



Glen Slater
eResearch 2020

By email: ge.slater@gmail.com

16 March 2016

Economic assessment of the value of open research data

Dear Glen

Thank you for engaging PwC to assist with developing the economic impacts analysis for eResearch 2020. This letter provides initial findings from our analysis, in accordance with our letter of engagement dated 06 October 2015. The purpose of the engagement was to calculate a potential economic value for the proposed National Research Data Programme, building on similar overseas analysis. The results should be read in conjunction with the disclaimer in Appendix A to this letter.

Method in Houghton & Gruen (2014)

You provided us a report by Houghton & Gruen, 'Open Research Data: Report to the Australian National Data Service', dated November 2014 (hereafter, HG). We reconstructed the calculations in that report in order to develop the method for assessing impacts in New Zealand.

- The report does not contain sufficient detail to replicate fully the calculations they made. We have therefore used our judgement to replicate some of the calculations.
- We have developed calculations that match the step-by-step process in section 3 of the report. Our method generally ties, with our final result being 8% higher than the reported result in HG.
- The Solow-Swan model referenced in section 3 and the report appendix is not used to make calculations in HG. We have assured ourselves of this by replicating exactly the calculations in Houghton & Sheehan (2009), the cited source.

As a result of replicating the HG calculations, we are confident that we can apply the same method to New Zealand data, within a reasonable level of materiality.

Results with New Zealand data

The results from this sort of calculation depend on the assumptions used. The assumptions for our calculations are as follows:

Table 1 Assumptions for analysis

Parameter	Value	Source
Value of New Zealand government science spending	\$1,410	MBIE
Labour cost as share of science spending	48%	HG
Return on investment	40%	HG
Time spent working with data	46%	HG
Share of research with domestic impact (localisation)	66%	HG
Depreciation rate	10%	HG
Useful life of data	10 years	HG
Discount rate	8%	New Zealand Treasury
Average time savings from data repositories	37%	HG
Respondents who could not have obtained data otherwise	52%	HG
Current data curation (Low assumption)	10%	HG
Current data curation (High assumption)	20%	HG

As we discussed, there are other possible assumptions that one could make for these parameters. The two in particular to which I would draw your attention are the following:

- Return on investment – the Treasury report by Hall & Scobie (2006) found a 17% return on investment to research spend. Given the wide range of prior estimates of returns and the time lag structure in their model, this value is not likely to have changed significantly since publication. This is a detailed report and there are several caveats to that RoI. However, it is a well-known piece of New Zealand research.
- Localisation – the results in Coe & Helpman (1995) suggest that this could be 55% in New Zealand.

Using New Zealand inputs to the HG calculations, we calculated a number of values and impacts:

- The localised (domestic) benefits of data produced annually in New Zealand would be \$283m to \$589m. These figures represent the net present value (NPV) over 10 years from one year's current investment in data production.
- The cost savings from increased research efficiency would be \$76m to \$158m annually. The increase in localised benefits from greater efficiency, calculated as a return on investment, would be in the order of \$105m to \$218m. The costs savings plus the return on investment, what HG refer to as the 'total impact' (p. 20), would be \$181m to \$376m. This is the net present value over 10 years, and accounting for the localisation effect.
- The localised value of reusing data would be \$54m to \$113m in impacts. This value is based on returns to research over a 10-year period.

- The total NPV of the localised economic impact of one year of good data management, following the HG method, would be \$235m to \$490m.
- Assuming that 20% of this data curation activity is already undertaken, the potential benefit of better data curation would be \$188m to \$392m.

Please note that these economic impacts or benefits represent the discounted stream of benefits over 10 years, from making improvements to just a single year's investment in data activity.

For convenience, the data are summarised in the table below.

Table 2 Summary of results

	Labour spending only (\$)	Total spending (\$)
Domestic value of data produced annually in NZ	283	589
Value from cost savings	76	158
Value from greater efficiency	105	218
Cost savings + greater efficiency	181	376
Value of reusing data	54	113
Total additional impact	235	490
Assume 80% benefit	188	392

As we have discussed, there is the possibility that the HG method does not fit the Cost Benefit Analysis method preferred by the New Zealand Treasury. Our modification to the HG method treats the value of better management as the increased return on research efficiency, without also adding in the cost savings. Using the assumptions in Table 1, including an RoI of 40% and localisation factor of 66%, the alternative method produces a final potential benefit of better data management of \$159m to \$331m (compared to \$188m to \$392m, above).

A more conservative approach is to use the alternative values for New Zealand: 17% return on research and 55% localisation parameter. The value of one year's improvement in data management, in that case, would be \$56m to \$117m.

In addition, the value of a better data management programme would extend over time. The annual values presented above are a snapshot, but the change over time is important. We have calculated the value of a programme that lifted good data management from 20% of the sector to 80% over 10 years, using the same 8% discount rate:

- Based on the total economic impact of \$235m to \$490m, the value of that 10-year improvement would be \$461m to \$961.
- Based on the most conservative set of assumptions, which produced a total economic impact of \$56m to \$117m, the value of the 10-year improvement would be \$110m to \$230m.



Glen Slater
16 March 2016

We hope that this information is helpful to you. We would be happy to run through our calculations with you, and we can also model a few different scenarios. Understanding both optimistic and conservative scenarios might be useful in communicating with different audiences in the future.

Yours sincerely

A handwritten signature in blue ink, appearing to read "W. Kaye-Blake", with a long horizontal flourish extending to the right.

Dr Bill Kaye-Blake
Director
Consulting

Appendix A – Disclaimer

This letter has been prepared solely for the purposes stated therein and should not be relied upon for any other purpose. We accept no liability to any party should it be used for any purpose other than that for which it was prepared.

To the fullest extent permitted by law, PwC accepts no duty of care to any third party in connection with the provision of this report and/or any related information or explanation (together, the “Information”). Accordingly, regardless of the form of action, whether in contract, tort (including without limitation, negligence) or otherwise, and to the extent permitted by applicable law, PwC accepts no liability of any kind to any third party and disclaims all responsibility for the consequences of any third party acting or refraining to act in reliance on the Information.

We have not independently verified the accuracy of information provided to us, and have not conducted any form of audit in respect of the organisation for which work is completed. Accordingly, we express no opinion on the reliability, accuracy, or completeness of the information provided to us and upon which we have relied. Responsibility for the reliability, accuracy and completeness of such information therefore remains with eResearch 2020.

The statements and opinions expressed herein have been made in good faith, and on the basis that all information relied upon is true and accurate in all material respects, and not misleading by reason of omission or otherwise.

The statements and opinions expressed in this report are based on information available as at the date of the report.

This modelling underlying this report has been tested by PricewaterhouseCoopers for logic and consistency, but has not been reviewed by an independent party, and we are unable to guarantee that the modelling is 100 per cent free of errors. We reserve the right, but will be under no obligation, to review or amend our report, if any additional information, which was in existence on the date of this report was not brought to our attention, or subsequently comes to light.

PricewaterhouseCoopers accepts no liability for the performance of the indicative results of this report, or any of the consequences of eResearch 2020’s commercial decisions.

This report is issued pursuant to the terms and conditions set out in our contract dated 06 October 2015.