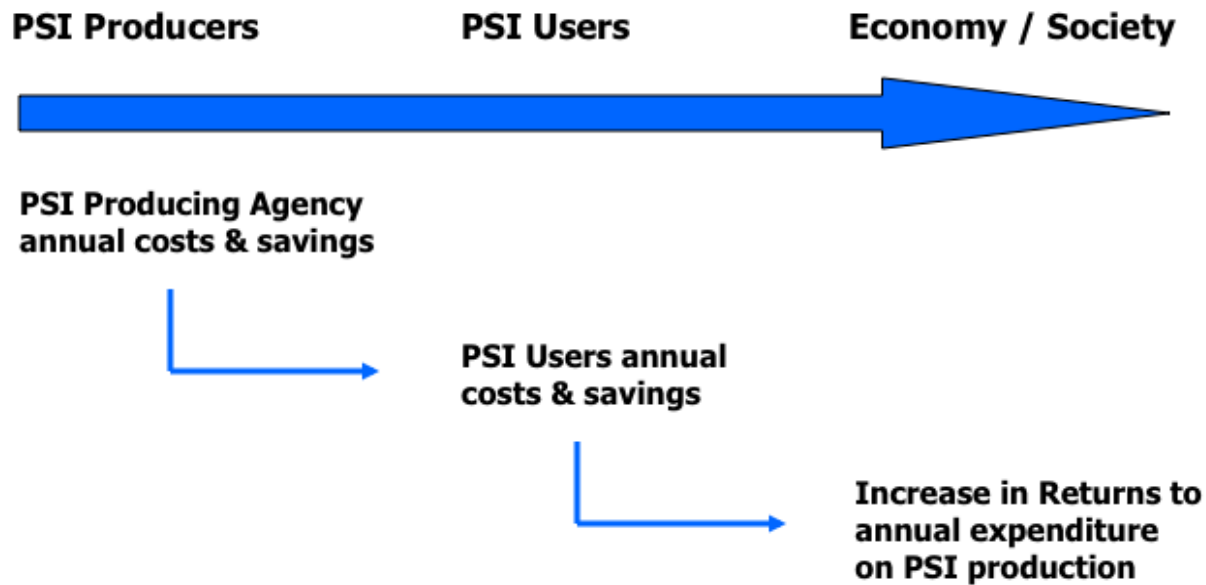
Public Sector Information (PSI)

- Presentation outlines a method for cost-benefit analysis at the agency level*, and explores the data requirements.

- Focuses on three main elements:
 1. The costs and savings experienced by PSI producing **agencies** involved in the provision of free and open access to information;
 2. The costs and savings experienced by the **users** of PSI in accessing, using and re-using the information; and
 3. The potential **wider economic** and social impacts of freely accessible PSI.

* Need time series, before/after 'freely available'

A framework for estimating cost-benefit: producers, users and the wider economy



A framework for estimating cost-benefit: producers, users and the wider economy

$$\frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Agency savings} + \text{Users' savings} + \text{Increased returns to annual expenditure on PSI production}}{\text{Agency costs} + \text{Users' costs}}$$

$$\text{ABS case}^* = \frac{\text{Cost of producing products (ABS)} + \text{Users' savings on products (now free)} + \text{Economy}}{\text{Lost revenue (ABS)}}$$

* Similar variables elsewhere

Overall, for the ABS

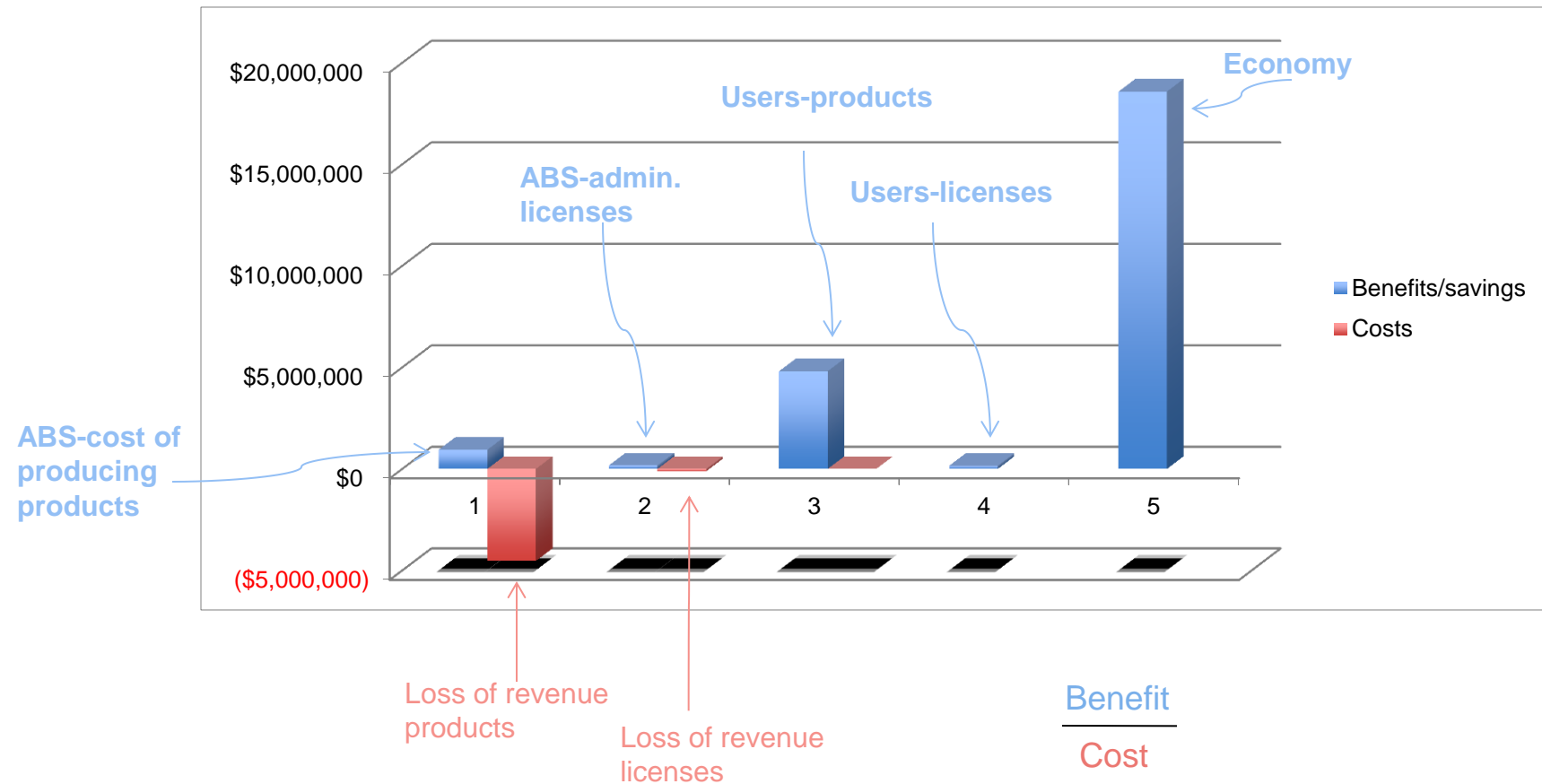
$$\frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Agency savings} + \text{Users' savings} + \text{Increased returns to annual expenditure on PSI production}}{\text{Agency costs} + \text{Users' costs}}$$

$$\frac{\$24,666,100}{\$4,645,300} = \frac{(\$943,800 + \$171,700) + (\$4,810,600 + \$156,700) + (\$18,583,300)}{(\$4,530,300 + \$115,000) + (\$0)}$$

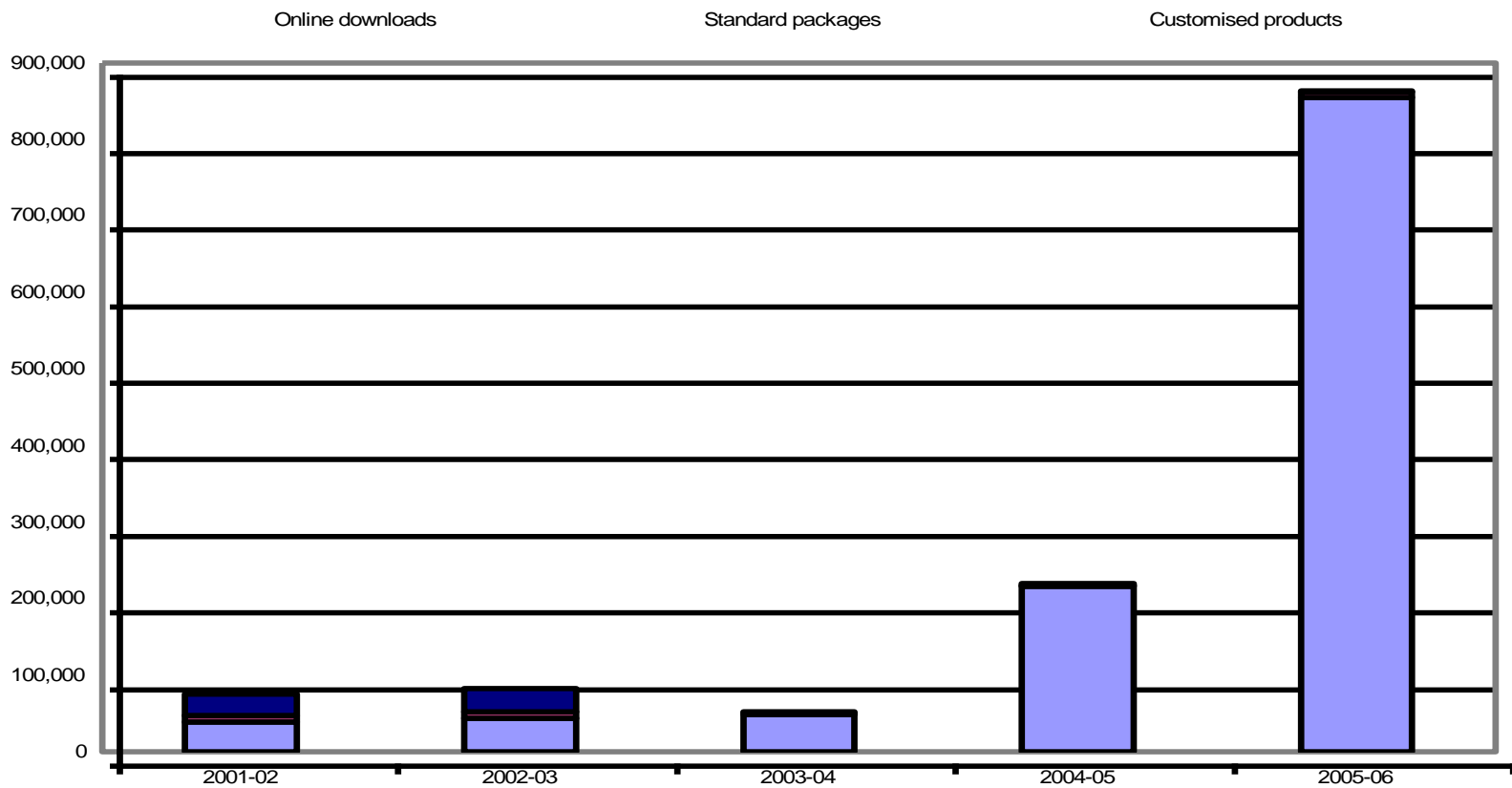
= 5.3* and ongoing

* Estimate (ignore apparent precision in \$)

Overall, for the ABS



Fundamental spatial data: scheduled datasets delivered by government agencies, GA and others



Fundamental spatial data

Annual costs and benefits, circa 2005-06

<i>Item</i>	<i>Value</i>
Agency and user impacts:	
Agency impacts (net)	-923,900
Users impacts (net)	1,676,100
Welfare estimates:	
Surplus from lost revenue (average annual 2001-02 to 2005-06)	2,165,000
Returns estimates:	
Increase in returns (downloads per page published 3yr 2003-04 to 2005-06 and 2006-07 to 2008-09)	15,496,200
Overall impacts estimates:	
Estimated total costs (revenue foregone)	1,300,000
Estimated total benefits (savings and returns)	17,548,400
Benefit/cost ratio	13

Source: Various sources. Author's analysis.

Ratios cannot be compared across agencies or data types

Hydrological data

Insufficient data for a complete case study

- Hydrological data has been fragmented and incomplete, and it is too early to measure the impacts of making National Water Account data freely available centrally.
- However, the *Victorian Water Resources Data Warehouse* provides an example of the cost and use impacts of making water data freely available online.
- Looking at cost per use since the late 1990s, it is clear that Victorian water data use increased and provision costs per use have fallen dramatically.

Benefits outweigh the costs:

- Even the subset of benefits that can be measured outweigh the costs of making PSI more freely and openly available.
- It is not simply about access prices, but also about the transaction costs involved.
- Creative Commons licenses, ready discoverability and data standards are crucial in enabling access that is truly open (*i.e.* free, immediate and unrestricted).

- The direct and measurable benefits of making these forms of PSI freely available outweigh the costs.
- Ongoing (not one-offs!) ***
- The wider benefits are more difficult to measure, but are likely to be substantial.
- Adding the longer-term benefits that we cannot measure or foresee, the case for more open access to these more commercial forms of PSI is strong.

- Individual cases vary greatly, making generalisation extremely difficult.
- What can be generalised is the methods of analysis. For example, it would be useful to:
 - Further develop the framework for estimating cost-benefits outlined in this study to produce a tool tailored to the analysis of the impacts of **making PSI freely available**; and
 - Combine the frameworks and models into a similar tool tailored specifically to estimating the impacts of **research data curation and sharing**.

This project was commissioned by the Australian National Data Service (ANDS). ANDS is supported by the Australian Government (DIISR) through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative.

Full study available at ands.org.au/working-with-data/articulating-the-value-of-open-data/costs-and-benefits-of-data-provision