Data Driven Innovation

IN NEW ZEALAND

covec
sapere research group
Innovation Partnership

Table of Contents

Sapere Research Group is one of the largest expert consulting firms in Australasia and a leader in provision of independent economic, forensic accounting and public policy services. Sapere provides independent expert testimony, strategic advisory services, data analytics and other advice to Australasia's private sector corporate clients, major law firms, government agencies and regulatory bodies.

Hayden Glass is a Principal with Sapere who specialises in regulatory and competition economics and advocacy, particularly in the telecommunications and Internet markets. He has a background in law, economics, and commercial strategy, extended experience in public policy debate on telecommunications issues, and a particular interest in the commercial and economic impacts of the Internet.

Anna Livesey is a Senior Managing Consultant with Sapere who works in complex policy areas including telecommunications, cybersecurity, local government, housing, machinery of government and education. She has experience across the public policy cycle including policy development, stakeholder consultation, business case preparation, performance audit and evaluation.

Preston Davies is a Principal with Sapere and an economist with wide interests including public policy, strategy, economic development and transport. Key areas where Preston has provided advice include Maori economic development, land transport, airport pricing, regional development, competition in the banking industry, regulatory reform, economic performance, economic impact assessment, electricity demand, biofuels policy, education, and health labour markets.

Covec is an economics consultancy that provides analysis and advice to policy agencies, local government and a range of private sector clients to assist in complex decisions.

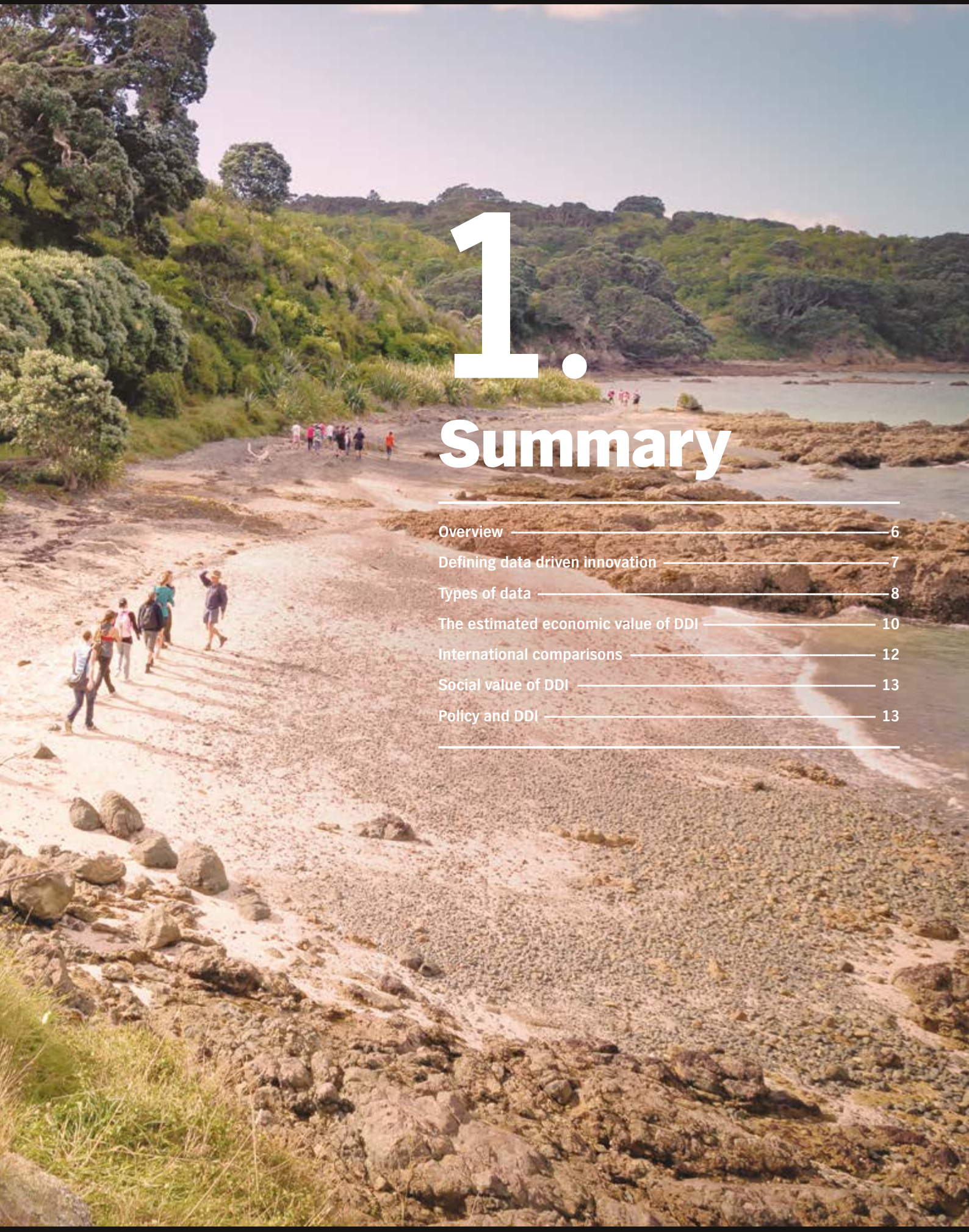
Aaron Schiff is a consulting economist who helps organisations to understand complex issues and make better decisions through the use of data analysis, economic modelling, and data visualisation. He has particular expertise in the economics of networks and innovation.

This report was commissioned by Innovation Partnership members;
Business New Zealand, Chorus, Crown Fibre Holdings, Google New Zealand, Retailers Association, Internet New Zealand and Xero.

The views in this report are those of the authors, not the funders.
Any errors or omissions remain our responsibility.

Counting the environmental cost

Companies are reducing their carbon footprints by up to 10% thanks to using the FoundationFootprint platform. Where previously understanding a company's impact on the environment took multiple spreadsheets and an immense time commitment to understand, a company can understand its energy efficiency and environmental impact in a simple to use and timely way. The powerful system combines information from multiple sources such as energy use, vehicle mileage, paper consumption and flight miles to create two key measures: dollars spent and carbon produced.



1. Summary

Overview

Data driven innovation (DDI) represents a multi billion dollar opportunity for New Zealand. In 2014 we estimate that New Zealanders shared $2.4 billion of value by harnessing data to make smarter decisions and develop new products and services.

Consumers enjoyed substantial benefits, with $1.3 billion accruing to them in the form of lower prices. Businesses obtained $1.1 billion of value by using data to drive higher revenue, cut costs and streamline operations.

The data revolution is also empowering the public sector to deliver better services through improved targeting of limited public resources and smarter policies. From health to education, data driven innovation can generate tremendous social value.

Take Farmax, a new data tool for farmers that predicts budgets and profits by modelling everything from grass cover to stock performance. Farmax estimates that using its platform to make better decisions has helped to contribute $450 million to bottom lines across its user base.

Or Christchurch Hospital, where a new data-centric approach to patient management called HealthOne has helped to reduce patient acute medical admission rates to 30% below the national average, while improving health outcomes.

Data driven innovation benefits business, consumers, government and taxpayers. But New Zealand is just getting started. We estimate that data driven innovation could easily deliver $4.5 billion in benefits within five years with higher uptake from business and government.

To see the potential, we need only look to Australia where the impact of data driven innovation on the economy is already three times greater than ours in percentage terms. Adoption is the key reason for this difference – if organisations in New Zealand adopted data driven innovation at the same rate as their Australian counterparts, the value we currently generate would jump from $2.4 billion to $4.8 billion.

The benefits of higher uptake of data driven innovation would be greater competitive advantage in key industry sectors, new, improved and customised goods and services, and greater well-being through better outcomes across the public sector.

Whether we seize this opportunity and harness the potential of the data revolution is up to us. The government can also offer encouragement with smart regulation, support for open data initiatives and by demonstrating practical benefits of DDI. This report sets out the challenges ahead, and points the way forward to a data driven future that fuels growth and creates value for all New Zealanders.

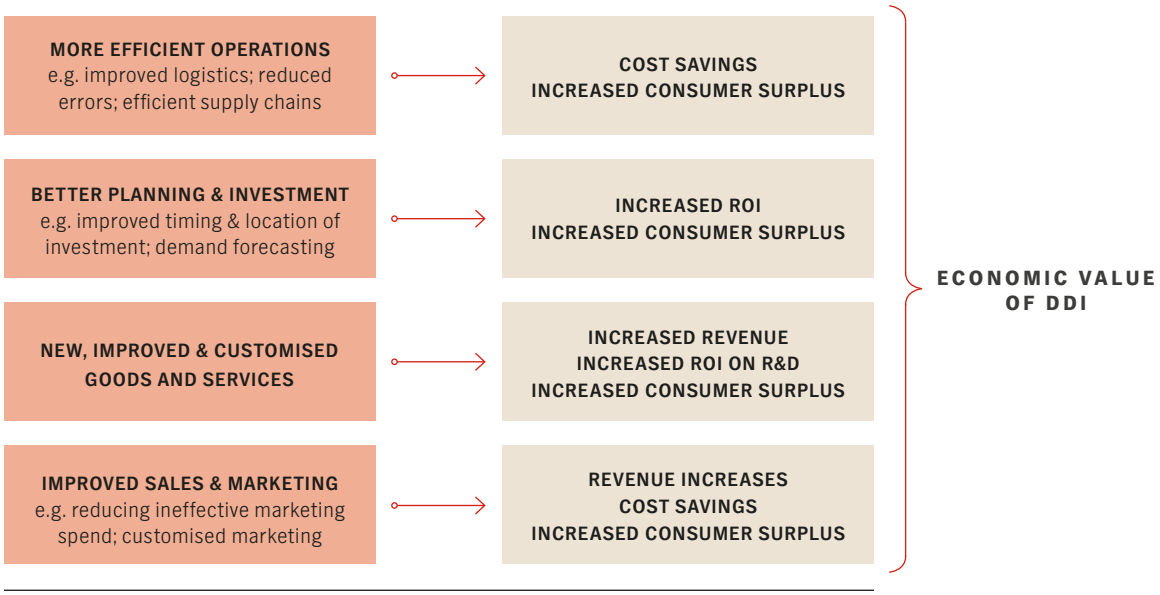
Defining data driven innovation

DDI is the innovation and consequent economic and social value that arises from the use of data analysis by private and public sector organisations to make better decisions and create new products and services. DDI can support a very wide range of innovations, including improving firm operational efficiency, developing new products and services, making better investment and strategic decisions, and more effective government interventions.

Figure 1 illustrates the main sources of economic value arising from DDI. Social value is also significant and can arise from better policy design and improved targeting of limited government resources. Together with positive spillovers from DDI in the private sector, DDI can generate significant social value in addition to economic value.

FIGURE 1

ECONOMIC LINKAGES AND VALUE OF DDI



SOURCE: SAPERE AND COVEC

Data and analytics are being used by businesses and government departments in a wide range of innovative ways. Some local examples include:

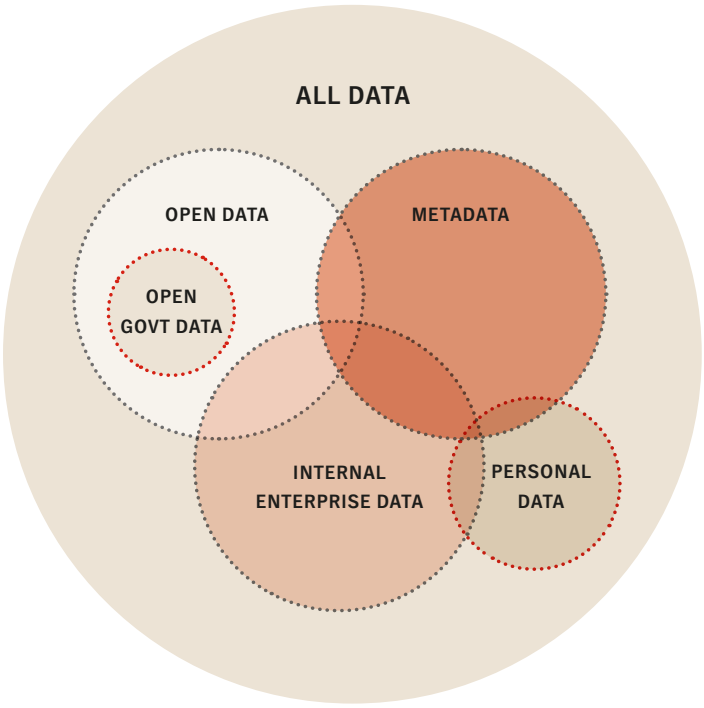
- Retailers are using sales data and cloud-based point-of-sale software to understand the popularity of different products and ensure that they have the right levels of stock at the right times.
- The New Zealand Insurance Claims Register is a database of all claims lodged with participating insurance companies. This can be used to assess the validity of a claim and/or decide whether or not to offer cover. The register is also used to detect insurance fraud, and enable better pricing of insured risks.

- Frontline health providers such as primary healthcare organisations analyse data to proactively determine the treatment needs of their patients and work to prevent the need for more costly and traumatic treatments including hospital admission.
- Logistics and transport operators use data on road and traffic conditions to optimise their operations and routes. Operators can determine the best route for pick-up and delivery given traffic conditions, road repair work and volumes to be collected. This reduces transport times and vehicle emissions. They can also collect performance data to benchmark the performance of the truck fleet and drivers.

Types of data

In general terms, data is structured information where the structure is imposed through data definitions that specify what is being measured and qualitative characteristics of the measurements. We discuss below some types of data, which are illustrated conceptually in Figure 2.

FIGURE 2
TYPES OF DATA



SOURCE: ADAPTED FROM PWC (2014)

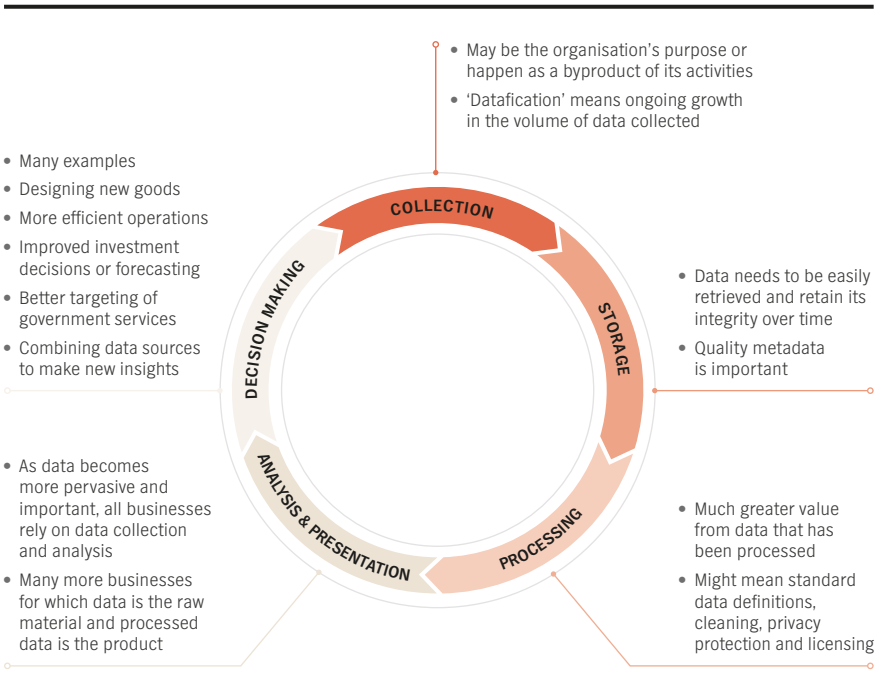
Datasets do not need to be large to be useful - a fact that is particularly relevant given the size of many of the datasets available to New Zealand organisations. Depending on the context, a few dozen data points may be sufficient to reveal general trends and patterns, although larger datasets may be needed for more sophisticated analysis. The real issue is not the size of the dataset, but smart innovative use of the data that is available, applied to a particular problem.

For example, retailers can use information published for free by Statistics New Zealand to understand characteristics of people and households in their local area. Tourism operators can use information on international visitors to understand who visits, from where and how much they spend. Paymark offers a service to any business using its EFTPOS network that allows comparison of turnover performance to other businesses that sell similar things or are in the same area. This report includes six case studies of innovative uses of data in New Zealand.

Value is created from data via a number of processes in a data "value cycle" (Figure 3). Data must first be collected or acquired and stored in suitable formats before it can be used. Most of the value of data then arises from its processing and use in analysis and presentation. The use of data may be as an input to decision-making, or as a product or service that is largely comprised of data or involves data analysis (e.g. forecasting, or business intelligence services). One important characteristic of data is that it is not used up when it is used; the same data can be useful to many people for many different purposes, regardless of who originally collected it.

FIGURE 3

THE DDI VALUE CYCLE



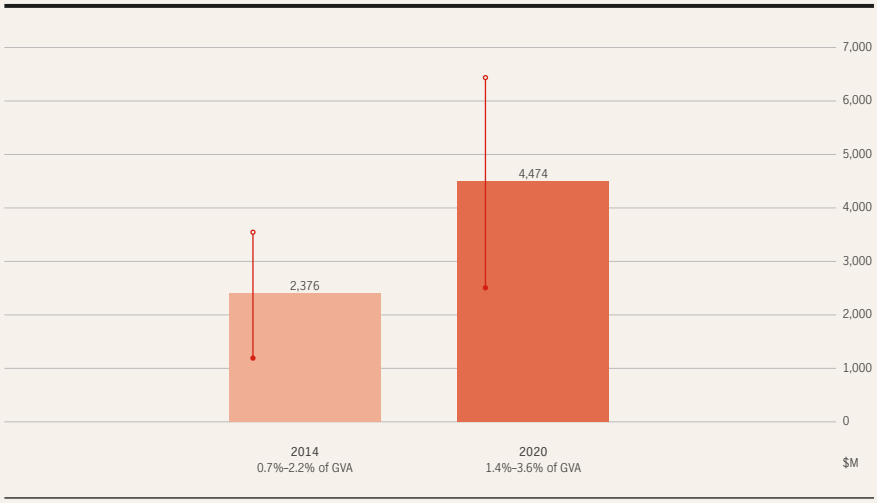
SOURCE: SAPERE AND COVEC, OECD (2014)

The estimated economic value of DDI

In New Zealand we estimate that DDI contributed $2.4 billion to gross value added (GVA) in 2014 (Figure 4). This represents 1.4% of total economic activity across seven major industry sectors. By 2020, with higher adoption of DDI by businesses, we project this to grow to $4.5 billion GVA in real terms, or 2.5% of projected economic activity across the seven sectors. In dollar terms, we estimate the largest impacts of DDI in New Zealand are in the transport and logistics, retail, and other services sectors (Figure 5). In part this reflects relatively high estimated rates of DDI uptake in these sectors, as well as cost savings and revenue opportunities for the adoption of DDI. Within government, DDI can also create significant social value, helping to design better policy interventions and target public resources to where they are most needed.

FIGURE 4

ESTIMATED TOTAL IMPACTS OF DDI ON GVA IN NEW ZEALAND



SOURCE: SAPERE & COVEC

Our estimates are based on modelling the impacts of DDI on GVA that arise from:

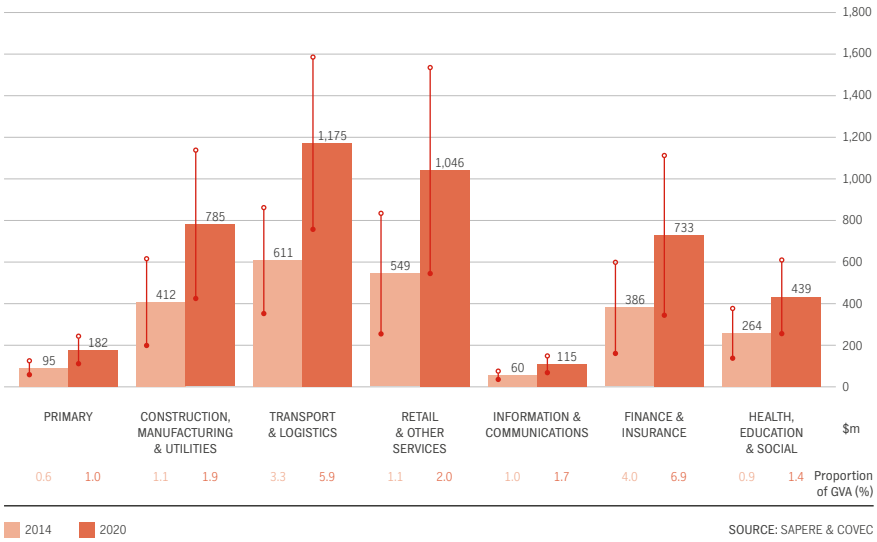
- Cost savings, such as from operational efficiencies and better investment decisions – We assume that these savings are partially retained by organisations and partially passed on to consumers in the form of lower prices.
- Increases in revenues, such as from more efficient marketing and the development of new products and services – We assume that these gains are all retained by firms, and we do not attempt to estimate the value to consumers of new products and services arising from DDI.

We estimate substantial consumer benefits from DDI as competition forces businesses to pass on some of their cost savings. Overall we estimate that around 56% of the estimated benefits (around $1.3 billion in 2014 and $2.5 billion in 2020) accrue to New Zealand consumers in the form of lower prices. Across sectors, our estimates of the economic value of DDI vary according to the overall size of the sector, as well as estimated uptake of DDI and its impacts on business revenues and costs in each sector.[1]

We estimate that around 56% of the estimated benefits (around $1.3 billion in 2014 and $2.5 billion in 2020) accrue to New Zealand consumers in the form of lower prices.

FIGURE 5

ESTIMATED IMPACT OF DDI ON GVA BY SECTOR IN NEW ZEALAND



SOURCE: SAPERE & COVEC

It is important to note that we only estimate impacts of DDI on economic activity. DDI may create significant spill-over benefits that are not captured in traditional measures of economic activity, such as reductions in environmental emissions and improvements in people's health and well-being. We have evidence that these effects exist, but cannot quantify them with precision. It is also likely that we have excluded some uses of data from our analysis. This, together with the fact that we ignore consumer benefits from new products and services, means that our methodology probably generates an under-estimate of the overall value of DDI in New Zealand. We include a Technical Appendix that has additional information about how we calculated our results, and the other limitations of our analysis.

[1] We have estimated the GVA impact of DDI at the point of use of DDI, rather than the sector by which DDI services are provided. This partly explains the relatively small impact of DDI in the ICT sector.

International comparisons

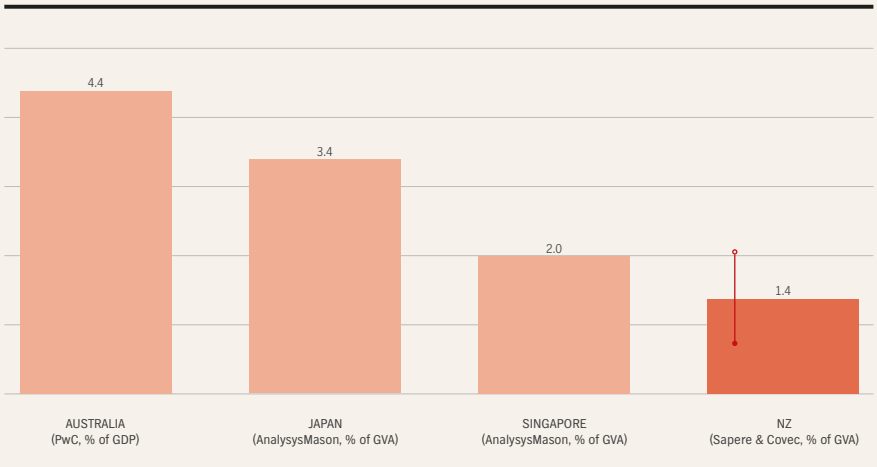
Our estimates of the impact of DDI in New Zealand are significantly lower than comparable international estimates (Figure 6). Our lower estimates partly reflect the nature of the New Zealand economy: we have a comparatively small manufacturing sector where studies in other countries have identified extensive opportunities to use data to improve processes and develop new products.

However the main reason for our relatively low estimates is an overall low rate of uptake of DDI by New Zealand businesses. For example, we estimate overall DDI adoption rates in New Zealand in 2014 to be around half those estimated by AnalysysMason for Japan. If New Zealand businesses were to adopt DDI at the same overall average rate estimated by PwC (2014) for Australian businesses, we estimate that the value of DDI in New Zealand would double to around $4.8 billion.

A number of businesses we interviewed indicated that adoption of DDI in New Zealand is around five years behind that in other leading countries. A common anecdote was that while there are many people in New Zealand organisations that understand and use data, relatively few organisations have embraced data as a basis for making decisions at the upper-management and board levels. Instead, many New Zealand decision-makers prefer to rely on gut feel and experience when making important decisions.

FIGURE 6

INTERNATIONAL COMPARISON OF DDI IMPACTS



SOURCE: ANALYSYSMASON (2014A, B), PWC (2014), SAPERE & COVEC

Note that the Australian study estimates a GDP impact while the other studies focus on GVA. The primary difference between GDP and GVA is that the latter does not take account of taxes and subsidies applied to goods and services.

Social value of DDI

The effects of DDI extend beyond economic activity. For example, there are significant opportunities for government departments to use data to improve the wellbeing of New Zealanders through improvements in public services, and a better understanding of the real extent of problems that government agencies can best solve. To illustrate the potential: in New Zealand there are approximately 180,000 acute hospital admissions per year of adults aged over 65. If, for example, 10% of these admissions could be proactively prevented by identifying individuals at risk of hospital admission and making changes to their regular healthcare in order to prevent the problem from becoming serious enough to require admission, total costs of around $97 million per annum could be saved. This would free up healthcare resources for other purposes, such as increased elective surgery, generating further benefits for other patients. In addition, some people would not have to endure being admitted to hospital and this is a benefit to them quite apart from the cost savings.

The OECD (2014, page 21) refers to work in Finland to link patient data on the content, quality and cost-effectiveness of treatment of a set of selected diseases "across the whole cycle of care from admission to hospital, to care by their community doctor, to the medications prescribed and deaths. The results of the analysis are made publicly available and have empowered patients and led to improvement in the quality of hospitals in Finland."

Policy and DDI

DDI can be enabled or constrained by policy settings. Policies such as a drive towards open government data, government leadership on the smart use of data, or specific support for data-related training and education can actively promote DDI.

Policies that restrict data collection, sharing, use and re-use can constrain DDI. These policies can serve important purposes: in particular, constraints on collecting and sharing data are central to the protection of privacy.

But it is not a simple trade off. Clear rules can also actually help DDI, by reassuring data subjects that they will not suffer harms as a result of providing data to government or firms. We argue that a good understanding of the specific harms we are trying to prevent through policy is an important starting point for data-related regulation, and allows clearer ground to be found between the desire to protect individuals from harms and the enabling of DDI.

The challenge for policy-makers is to understand and respond to the data-related harms that society wants protection from, and provide a balanced regulatory framework that enables the maximum social and economic benefit from DDI.

Framework for data regulation

Regulatory frameworks affect DDI adoption in a variety of ways. Analysys Mason (2014) express this as 'enabling' or 'constraining' policies in Figure 7 below

FIGURE 7

CHARACTERISTICS OF DATA REGULATORY FRAMEWORKS

ENABLING POLICIES	CONSTRAINING POLICIES
Positive Policies: • Direct funding • Ecosystem development • Open data • Skills development • Technical standards	*Activity-specific:* • Collection • Storage • Combination/repurposing • Analysis • Use
Removal of barriers: • Regulatory clarity • Trust-enabling data protection • International regulatory interoperability	*Non activity-specific:* • Inter-company transfers • International transfers • Security

SOURCE : ANALYSYSMASON (2014A)

They also make the important point that the description of 'enabling' or 'constraining' policies:

> *is by no means intended to suggest that only 'enabling' policies are desirable while 'constraining' ones are always undesirable. Indeed, the importance for citizens of issues such as privacy protection means that some constraining policies may always be necessary. However... policymakers can aim to address the intended outcomes of data protection through alternative policies that minimise potentially negative effects on innovation.*

In our context, it appears that New Zealand regulation of data collection and re-use may constrain innovation as personal information cannot be used for new purposes other than those disclosed at the time the data was collected.

One insurance company told us that in the aftermath of the Christchurch earthquakes there was less information sharing than there could have been because organisations were concerned about privacy risks and were highly risk averse. The net effect was that the same reports got repeated but for different clients. Our interviewee thought that government action was required to resolve this because of the privacy issues involved.

Data regulation does serve a broader social goal but the fact it can act as constraint - including on innovation - means that greater effort should be made to design regulatory alternatives to achieve both goals.

Continue to push open data

New Zealand is one of the world's leaders in open government data. Government agencies continue to make substantial efforts to make their datasets more easily available and more useful, and the latest reporting shows that just over a third of core government departments are consistently releasing public data fully compliant with the standards the government has set. For example, land registry information published by Land Information New Zealand has been widely used by public and private sector organisations involved in urban planning and property development services.

There is still a considerable gap between the government's aspirations for open data and the reality. Many government datasets available online are not in a useable format, not usefully processed, or not published regularly. The latest government report on progress on its open data efforts (New Zealand government, 2014) says that in 2014, nearly 80% of government agencies released datasets, but only 45% of agencies released datasets in open formats.

Plus, making available data for public release is only the first part of the story. Even if data is open, it requires special skills to find and manipulate the data. To make use of data, a person has to know the data exists, know which agency collects it, navigate to a website and either download a spreadsheet or use a web-tool that is different for every agency to obtain the numbers. From there, the data will need to be loaded into a spreadsheet or some other software and manipulated to get the specific figures of interest. While such data may be "open" in the sense that it has been published in an open format on a website, this does not mean the data is accessible to the majority of potential users. This remains unfinished business for all open data programmes: open data does not yet mean useable data.

Consider a Chief Data Advisor

A government-appointed data specialist could help to improve data policy and the use of data within government. This could be a "Chief Data Advisor", in a similar role to the existing Chief Science Advisor. A Chief Data Advisor could provide independent advice on data-related policy proposals and consider the impact on DDI. This could help to balance the policy-making system on matters relevant to the regulation of data collection and use. The Chief Data Advisor could also advise government agencies on making better use of data within their own operations and help to increase the quantity and quality of open government data.

Help the development of the DDI ecosystem

The efforts of the Data Futures Forum in 2014 were a helpful and innovative way to progress issues very relevant to DDI. The government's response is very supportive of the Forum's proposals.

There are a range of other broader ways that government effort could directly promote greater use of DDI. These include working to improve the availability of people with data analysis skills, by making sure that the education system is set up to create graduates

One insurance company told us that in the aftermath of the Christchurch earthquakes there was less information sharing than there could have been because organisations were concerned about privacy risks and were highly risk averse.

who can participate in DDI, and by checking that the immigration system is able to support bringing in people from other places: these skills are in demand worldwide.

Policy settings that promote entrepreneurship generally, such as seed or grant funding for high-tech companies, incubators and accelerators to help firms grow, and efforts to assist internationalisation and access to the capital markets will all support new data analytics businesses. The OECD (2014) refers to a range of other relevant issues that impact on DDI, including investments in broadband, access to cloud computing and analytics platforms, and resolving tricky issues about what data "ownership" means.

All that said, effective use of data to transform businesses and the operations of government is the main long-term challenge. To truly gain the potential benefits of DDI, businesses and government entities will need to change how they operate, from decision-making to organisational structures, in light of the huge increases in the volume and usefulness of data for decision-making in recent years, and the ease with which this data can now be stored, analysed and shared. The extent to which businesses and government agencies take advantage of these trends will determine whether New Zealand makes the most of data driven innovation.

Big dairy meets big data

Fonterra has been milking international dairy information to forecast its prices – but this data now allows it to do even more. In 2009 the Global Dairy Intelligence team created the Marketing Analytic System (MAS), which at first focused on global supply and demand but has since expanded to include everything from on-farm economics and trade data to global production and consumption. The complex system sources data from more than 1200 points worldwide and its use has far reaching benefits – and the potential to benefit the entire New Zealand economy.

2. The value of data-driven innovation in New Zealand

This chapter presents our estimates of the value of DDI in New Zealand, and explains how they vary across sectors and how they compare with estimates from other countries. We begin with an explanation of how DDI creates economic value.

Defining DDI

In the past decade there has been a step-change in the volume and usefulness of data, driven by improvements in the economics and performance of computer hardware, software, and networks used for collecting, storing, analysing and sharing data.

The step change in the volume of data has largely been associated with machine-generated data. This is data that machines create automatically, either:

- As part of their core function, e.g. GPS units, weather sensors, retail loyalty cards, and radio frequency ID (RFID) chips exist mainly to create data; or
- As a by-product of their core function, e.g. cellsites exist to provide mobile telephony services but also generate meta-data that can be analysed, for example to understand populations and movements within cities to aid transport planning.

The step change in usefulness of data has been driven by reductions in the cost of data storage and analysis, and improvements in the ease with which data can be analysed.

Data is a broad concept, and potentially DDI could include almost all innovations, because the majority are likely to involve some sort of analytical process rather than arise just by chance. However, in order to reach more meaningful conclusions about the value and implications of DDI in New Zealand, we have restricted our analysis to the effects of this step-change in the volume and usefulness of data. The OECD (2014 page 11) defines DDI as "the use of data and analytics to innovate for growth and well-being".

In essence, we have studied the value to New Zealand of moving to a state where innovation and decision-making within firms and government agencies is based much more on data analysis than on hunches or theories.

The DDI value cycle

Value created from DDI can be broken down into a number of steps, which we think of as a cycle (see Figure 3 on page 9).

The process of generating value from DDI starts with the **collection** of data, which may be either directly captured machine-generated data or gathered as a result of observation of the real world. An important part of the data collection stage is the specification of data definitions: what exactly will be measured, in what units, and how frequently will measurements be taken? Data definitions will shape and constrain the analysis that is subsequently possible. Thus the value that can be generated from analysis of a dataset depends on the quality and clarity of data definitions, and the consistency with which these definitions are maintained over time.

Once data has been collected it must usually be **stored**. Digital data storage involves defining formats and establishing processes for data backup, security, and privacy. The quality of analysis of a dataset depends on maintaining the integrity of data over time and making it readily accessible in suitable formats.

Substantial potential value can be added to data through **processing** to tidy it up and perhaps transform it for subsequent analysis where that potential value is realised. This might involve imposing standard data definitions, fixing errors, interpolating missing values, or manipulating the data to get it into a form that is more easily analysed.

Together, the collection and storage of data enables innovation through subsequent analysis and application of data to end uses. Within the **analysis and presentation** stage, raw data is turned into useful information through the application of statistical analysis and models of varying degrees of sophistication, and data visualisation. Analysis and presentation may involve the combination of datasets and/or re-purposing datasets to new uses.

The final stage of the DDI value cycle involves application of the analysis to **end uses**. In many cases this is an input to decision-making, e.g. designing products and services and setting prices, planning marketing campaigns, making operational decisions, and making investment decisions. In such cases the value created is the ability to make "better" decisions.

How DDI creates economic value

The value of data in decision-making arises from the fact that most decisions are made in the face of incomplete information or some uncertainty about future outcomes. Data and analysis reduces this uncertainty to some extent, allowing people to make choices that have better outcomes (on average) than they otherwise could.

DDI generates economic value through a number of channels; these channels underpin our economic analysis. For example, improved sales and marketing will lead to new products and services, increasing revenue for firms and choices for consumers. Cost savings as a result of DDI will increase business profits and allow resources to be re-allocated to other productive uses. Competition will also lead to some of the cost reductions being passed through to consumers, increasing consumer welfare.

In the longer term, data and analysis will improve the quality of planning and investment decision-making. There will be improvements in the timing and location of investments. Better demand forecasting will allow firms to have more confidence in investments in specific locations or for new products. This will allow scarce capital resources to be allocated to better uses, and will increase return on investment. To the extent that additional investment leads to new products and/or stimulates competition, consumers will also benefit through an increase in consumer surplus.

There is some evidence in the economic literature of positive economic impacts from firms making decisions using data. For example:

- LaValle (2011) reports the results of a global survey of 3,000 businesspeople that found higher-performing organisations were more likely to apply analytics when making decisions in comparison with lower-performing organisations. The authors noted that top performers "put analytics to use in the widest possible range of decisions, large and small".

- Brynholffson et al (2011) found that the use and application of "data-driven decision-making" makes firms 5-6% more productive based on survey data of the business practices and information technology investments of 179 large publicly traded firms.
- Barua et al (2011) examine the impacts of effective data use for businesses based on a sample of 150 Fortune 1,000 companies. The headline result is that relatively small improvements in firms' effectiveness in the use of data led to significant improvements in financial returns.

Aggregate impacts

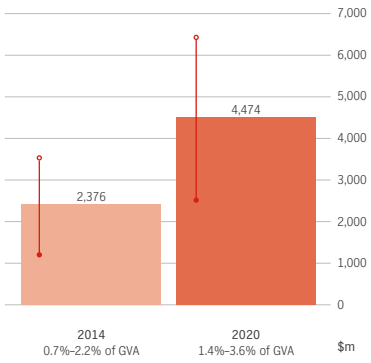
We estimate that DDI contributed $2.4 billion to GVA in New Zealand in 2014, with an expected range from $1.1 billion to $3.6 billion (Figure 8). This represents around 1.4% of total economic activity across the seven sectors in our analysis. By 2020, with higher adoption, we project this to grow to $4.5 billion, with an expected range from $2.5 billion to $6.5 billion in real terms, or to between around 2.5% of total economic activity in 2020 across the seven sectors.

The projection for 2020 should be taken as indicative of short to medium term effects, rather than a precise projection for that year. We expect DDI uptake to increase in New Zealand in the short to medium term, but eventually diminishing returns will set in and the increase in the value of DDI that we are projecting between 2014 and 2020 will not continue indefinitely. In addition, given available information, it is difficult to be more precise about the path of uptake of DDI between 2014 and 2020, or beyond 2020.

Note that we only estimate impacts of DDI on economic activity, and our methodology excludes some benefits and uses of DDI, so we believe we have under-estimated the overall value. The Appendix has more details of our methodology and its limitations.

FIGURE 8

ESTIMATED TOTAL IMPACTS OF DDI ON GVA IN NEW ZEALAND



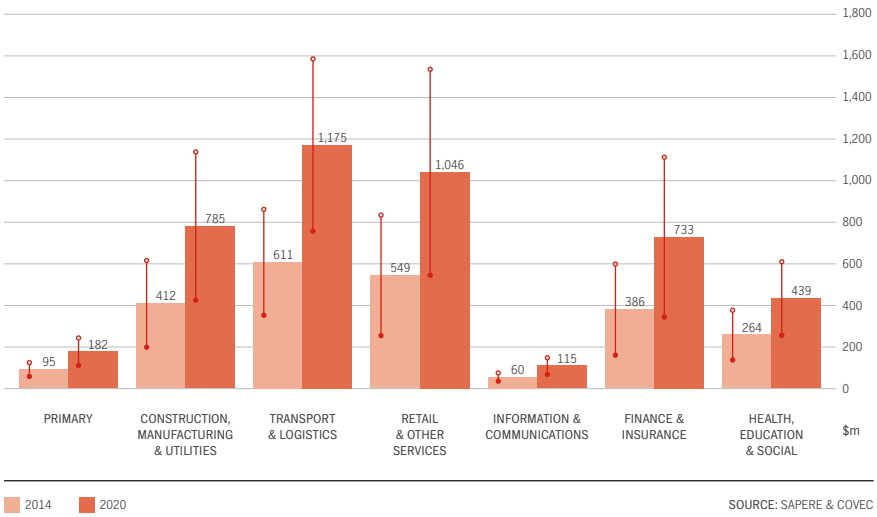
SOURCE: SAPERE & COVEC

Impacts by sector

Figure 9 breaks down our estimated impacts of DDI across the seven sectors in our analysis. The results vary across sector depending on the size of each sector and the assumed uptake and revenue and cost impacts of DDI in each.

FIGURE 9

ESTIMATED IMPACT OF DDI ON GVA BY SECTOR IN NEW ZEALAND



SOURCE: SAPERE & COVEC

The greatest estimated impacts in dollar terms are in the transport and logistics sector, and in retail and other services. Relatively large impacts are also estimated in the construction, manufacturing and utilities sector and the finance and insurance sector. This reflects the richness of datasets in these sectors, the potential impacts that data can have on business performance, and the propensity of businesses in these sectors to use data and analytics to make decisions.

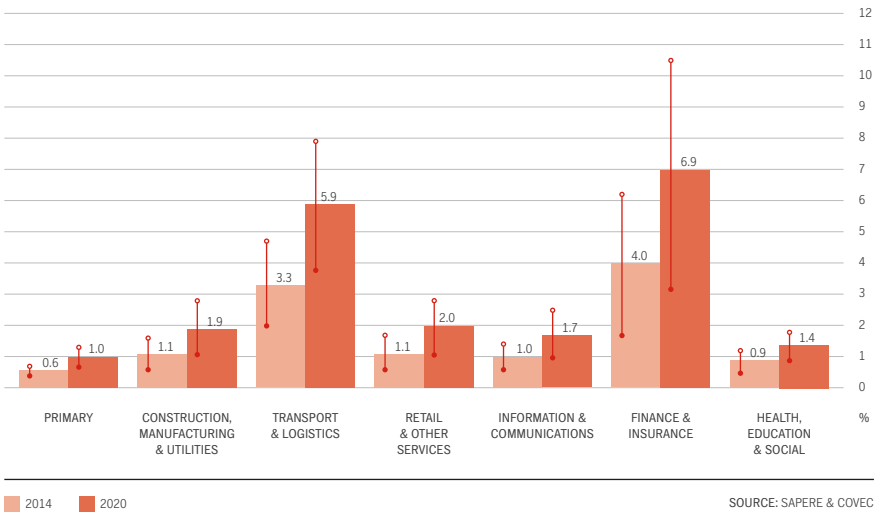
The relatively small impacts in the primary sector reflect the size of this sector as formally measured[2], and because we think there is limited uptake of DDI within this sector based on our interviews with New Zealand primary sector organisations. However there are significant potential benefits of DDI in the primary sector – for example people we spoke to in the dairy sector suggested that DDI could generate efficiency improvements on the order of 10%, from a combination of cost reductions and improvements in yield. The low results in the information and communications sector mainly reflect the relatively small size of this sector as a proportion of total economic activity in New Zealand, and the fact that we attribute the value of DDI to the sector in which it is undertaken, rather than the sector in which DDI-related services are supplied.

[2] Note that a large proportion of economic activity related to primary production is reflected not in the primary sector statistics (which generally cover on-farm efforts) but in the manufacturing statistics, and particularly the sub-sector "Food Product Manufacturing", which incorporates most of Fonterra's operations, for example.

Figure 10 expresses the sector results in relative terms as a proportion of total GVA in each sector. The greatest impacts are estimated in the finance and insurance, and transport and logistics sectors. Results across other sectors range from around 1 - 2% of sector GVA, with the exception of the primary (agricultural and natural resources) sector, which again reflects the relatively low estimated uptake of DDI in that sector.

FIGURE 10

DDI IMPACT RELATIVE TO TOTAL GVA BY SECTOR IN NEW ZEALAND



SOURCE: SAPERE & COVEC

Our New Zealand-specific results are rather different from the work of the OECD (2014, page 5). It says "the ICT sector is still the largest uses of advanced analytics according to some estimates", and:

> "[t]he most data-intensive sectors outside the ICT sector are the financial sector, and professional and business services sector... However, public administration as well as education and health services are the sectors were [sic] the adoption of data analytics could have the highest impact in the relatively short run. These sectors employ the largest share of occupations which perform many tasks related to the collection and analysts of information with, however, a relative low level of computerisation".

Consumer and producer impacts

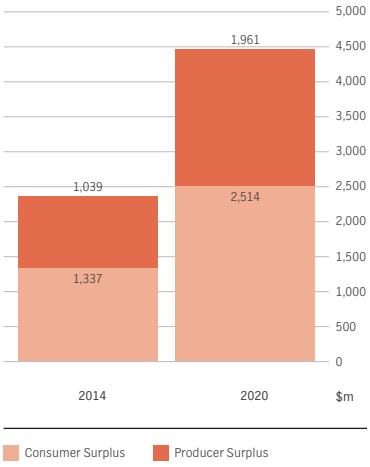
We have also disaggregated the impacts of DDI into gains for consumers and gains for producers. We assume that some proportion of cost reductions created by DDI are passed through to consumers in the form of lower prices, with the degree of pass-through varying by sector depending on the intensity of competition.

Figure 11 summarises the mid-point estimate of total consumer and producer benefits of DDI. In 2014 we estimate that DDI generated $1.3 billion of benefits for consumers (with an expected range from $0.7 billion to $2.0 billion) and $1.0 billion of benefits for producers (expected range from $0.4 billion to $1.6 billion). By 2020 we project these impacts to increase to $2.5 billion for consumers (range between $1.5 billion and $3.5 billion) and $2.0 billion for producers (range between $1.0 billion and $3.0 billion).

Overall we estimate that around 56% of the benefits of DDI accrue to consumers in the form of lower prices and the remaining 44% are retained by firms in the form of higher profits, although this varies by industry depending on competitive intensity (Figure 12).

FIGURE 11

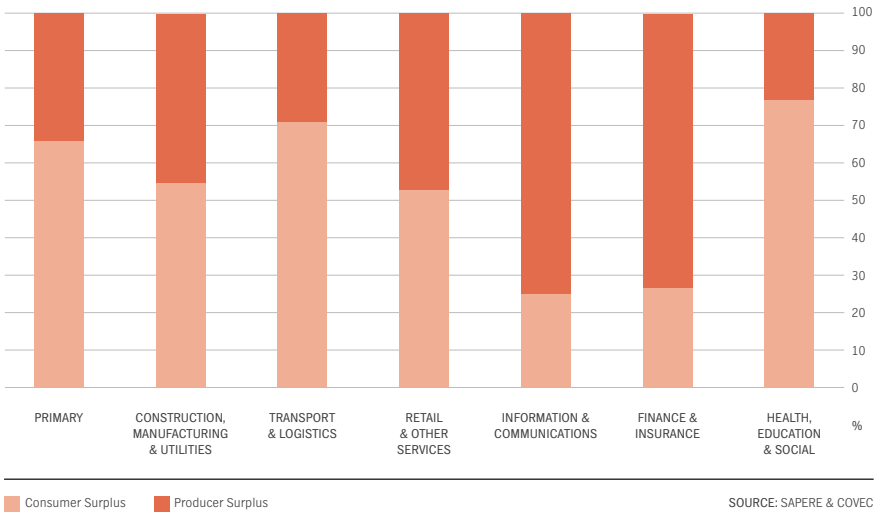
ESTIMATED EFFECT OF DDI ON CONSUMER AND PRODUCER SURPLUS IN NEW ZEALAND (MID CASE)



SOURCE: SAPERE & COVEC

FIGURE 12

SPLIT OF IMPACTS OF DDI IN NEW ZEALAND BETWEEN CONSUMERS AND PRODUCERS BY SECTOR



SOURCE: SAPERE & COVEC

International comparisons

If New Zealand businesses were to adopt DDI at the same overall average rate estimated by PwC (2014) for Australian businesses, we estimate that the value of DDI in New Zealand would double to around $4.8 billion.

Our estimates of the impact of DDI in New Zealand are significantly lower than comparable international estimates (Figure 13). Our lower estimates partly reflect the nature of the New Zealand economy: we have a comparatively small manufacturing sector where there are extensive opportunities to use data to improve processes and develop new products. However the main reason for our relatively low estimates is an overall low rate of uptake of DDI by New Zealand businesses. For example, we estimate overall DDI adoption rates in New Zealand in 2014 to be around half those estimated by AnalysysMason for Japan. If New Zealand businesses were to adopt DDI at the same overall average rate estimated by PwC (2014) for Australian businesses, we estimate that the value of DDI in New Zealand would double to around $4.8 billion.

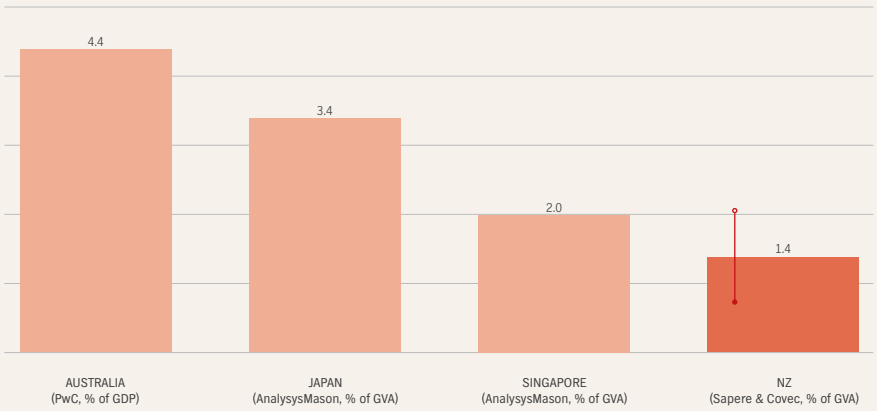
A number of the businesses that we interviewed as part of our research indicated that adoption of DDI in New Zealand is around five years behind that in other leading countries, such as the United States. A common anecdote was that while there are many people in New Zealand organisations that understand and use data, relatively few organisations have embraced data as a basis for making decisions at the upper-management and board levels. Instead, many New Zealand decision-makers prefer to reply on gut feel and experience when making important decisions.

We estimate that DDI has been adopted by only 10 - 15% of New Zealand organisations in 2014. Over time we expect adoption to increase as local and international competition drives organisations to adopt new technologies and improve their internal processes to make more use of data in decision-making. We have assumed that the overall average adoption rate by 2020 has increased to 20 - 25% of New Zealand organisations.

This helps explain why the current value of DDI in New Zealand is lower than other countries. It also means there is considerable potential to generate value from greater business uptake of DDI in the future.

FIGURE 13

INTERNATIONAL COMPARISON OF DDI IMPACTS



SOURCE: ANALYSYSMASON (2014A, B), PWC (2014), SAPERE & COVEC

Note that the Australian study estimates a GDP impact while the other studies focus on GVA. The primary difference between GDP and GVA is that the latter does not take account of taxes and subsidies applied to goods and services.

Use of data in the public sector

Our analysis does not allow us to estimate a value for DDI in the public sector directly: it is combined in our analysis with private organisations. But clearly there is significant scope for using data in the public sector to improve social outcomes. All governments face the challenge of allocating a limited amount of funding and resources to where these are needed most. Analysis of data can help to identify the areas of greatest need and re-allocation of expenditure from interventions that are less effective to those that are more effective and will increase citizens' overall well-being.

There are a growing number of examples of the use of analytics by government agencies in New Zealand. Some of the newer examples are driven by the collation and availability of larger datasets that give information about individual people, businesses, or relatively small geographic locations. One example is the Integrated Data Infrastructure (IDI) administered by Statistics New Zealand and containing data on individuals from a number of government departments.[3] Microdata about businesses is available in Statistics New Zealand's Longitudinal Business Database (LBD).[4]

Both the IDI and LBD allow researchers and policy analysts to study the effects of policy on individual people or businesses, and to undertake cross-sectional analysis on the determinants of individual health and well-being, and business performance. This provides a much richer source of information than aggregated statistics: the analyst move beyond simple statistics such as averages to the *distribution* of variables like health status or business profitability within the population. This allows better understanding of outcomes for individual people or businesses.

[3] See http://www.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx.

[4] See http://www.stats.govt.nz/browse_for_stats/businesses/business_characteristics/longitudinal-business-database.aspx.

Smarter means healthier

You can't put a price on health. But you can improve the way it's delivered. By using data to study the link between admissions and certain types of conditions, Christchurch Hospital has managed to eliminate patient 'gridlock' during its busy winter months – and its subsequent effects on care, waiting times and staff frustration. Following this success, the Canterbury District Health Board has used data to reduce patient acute medical admission rates to 30% below the national average, while improving health outcomes.

3. Interview themes

In this chapter we present our analysis of the themes from the 48 interviews we conducted part of this project. The interviewees were a mix of private businesses and government agencies in New Zealand; some interviewees were data and analysis service providers, but most were consumers of data and analysis services provided internally or by others.

This was not a formal survey so we did not undertake any statistical analysis, and the responses should be compared with relevant data to check how far any particular response can be extrapolated. But the interviews were structured, enabling us to identify common threads.

The questions were about how data influences organisational performance, how sophisticated use of data is, and what government policy or organisational issues influence these issues. Interviewees also nominated a wide range of uses of DDI. We present examples of these uses throughout this report to aid understanding and give some qualitative depth.

The value of DDI

All our interviewees thought that DDI generates significant value. After asking specifically about the value for their organisation, we explained to interviewees that international studies showed a 2-5% profitability gain from DDI and asked what they thought of that range. Overall, the responses suggest that this is an understatement of the value that is possible, although actual value depends on take-up, which we discuss further below.

In agriculture, even our most pessimistic interviewee said that a 2-5% efficiency gain was plausible from a mixture of cost reductions and production increases, noting that the difference between lowest quartile and highest quartile farm productivity could be as much as 20-30%. The dairy sector interviewees nominated higher impacts, closer to 10%. For example, one farmer told us that just more efficient decisions around irrigation plus the avoided costs of expensive manual monitoring could deliver a 5-10% net improvement in returns. One major logistics operator told us that its internal estimates of the payoff in avoided costs from trucks being out of action, goods spoiling, and inefficient use of assets and time delays was in the order of 10-15%. Others were more bullish, with one agriculture sector interviewee saying that "data collection and use will be the major, and possibly only, driver of growth for agriculture in the future"

The data and analysis businesses that we interviewed also nominated larger impacts. A best case result was said to be a 30% reduction in advertising spend or a 30% increase in revenue from the same spend from using data to target marketing effectively. Another service provider thought marketing analytics could easily improve returns by 5% initially and as much as 20% over time. A third said that members of its loyalty programme were seeing a 5-10% revenue uplift from the data-driven communication enabled by their platform. At the most extreme amongst our interviewees, one retail service provider we spoke to said that "the value of data is the difference between business success and failure", suggesting the 2-5% impact is far too low in some cases.

It was typical that interviewees did not really know how to quantify the benefits of DDI, and they generally had not quantified the benefits themselves. Several said that business cases for investment were justified on the basis that the benefits were so large relative to costs that the project only had to be slightly successful for it to be worthwhile. One insurance company told us that it had done a cost-benefit analysis of one of its major DDI efforts intended to reduce insurance fraud and the direct benefits were calculated at four times the costs.

That said, it is hard to attribute all of the value to data analysis, and a lot of the value of DDI comes about through happy accident or as a by-product of another effort. This echoes a point made by other interviewees; making smarter use of data is a gradual process rather than a one-off investment, and so there was no single specific investment case to consider.

Interviewees also nominated a range of other benefits from DDI apart from its financial payoffs. We have included a case study of FoundationFootprint's services that help businesses reduce their carbon usage through analysis of their energy consumption data. More generally, building the capability to do data analysis means that firms benefit from lower cost innovation and experimentation, and they can build better products over time. Several interviewees talked about the savings in operating costs from investing in data analytics systems and processes, meaning that projects that could never be justified on their own could be pursued in a much more efficient way.

Data collection and use will be the major, and possibly only, driver of growth for agriculture in the future.

Take-up of DDI

Although all interviewees agreed that there was a lot of value available from DDI, they also universally said that take-up of DDI at present amongst New Zealand businesses and government agencies was relatively low. "It is amazing how poorly the agriculture sector uses data, especially monitoring" said one interviewee with experience in financial services. His view was that 1-3% might be a more accurate figure for the current impact of DDI, given the rudimentary understanding and utilisation of data compared with what is possible.

We were told that New Zealand is 2-3 years behind the USA or Australia in terms of the use of DDI in the retail sector. The main gap is just adoption by organisations. There were said to be some good technical people within organisations but the organisations themselves just do not understand data and what it can do for their performance. As one interviewee put it, "there are pockets of capability within organisations but organisations as a whole have not embraced data driven decision-making". We were told, for example, that take-up of data-driven marketing communications in retail might be around 30% of large organisations, and is even lower for smaller operators.

Other interviewees saw larger gaps. One large retailer suggested that New Zealand and Australia are as much as seven or eight years behind the USA. One retail service provider said it "could not name one organisation in New Zealand that is really data-driven", and that often the use of data is to justify decisions rather than to make them. Another interviewee noted that firms have yet to embrace data analysis in their regular decision-making: there are many people that understand data but only a few are data-oriented from the very top. Another thought that his sector was "at least ten years away from real change".

Policy

There were relatively few specific issues with existing policy raised by interviewees as barriers or enablers to use of DDI.

Framework for data use

Three interviewees referred to the lack of a structured ethics framework around approving new uses of data or new data projects. Two drew a comparison with health research, which has a well-established process for formal approval of research. One suggested an open public catalogue of data that is collected and the uses to which it is put to create more pressure and market discipline through transparency. This was thought increasingly important as the data aggregation market is growing strongly and there is a clear trend for organisations to try to monetise their data in the absence of governance and rules.

Several interviewees mentioned the commercial benefits of good privacy protections, and pointed out that it was in the interests of the data analysis industry to promote best practice and to control rogue operators in order to reduce the risks created by privacy breaches to the industry as a whole. Generally privacy regulation was seen as a business as usual issue that needed to be managed, not a major inhibitor to business operations. Several interviewees pointed out that firms collecting and using data operate by the consent of those providing personal information. If privacy controls are too lax or data is being misused then people may refuse to provide information or provide false information that makes analysis harder.

Some interviewees said that they saw growing demand by consumers for privacy protection; one noted that it may be harder in this environment for small firms to meet consumer demands. One major retailer told us their privacy policy is tighter than that required by legislation and restricts third party data sharing. They saw this as a differentiator in the market; protecting the value of their brand and retaining the trust of their users. Other interviewees pointed out that sensitive information abounds, even in places that are unexpected. For examples from transport: vehicle crash data, tolling system data, and data from the integrated public transport ticketing system in Auckland all contain potentially sensitive information.

None of the few interviewees who addressed this question thought that market disciplines alone would be sufficient for privacy protections. In particular there was a perceived risk that some users of data would not respect the privacy of data subjects and that this would create negative consequences for all users of data, including those that did treat personal information appropriately. There is a need for a regulator to support market behaviours but also to ensure there are minimum standards of practice.

Data localisation

Most of those we interviewed did not see data localisation (or 'data sovereignty') as a major issue, but it was common to hear from data analysis firms that customers required their data to be kept onshore, leading to higher costs. One estimated that local cloud storage is 10 to 20 times more expensive than storage from international providers and that these extra costs lead to organisations not storing and using data as much as they can.

Some interviewees hypothesised that these customer requirements were based on an erroneous view that overseas firms and laws will not protect New Zealand data as well as domestic institutions will. One interviewee suggested that the lack of coordinated international legislation to ensure protection of offshore data meant there was a greater emphasis on trust between commercial operators in different countries. Another interviewee emphasised the importance of government negotiations to work out safeguards and ensure using offshore cloud providers are following equivalent rules to those domestically.

Concerns about future uses of data

Many interviewees noted that the capability to analyse data continues to rapidly advance. In this environment, it is hard to be sure what could be done with any particular dataset in the future. Our interviewees were split between those emphasising the benefits of this uncertainty, and those concerned about the risks.

Some were especially concerned about the use of data collected for one purpose for another purpose, e.g. student educational data designed for educational purposes used to performance-measure teachers or unfairly rank schools. One interviewee urged caution particularly about predictive analytics for public services, saying that it needs to be better than other available methods of identifying at risk groups to justify the risks of its use, i.e. people missing out on assistance that would have been helpful to them. This interviewee emphasised the need for a moderated, practitioner-based approach rather than something fully computerised. This interviewee also specifically mentioned re-identification as a big issue, particularly since very little data is necessary to identify someone.

Other interviewees were concerned about future uses of data for the opposite reason, i.e. the risk of failing to take advantage of the opportunities of data because of actual privacy restrictions or because uncertainty about the rules leads to firms not sharing data even when they are able to do so.

One interviewee in the health sector said that the greatest issue was the lack of ability to use all the data that is being collected for all the things it could be useful for, and suggested building datasets that would enable researchers to work with real data on real health system problems. This interviewee pointed to specific challenges in health where some useful data is collected only at the national level and it cannot be shared

Many interviewees mentioned the need for organisations to sharply improve their ability to analyse data and lift their maturity in the use of data in order to make the most of the more data-rich business environment.

regionally with personal information included, but that personal information (in the form of the NHI, a unique patient identifier) is required to link datasets together to make them more useful.

One interviewee talked about privacy impact assessments as a tool to try to assess some of these costs and benefits. Using an example of data matching between two sets of personal information, for example, an assessment could look at whether the potential for savings in operational processes from a combined dataset are outweighed by the risks and potential harms of false matches. More generally the risks from errors can be mitigated by good processes, and any privacy assessment should itself be flexible enough to adapt to the question at hand. We were told that organisations can do these types of assessments before any new use of a database, a technology innovation or a new business process. If privacy issues are found, an organisation will need to figure out how to mitigate them or how to work in another way to get the same results without risks to privacy obligations.

Foundations of DDI

Many of our interviewees shared their thoughts on what was driving the growth in DDI and what was necessary for New Zealand to make the most of it.

Drivers

Several interviewees noted the recent change from scarcity to abundance in the volume of data that is collected and is able to be practically analysed. A typical respondent was one who noted that the requirements for data have not changed much over time and that his organisation still relies on some long-standing datasets, but that there is more and more data available, and more use is being made of data that has previously been collected.

Many interviewees mentioned the need for organisations to sharply improve their ability to analyse data and lift their maturity in the use of data in order to make the most of the more data-rich business environment. One said that its main problem is in dealing with the volume of data created by the users of its online service. Data storage is not the problem. The problem is figuring out how best to use all of that data, and ensuring there are analytic tools available and that the organisation has the right ones in place to make sense of the data it has.

We were told that the change in data availability amounts to a paradigm shift that should alter how firms think about what they do, and that the current methods of operating are highly inefficient: firms are not used to working at speed of data, or with the precision and clarity that data can provide, and analysis and interaction methods within firms are still very primitive, i.e. data is collected and analysed in spreadsheets and shared by email. “Data is getting ahead of policy”, said one government interviewee. It means that a new approach to government policy-making is possible. Rather than making a best guess at a single policy solution and then coming back and reviewing

some years later to see how it is working, it is possible to "beta-test" policy, i.e. to make real-time changes in policy settings based on a flow of data on how the policy is working. This is pushing government agencies to innovate and continuously improve rather than just set and forget.

Some interviewees mentioned increasing demand for data-driven decisions, especially in government policy. One agricultural sector interviewee said that changes in organisational structure on farms have changed demand for data: a move from a single herd/single farm model to multi-herd across multiple sites generates the need for data for management. In previous work we have been told both that corporate farms require more data because the shareholders are not on the farm, and that only computerisation can allow farmers to make good management decisions when their herd grows beyond a few hundred animals.

Skills and culture

Many interviewees said that a lack of people with the right skills was the real constraint on DDI. "Data is not that useful unless you have people that understand it", said one. Another reported a "severe shortage" in people with the capabilities required. One questioned whether there should be greater public sector investment in empirical research, and more emphasis in universities on teaching these skills. Another thought that only large organisations would ever need these types of skills.

Relatedly, many interviewees referred to the need to build an organisational culture of measurement and recording, and a reliance on data in making decisions. One major retailer spoke of the "need to democratise data within organisations and make it useable by anyone". Another said that leadership around data-driven decisions was crucial: if the CEO makes it clear that data is required to support every decision, then the organisation can orient around measured reality rather than around hunches about the way the world actually is.

Our interviewees spanned a spectrum in their views on how important data should be. Some said that data could complement hunches or human leadership and decision-making. "Data does not replace heuristics and intuition but supplements it". One pointed out that "data has its own noise", but it can provide a useful test of other, softer, thinking. Others said that sharing data across the organisation was crucial, that it was very hard to make good choices that were not supported by data. Without data, in this view, decision-making just descended into politics and competing views of the world that could not be rationally resolved. One interviewee volunteered that the public sector was less advanced than the private sector in its use of data because politics can work against data-driven decisions.

Terms for sharing

Many interviewees emphasised the importance of getting the terms of access to data right because it directly affects the uses that can be made of data, and therefore its value. One said that the main driver behind innovation is playing, and when data is easy to access and available, people play with it and surprising things emerge. We were also told that the rules for access to data can influence data quality: if data is being used, and particularly if it is being paid for, incentives are stronger to improve data quality.

We spoke to two government agencies that had moved away from charging for data towards making it free in order to let users easily access and use information. The result, unsurprisingly, was much greater demand and from unexpected places; under the charging model it was not straightforward for the agencies to find out that the unmet demand even existed. Similarly, we spoke to a company that provided information to others in the insurance industry. At first it had one of its major datasets managed by a third-party with access on a pay-per-click model. Over time it was found that this pricing discouraged use of the dataset, which undermined the purpose of creating the dataset in the first place. Changing to a membership fee where members can access data freely in individual records or in bulk has sharply altered how the data is used.

Some interviewees encouraged central government agencies to do more on open data, arguing that an agency just making available a spreadsheet on its website does not mean that data is accessible or useable. Government data that is meant to be open is often made available in different places, is stored in different formats, is accessed with different tools, and uses quite different definitions. This creates opportunities for aggregation businesses that combine, clean and document public data and then make that available for a fee. One interviewee argued that there is a significant societal waste from these efforts that could be avoided from better open data practices by government agencies.

One interviewee criticised the lack of clarity in the government's position on data sharing, and the lack of consistent policy across government departments. He argued that a lack of clarity makes departments risk averse, and therefore limits the application of analytics in the government sector. On the other hand, we spoke to one organisation that felt that it had got a competitive advantage from having made the effort to bring together varying data sources relevant to their sector, in this case a dataset of local council data on unsafe buildings that would seem ripe for nationwide standardisation.

Commitment

We did not ask directly about spending on DDI, but many large organisations reported substantial investments in data systems and changes in processes to generate value from data. One organisation, realising that its information on its assets was poor, had 12 people in the field collecting data for two years. One large government agency told us of a six year journey to improve the quality of data collection, to develop new skills to analyse and process it, to revamp how data was shared across the organisation, and to establish processes for how new data-intensive projects were commissioned. Others spoke to us of huge improvements in capability generated by data investments: building a dataset and the tools to analyse it gives endless possibilities to make use of that data that could never be justified if the data collection and storage and analysis was just a one-off project.

Several interviewees spoke of an organisational commitment to data as a gradual journey, with payoffs from early steps justifying decisions to move further ahead. One organisation started using data systematically when it was realised that existing systems did not allow decisions to be made quickly enough or with an accurate understanding of the real value at stake. Another interviewee stressed the importance of learning from data over time, regularly comparing predictions to actuals so that continuous improvement in data modelling is possible.

Several interviewees talked of their efforts to standardise the collection of data in order to make it more useful. One said that it had moved from highly localised ownership of data collection and use to a centralised model because it had learned that coordination issues and information gaps could not be resolved in a decentralised model. Extra structure and discipline in the collection of data and more systematic processes around its use had enabled the organisation to use data to optimise its operations. Others told of the benefits of automated data collection in preference to manual methods. Unless it is practical and straightforward to collect data, there is a barrier to collection and thus there is more likely to be missing data and errors and bias in the numbers.

One very advanced company said that it was now looking to remove spreadsheet-based tools from use because they result in business intelligence being restricted to the relatively few people who can use spreadsheets effectively. New browser-based tools mean that high-quality data analysis can be quickly created and easily shared in formats that anyone can understand.

Control

Several interviewees pointed out that sharing data holds the promise of making life a lot easier and lowering everyone's costs. For example, one dairy sector interviewee told us that many different organisations wanted essentially the same dataset on farm performance, but that they were all required to collect it separately because arrangements had yet to be put in place to enable sharing.

That said, there may be substantial commercial interests against wider data sharing. Particularly in an environment of uncertainty, where no one is sure where the value will be, waiting to see what happens rather than sharing data with organisations that might end up as one's competitors might be the preferred strategy. As one farming sector interviewee said "the tension is about who is going to make a buck out of data, not about policy issues". We were also told that farmers are worried about maintaining control over any data they share: they do not want the data they collect being shared with regulatory bodies in an uncontrolled fashion, for example.

There can also be uncertainty or diversity in arrangements as to who owns the data and who gets to decides what happens to it. Again in the farming sector, we were told there were different views about who owns the data with some saying farmers own their own data and can do with it what they will, and others arguing that data belongs to the collector and that this is not necessarily the farmer, even if the data relates to his/her farm, but might instead be a service provider. We were told a similar situation can arise with public transport data because fragmentation of provision and agreements with different providers may not make it easy to collect system-wide performance data.

One insurance company told us that data sharing arrangements need to be sorted out in advance for natural disasters to avoid delays and reduce duplication. We were told that in the aftermath of the Christchurch earthquakes there was less information sharing than there could have been because organisations were concerned about privacy risks and were highly risk averse. The net effect was that the same reports got repeated but for different clients. Our interviewee thought that government action was required to resolve this because of the privacy issues involved.

Quality

Many interviewees emphasised the importance of data quality. Two said that all the difficult work in data analytics is in tying datasets together and working with various data systems to build a single view of the truth. No organisation, we were told, has a single data system; there are different systems for different purposes and it is not always easy to tie together the data that emerges from them, even when it appears to be about the same thing.

We were told specifically of gaps in quality in particular datasets because aggregated data had lost some of the specific dimensions that made the unit record data useful, because the data itself did not measure exactly what is of most interest, or because data definitions had changed over time and those adjustments had not been properly reflected in the datasets.

One government agency told us that it was increasingly using data collected by others in the course of their business to fulfil its data needs instead of collecting data itself. The benefit is substantially lower costs and freedom to focus more on data analysis; the data for the organisations that collect it as part of their business is essentially a by-product of their operations. But using others' data does not remove the need to be sure about the provenance and quality of information. Organisations need to understand what has been done to collect the data, its characteristics and the purpose for which it was collected.

Smarter spending

Consumers love getting rewards but are tired of carrying a wallet full of loyalty cards. Retailers love having practical data about its customers' spending habits. Both retailers and consumers can benefit from transaction tracking on and with more than 900 million EFTPOS transactions made in New Zealand every year, it makes sense to link loyalty rewards to people's existing debit and credit cards. EFTPlus has created a virtual loyalty programme which gives retailers powerful insights about the spending habits of their customers as well as their competitors', while customers are offered personalised special discounts and other offers.

4. Policy issues

In this chapter we consider how New Zealand's policy settings affect the realisation of value from DDI, look at some of the issues involved in making policy changes, and present some possible ways forward.

DDI can be enabled or constrained by policy settings. Policies such as a drive towards open government data, government leadership on the smart use of data, or specific support for data-related training and education can actively promote DDI.

Policies that restrict data collection, sharing, use and re-use can constrain DDI. These policies can serve important purposes: in particular, constraints on collecting and sharing data are central to the protection of privacy. But it is not a simple trade off. Clear rules can also actually help DDI, by reassuring data subjects that they will not suffer harms as a result of providing data to government or firms. We argue that a good understanding of the specific harms we are trying to prevent through policy is an important starting point for data-related regulation, and allows clearer ground to be found between the desire to protect individuals from harms and enable DDI.

The challenge for policy-makers is to understand and respond to the data-related harms that society wants protection from, and also provide a regulatory framework that enables the maximum social and economic benefit from DDI.

Data-related regulation

Personal information

By and large, in New Zealand the debate about data-related policy is about policy related to "personal information", and the collection, storage, sharing, use and re-use of personal information is regulated by the Privacy Act 1993. The Act defines "personal information" as "information about an identifiable individual". Personal information is different from data about the weather or the state of a piece of infrastructure, in that the subject of the data has an interest in how that data is collected, stored, shared, used and re-used. In this sense, it is helpful to remember the context of personal data as set out in Figure 2 on page 8.

The broad international legal and social consensus is that personal data needs to be collected, stored, shared, used and re-used in ways that protect the interests of the data subject.

New Zealand's one-stop-shop approach to data regulation contrasts with the approach of the European Union, for example, which has a raft of data-specific regulations including specific rules to protect personal data in regard to employment relations, telecommunications services and credit rating agencies among other areas.

The advantage of the New Zealand regime is that entities need only become familiar with one piece of legislation, and can look to that legislation for guidance on dealing with all types of personal information. The disadvantage is that the Privacy Act has not been designed with DDI in mind, which means that it may not provide a complete framework for balancing the economic and social interest in enabling DDI with the personal privacy it seeks to protect. We think there could be gains available from alternative approaches, which we outline further below.

New Zealand also has laws and regulations that relate to other types of data or information: for example, firms have responsibilities regarding retaining their financial records for tax purposes, and some government information is classified for security reasons, but these are less likely to pose significant issues for DDI, and we do not discuss them in this report. In addition, companies and government agencies have their own policies on data collection, storage and use that reflect their view of the balance of interests for them and their stakeholders. For example, the New Zealand government has a long list of questions that agencies are required to work through if considering using cloud computing services. We talk about the impacts of some of these corporate and agency policies in what follows.

In DDI, as in other areas, clear and balanced policies provide certainty to firms and individuals, and thereby help encourage innovation and investment. For individuals, understanding that there are protections on what kinds of information about them can be collected, stored and used, and what protections there are against misuse creates certainty and security, and will tend to lead to an environment where individuals are more comfortable sharing information.

Table 1 compares New Zealand's personal information privacy rules with other countries, based on analysis by AnalysysMason supplemented by our research. The darker pink cells indicate laws that tend to present more of a "hurdle" to DDI, and the lighter cells indicate less restrictive laws. Compared with other jurisdictions, New Zealand has significant regulation around personal data collection, retention and repurposing, but does not specifically limit combination, profiling or international transfers, so long as the requirements of the Privacy Act are met. These requirements apply to New Zealand business and government agencies regardless of whether the data is held in New Zealand or offshore.

TABLE 1

COMPARING NEW ZEALAND'S PERSONAL INFORMATION POLICY SETTINGS WITH OTHER COUNTRIES

	COLLECTION		RETENTION	
New Zealand	Collection should be directly from the individual concerned, and the information collected must be necessary for a lawful purpose. The individual needs to be made aware of the fact that the collection is occurring, and a range of other facts, including the intended purpose of the information	3	Limited to the duration needed for "the purposes for which the information may lawfully be used". In most cases, this means limited to the duration needed for the original purpose	2
Singapore	Consent needed unless clearly in individual's interest and consent impractical or would not reasonably be withheld	2	Limited to duration needed for original purpose	2
South Korea	Consent needed unless interests of provider overrule those of user and data is not sensitive; collection to be minimal and service cannot be denied if consent withheld	4	Limited to period specified when obtaining consent	4
USA	If consistent with the context of the service provider's relationship with user and data is not sensitive	2	Limitation to duration needed for original purpose	2
EU	Consent needed except where interests of provider overrule those of user and data is not sensitive	3	As above	2
OECD	Consent where appropriate	1	As above	2
Australia	Differentiates between personal information, which may be collected without consent if it is "reasonably necessary" for the entity's functions, and sensitive information, which requires explicit consent. Sensitive information includes information about a person's race, political views, sexual orientation or criminal history. Collection of personal information should generally be directly from the individual and the individual should be notified of the collection and purpose for which the information is collected	1	Must be destroyed after it is no longer needed for the original purpose, a related secondary purpose or another purpose for which the information subject has given permission	2

Hurdle to DDI | Some challenges | Fewer challenges | Less restrictive laws

COMBINATION AND REPURPOSING		PROFILING AND DISCRIMINATION		TRANSFERS	
The general rule is that personal information obtained for one purpose cannot be used for any other purpose without permission. There are some limited exceptions. No specific rules on combination of data. Data can be de-identified and then combined or re-purposed. If analysis or combination of anonymised data results in re-identification of individuals, then the rules that cover storage and re-use of personal data apply to that re-identified data	3	Acceptable so long as no anti-discrimination legislation is contravened	1	Private entities can move data offshore provided the entity is still able to meet the requirements of the Privacy Act. Personal information cannot be transferred between entities without permission, but de-identified data can	2
If needed for agreed purpose; or clearly in individual's interest and consent impractical or would not reasonably be withheld	2	If needed for agreed purpose; or clearly in individual's interest and consent impractical or would not reasonably be withheld	2	If in individual's interest and consent is impractical; international transfers limited to organisations providing adequate protection	2
Combination not specified; assumed acceptable if consistent with context of the service provider's relationship with user	3	Not specified; assumed acceptable if needed for agreed purpose	3	Consent required for each specific third party	4
As above	3	If consistent with the context of the service provider's relationship with user	3	If consistent with the context of the service provider's relationship with user	2
As above	3	Opt out required	3	Provided adequate protections offered by other countries	3
Limited to original and compatible purposes	2	n/a	3	Where sufficient safeguards exist	2
Allows some secondary uses where they are related to the initial purpose. De-identified information can be repurposed. Combining data to re-identify individuals contravenes the general requirement to collect data directly from the individual	2	"Sensitive" information, which includes information about a person's race, political views, sexual orientation or criminal history cannot be used for targeting or profiling	3	Data can be transferred internationally, provided that the Australian entity is satisfied that the requirements of the Australian Privacy Principles will be met	2

SOURCE: ANALYSYSMASON (2014A), SAPERE AND COVEC ANALYSIS

The Privacy Act is built on principles. They include a principle of "data minimisation", which means that firms, government agencies or other entities collecting and handling data should collect and hold as little data as possible, and principle of "clarity of purpose", which requires that entities define how data will be used before collecting it. Taken together, these two requirements mean that New Zealand's regulation of the **collection** of personal data is at the more restrictive end of the spectrum. The Privacy Act also requires that collection should be directly from the individual concerned, and the information collected must be necessary for a lawful purpose. The individual needs to be made aware of the fact that the collection is occurring, together with a range of other information, including the intended purpose of collecting the data.

Entities collecting personal data in New Zealand are allowed to **retain** it for as long as they need it for "the purposes for which the information may lawfully be used". In most cases, this means retention is limited to the duration needed for the original specified purpose.

In general, **combining** personal data with other data sets is allowed, so long as the combining is part of the original purpose for which the data was collected.

"**Re-purposing**" personal data, using it for a purpose that is not the purpose for which it was collected, is not allowed without consent. However, anonymised data (personal data with any identifying elements removed) can be re-purposed, as the process of anonymisation removes the data from the jurisdiction of the Privacy Act, which only covers personal information. Note that, if the re-purposing or combining of anonymous data led to data subjects being able to be re-identified, the Privacy Act provisions would again apply, and the data could only be used for the purpose for which it was originally collected.

The **international transfer** of personal data (for example, sending personal data offshore to be analysed or stored) is acceptable under the Privacy Act, so long as the requirements of the Act are still met.

These principles have been constructed for good reasons but the implications for DDI have not be considered as part of the policy development process. Innovation, by definition, means doing something new. Along the way to a successful innovation there are likely to be lots of unsuccessful attempts and experiments. For DDI, data is the raw material of innovation, the experimental medium. In this context, New Zealand's relatively restrictive approach to data collection, in particular the requirement that the entity define the purpose the data will be used for before it is collected, is likely to constrain innovation. Personal data, in effect, cannot be used to new purposes that were not anticipated at the time of collection.

The limitations on data retention could also present a barrier to DDI. In effect, entities have to pre-define their purpose for collecting personal data, and then destroy the data once that purpose is achieved. A more open-ended approach to retention would allow data to be retained for future innovative uses, which might not emerge until some time after the data is initially collected, although of course it might raise greater concerns about the possibility of accidental data release.

In our view some aspects of the privacy regime deserve reconsideration with an eye to the potential benefits of DDI and an understanding of the relationship between risk and incentives to innovate. This is particularly important when considering any amendments to the current Privacy Act.

International consistency in regulation

Increasingly, DDI is something that happens in networks that span multiple countries. For example, a data analytics start-up in New Zealand might store its data in Australia, get some of that data processed by a firm in Japan and seek customers in the USA. Making a similar point, the OECD (2014, page 14) describes the "data ecosystem" as being "inherently global [in] nature":

> "For example, data may be collected from customers or devices located in one country through devices and apps developed in another account. The data may be processed in a third country and used to improve market ... to other consumers around the globe. Furthermore the ICT infrastructure used ... will rarely be provided only within one national border, but also distributed around the globe".

This international approach to business can be constrained or blocked if international rules about data, particularly personal data, are not aligned, or are confusing for firms or individuals - or if highly complex or badly designed regulations are adopted from other jurisdictions.

New Zealand's data regulation framework in the Privacy Act is recognised as meeting "international best practice" by the European Union (EU). This means that New Zealand is viewed by the EU as a safe and legal destination for EU companies to send personal data to be processed. EU practice around data is some of the most restrictive in the world, so EU acceptance suggests that other countries might find New Zealand's data regulations acceptable. However, this is a doubled edged sword - adopting a highly restrictive standard from the EU (a region where New Zealand now has fewer trading partners) may not provide significant benefit. We are not aware of any comprehensive cost benefit analysis of adopting EU data practices in New Zealand.

International data regulations have particular relevance for cloud computing, which can give smaller firms ability to access sophisticated IT services at an affordable price, and without needing expertise to develop and manage such services. KPMG (2012) recently estimated that cost savings from widespread cloud adoption could add between 0.15 and 0.2 of a percentage point to Australian GDP.

New Zealand government agencies are required to make a case-by-case assessment of cloud computing as a solution, with the caveat that any government information classified above "restricted" (i.e., "confidential", "secret" or "top secret") should not be stored in any cloud service. The Department of Internal Affairs (2014) outlines the risk assessment required of Government agencies before they make a decision about moving to cloud computing.

The Productivity Commission (2014, page 228) raised concerns that DIA's advice is "almost exclusively about risk management"; and certainly a requirement to fill in a 104-question questionnaire does seem like a strong signal that agencies should be cautious. That said, DIA's own Government ICT Strategy and Action Plan to 2017 sees "cloud computing platforms providing the majority of government's computing resource" by 2017 (DIA, 2013). This seems to be a significant vote of confidence in cloud computing by the government, and raises the question of mixed messages from the DIA.

Even if the government is positive, New Zealand businesses appear to be cautious in using cloud services with offshore storage. In our interviews, we heard from a number of companies that they keep their data in New Zealand because, although this is more expensive than offshore options[5], they are unclear about the status of their data if they store or analyse it offshore.

Common concerns amongst our interviewees regarding using offshore cloud storage included:

- feeling comfortable that personal data, or information that is commercially sensitive will be appropriately secure;
- understanding the rules that apply to data in the offshore jurisdiction, and therefore being able to make an informed decision about whether those are appropriate for their data (e.g. what level of powers the government in the host jurisdiction has to intercept or seize data); and
- understanding the responsibilities placed on them by New Zealand law and feeling comfortable that they are complaint with them.

These responses suggest that firms are still learning how New Zealand's data regulation relates to that of other countries, and that greater understanding of this relationship could encourage competition between cloud services and efficiency improvements for firms.

The government may have a role to play in helping simplify matters for firms and reassuring consumers. The Productivity Commission (2014, page 12) recommended that the New Zealand government pursue "free-trade-in-data" agreements with other countries, with the aim of ensuring that the rights and responsibilities of data subjects, guardians and users do not vary depending on the physical location of their data.

There are already international efforts underway to establish cross-jurisdiction agreements on data regulation. For example, APEC's Cross-Border Privacy Rules (CBPR) system is designed to protect the privacy of consumer data moving between APEC economies by requiring companies to develop their own internal business rules on cross-border data privacy procedures. The system is still in development, with only a few countries having joined. New Zealand may consider signing up.

[5] Some evidence of the cost differential is given by the New Zealand government's own business case for using cloud services, in which it estimated net benefits over five years of using offshore cloud services were approximately 30% higher than using onshore cloud services (New Zealand government, 2012).

Decision making about data regulation

Policy-makers in data-related areas need to make a balanced assessment of the economic and social benefits of the use of data, as well as potential harms that misuse of data can generate if they wish to create a policy framework that enables and supports DDI. But overlooking the potential benefits in favour of regulation designed to protect privacy is a theme that is picked up in some of the international literature on data regulation, and echoed in recent New Zealand experience.

For example, Lerner (2011) finds that the European Union e-Privacy Directive, which regulates the electronic collection and use of personal data more tightly in the EU than in other countries, could have generated societal harm in Europe equivalent to a €1 billion reduction in R&D. Whether or not this potential lost investment is justified by other goals remains an open question, but negative effects on innovation as a result of data regulation do need to be considered as part of the policymaking process.

The Conference Board of Canada (2012), discussing the Canadian approach to personal data, says:

> "Canadian privacy regulators and advocates approach privacy issues chiefly from a legal perspective. The regulators have little capacity to consider the economic impacts of their regulatory activities. Although a legal perspective is valid, it needs to be complemented by economic and business perspectives that take into account how private information is used in a modern economy"

It seems that the same can be said about New Zealand. For example, the Productivity Commission in its report on productivity in the services sector (Productivity Commission, 2014), noted that:

> "a significant amount of ICT-related regulation is designed to minimise harm... a risk with a narrow 'harm-minimisation' approach is that it fails to take a wider economic perspective and recognise harms [caused by the regulations they have introduced] elsewhere in the community/economy".

What the Productivity Commission is saying is that policy-makers have focused on trying to prevent specific ICT-related harms, such as preventing privacy breaches by restricting collection of data without consent, and they have not sufficiently weighted the social and economic loss from reduced innovation that such regulations may create.

As noted above, the Law Commission (2011) has recently reviewed the Privacy Act in work that led to the new Privacy Bill. As far as we can see, the Commission's report does not explore the economic issues around privacy at all. The report considered compliance costs for businesses, but we did not see in our review any evidence that wider economic effects, including potential costs from restricting the use of personal data were considered. This is despite the Law Commission itself noting that the review was necessary because of changes such as (page 249 emphasis added):

> "the ease with which personal information is distributed around the globe, the pervasiveness and longevity of personal data, and the *creation of commercial opportunities* and *incentives to commoditise personal information* due to the *increasing economic value of consumer data*".

The Regulatory Impact Statement that accompanied the current proposals to change to the Privacy Act states that (Ministry of Justice, 2014):

> "analysis of options has been constrained by lack of empirical evidence about the direct and indirect economic costs and benefits of privacy law settings to individuals, business and Government".

These local examples point to a need for a more coherent view of the value of DDI in the New Zealand policy system, some efforts to estimate the economic costs and benefits of privacy controls in a way that is useful for policy, and consideration of both these matters as a standard part of analysis on data-related issues. Existing agencies making these policy assessments, such as the Law Commission, seem to be focused on legal and 'pure privacy' concerns, and may not have a sufficiently full understanding of the possibilities of DDI to provide a complete view on their own.

As one of our interviewees noted, the ease and speed of data collection can transform the way that policy is made in some areas. Instead of a long problem definition phase and consideration of options, with a monolithic big-bang response and a review in three years, policy-makers can monitor their areas of interest continuously, intervene more rapidly but with more subtlety, see quickly if the intervention is working as planned, and change it if it is not. Policy cycles could be measured in months rather than in years. However this kind of approach is not appropriate for all areas of government endeavour, since quality data gathering is far from universal, the costs of experimentation could be very high, and it is not necessarily desirable for governmental systems to always move rapidly. But in policy areas like the use of data where there is significant innovation and industry disruption, and where government has a major role to play, policy experimentation might deliver greater rewards than other approaches.

The risk with rules that reduce apparently undesirable activity is that they also curtail what could have been useful innovations, but policy-makers do not see the innovations that do not happen. This puts greater emphasis on thinking about the least intrusive and most focused way to resolve a perceived problem, and on thinking about potential upsides as well as evident downsides (and how big those upsides might need to be to be worthwhile).

In situations of great uncertainty, there is value in waiting to see what further information emerges so as not to curtail options and the productive efforts of private actors. But this need not mean that governments should not do anything in innovative policy areas. The new age of data means that governments can also experiment. Unfortunately government agencies tend to be focused on avoiding embarrassment and minimising waste. This is not always the most productive mind-set. Disciplined experimentation within bounds would likely be a better strategy.

In particular, a process of disciplined policy experimentation in some areas would stand in contrast to the traditional strategy thinking of "here is where we are", "here is where we want to get to", and "these are the steps in between". This approach works in situations of low change and where the impacts of pulling particular levers are well understood. These are not characteristics of policy making around DDI.

Policies that actively support DDI

Data Infrastructure and Open Data

Government agencies collect, store and manage a wide range of data, from individuals' health data to information about the state of our national parks. Government has a role in ensuring that New Zealanders have appropriate access to this data to enable us to move towards our economic and social goals.

The OECD (2014, page 24) suggests thinking about data as a type of infrastructure, defined as "the basic equipment and structures ... that are needed for a country ... to function properly", arguing that data has similar economic characteristics: One person's uses of data does not affect its ability for someone else to use it; demand for data is not for data itself, but for other activities that require data as an input; and data may be used as an input into a wide range of goods and services.

To this list of characteristics, we would also add that data can be cheap or free to create (often because useful data is a by-product of another process), and cheap to store, process and share. Like infrastructure, tools used to process and share data can improve as technology improves.

The OECD says that governments have a role in providing open, accessible data as "positive externalities in combination with open access can lead to a 'comedy of the commons', where greater social value is created with greater use" (OECD 2014, p.24). Note that this does not necessarily mean that all data should be open or free, but the OECD does argue that governments have a responsibility to ensure that access is fair and enables innovation and development.

The OECD talks about the benefits of open data, saying (2014, page 19):

> "The use of open data by citizens as provided by governments through their open data initiatives, for example, can increase openness, transparency and accountability of government activities and thus boost public trust in governments. At the same time, it can enable an unlimited range of commercial and social services across society... Estimates on the economic impact of PSI [Public Sector Information]... focus on the commercial reuse of PSI and thus do not cover the full range of (social) benefits."

and (page 28), connecting open data with the speedy technology change and uncertain development path of the ICT sector:

> "In particular, in environments characterised by high uncertainty, complexity, and dynamic changes open access can be an optimal (private and social) strategy for maximising the benefits".

The Productivity Commission (2014) noted that New Zealand creates a very small proportion of global ICT products. The main way it will benefit from ICT is from cleverly adopting and using ICT that is developed and produced overseas. The same observation applies to DDI: most of the benefits in New Zealand will come from adopting innovations that have been generated elsewhere. However, in order to adopt these innovations, New Zealand needs its own data infrastructure, and the necessary skills and understanding to take advantage of the opportunities.

The New Zealand government has had a focus on open data for some years and New Zealand is ranked among the top few countries in the world in terms of open data.

- Open Data Research's 2013 *Open Data Barometer report* ranked New Zealand as 4th in the world in terms of the overall openness of government data.
- The 2014 *Global Open Data Index*, ranked New Zealand 5th out of 97 countries.

These rankings reflect the New Zealand government's sustained effort to create an effective framework for open data, and encourage its release. However, there is still some distance to go to realise the aspirations that government has for the quantity, quality, accessibility and usability of the open government data. We see an opportunity for New Zealand to step up its efforts in the sphere of open data, to retain its globally-leading status and to continue to stand-out in an area where increasing numbers of countries are focusing efforts.

New Zealand government data is required to be managed in line with the 2011 "Principles for Managing Data and Information held by the New Zealand Government", which states that data should be open, protected (for personal information), readily available, trusted and authoritative, well managed, reasonably priced and re-useable.

Some of the interviewees we spoke to considered that much of what is considered "open data" by the government is online but often not in a useable format, not "clean" and is published irregularly so cannot counted on as a reliable input.

The latest report on progress with open data (New Zealand Government, 2014) supports these concerns:

- In 2014, nearly 80% of agencies released datasets, but only 45% of agencies released datasets in open formats. This is, however, a significant improvement from 2012, when less than 20% of agencies released information in open formats.
- Government agencies released 162 high value datasets in 2013, and 163 in 2014.

The total number of datasets on the open government data portal was 2,603 as at December 2014, although reporting indicates that the proportion of datasets that are listed on data.govt.nz is falling over time, meaning that this number is an underestimate

of the total amount of open government data available. There is no data available on what proportion of the total number of public datasets that could be published that figure represents. The amount of data available from open government portals is far higher in the USA (134k) and the UK (16k) but of course they are much larger countries.

Although there has been a significant effort to advance open data, making available data for public release is only the first part of the story. This is because even open data requires skills to find and manipulate the data that most people do not have. To make use of data, a person has to know the data exists, know which agency collects it, navigate to a website and either download a spreadsheet or work his/her way around a web-tool that is different for every agency to secure the numbers. From there, the data will need to be loaded into a spreadsheet or some other software and manipulated to get the specific figures of interest. While such data may be "open" in the sense that it has been published, this does not mean the data is accessible to most potential users. This remains unfinished business for all open data programmes: open data does not yet mean *useable* data.

We are aware of substantial ongoing effort to make more data more accessible. For example, Land Information New Zealand's LINZ Data Service makes a large quantity of geo-spatial data clearly licensed and freely available for use, download and re-use. There is a host of case studies talking about how others are using LINZ's data for their own initiatives.[6]

And we are also aware that there are still enormous opportunities to improve New Zealand's data infrastructure. For example, one of our interviewees is building an extensive asset database for Wellington City Council that tracks the location and condition of the $6.5 billion of physical assets that the Council holds. Combined with population and growth projections and spatial and financial data, we were told the Council projects it will save $60 million in unnecessary renewals spending over the next ten years. There are limitless opportunities for projects across the public sector that could provide significant opportunities for innovation and efficiency.

Ensuring the availability of appropriate data skills

DDI needs more than available data and appropriate regulation. There need to be people with the right skills, and businesses that understand the benefits of DDI. The OECD says (2014, page 8):

> "The anticipated benefits in productivity growth from DDI depend on a number of enabling and complementary factors, including in particular (i) the level of skills available to organisations, and (ii) the readiness of organisations to change their internal and external business processes".

At present the government does not appear to have a specific focus on ensuring that the education system creates graduates who can participate in DDI, by which we mean both graduates who are data-literate and can work with and understand data, and high-end data scientists. We note that this shortage of people with data-skills is world-wide, and only likely to get worse. This suggests that we will need to focus on how we can grow these skills at home, as we will not necessarily be able to import people to fill our gaps.

[6] See http://www.linz.govt.nz/about-linz/what-were-doing/success-stories

Businesses also need to be aware of the benefits of DDI to start thinking about how to take advantage of the possibilities and to start to hire the people with the right skills. Governments have a range of ways of supporting businesses to develop their capability, from general education programmes to highlight targeted ways of working with businesses that have been identified as likely to benefit from intensive support in a specific area.

Incentivising an innovation culture and encouraging business action

Firms are engaged in a learning process with regard to DDI. It will take time for them to figure out how DDI can help them, and time to invest in the two complementary intangible assets that the OECD says are essential to get value from ICT investments (ICTNET, 2012):

- Investment to develop organisational conditions that support the use and application of ICT in the firm's production process, and
- Investment in human capital, so that labour resources are able to use the ICT tools and applications efficiently, and to drive productivity improvements.

The Productivity Commission (2014, page 199) says that "New Zealand firms are slow adopters of ICT compared to other developed countries". The Commission points out (page 189) four factors in particular that might cause New Zealand firms to delay investments in ICT. These are smaller domestic markets, smaller firm sizes, distance to larger markets and low labour costs that make it profitable to use less capital-intensive business models. The Commission also mentions poor management quality and a general lack of competition in New Zealand, with the evidence for the latter stronger than for the former.

There could be a useful demonstration role for government. During our interviews we spoke to a wide range of people in the public sector using data in innovative ways. We encourage these efforts: Government needs to be not only a supporter of DDI, but should be one of the most enthusiastic innovators. It can use its position, its wide service provision responsibilities, and its broad overview of social and economic value, to innovate in ways that would be inappropriate or impossible in the private sector. The OECD comments that "government should lead by example by considering the policy opportunities in the context of the public sector which is one of the most data-intensive sectors in many OECD countries." (OECD 2014). A strong government DDI focus will generate demand and build capability in New Zealand's DDI ecosystem, as well as creating direct social and economic benefits for New Zealanders.

The work of the Data Futures Forum was a standout effort last year. The Forum put forward a range of ways the government could "incentivise an innovation culture" around data, which are all potential ways of supporting the DDI ecosystem. We support the approach promoted by the Data Futures Forum, of the government allowing more experimentation and greater risk tolerance. We also note that an experimental approach is particularly appropriate in an innovative and fast-moving sector like DDI.

One possibility the Forum raised was for shared effort on collaborative "catalyst projects" that create value by solving real data problems or making new data assets available. One example presented was looking at linking supermarket shopping data from loyalty cards with health outcomes data to build a research platform for studying the impacts of diet on health. At the heart of the catalyst approach seems to be an experimental decision-making model: projects could be monitored to see if they produce worthwhile results, and stopped if they do not. Exhaustive before-the-fact cost benefit analysis of new initiatives need not be required, reducing the requirement to quantify the potentially unknowable benefits of what is, by definition, an innovative activity. We understand that the government is considering its formal response to the Forum's recommendations at present.

In terms of direct support for DDI businesses, policy settings that promote entrepreneurship generally, such as seed funding, incubators, and ensuring appropriate capital markets to support growth, will also support new data analytics businesses.

New DDI business will have data in their DNA, but effective use of data can be a challenge for more established businesses and for government. To truly gain the potential benefits of DDI, many businesses, and government entities, will need to change how they do things, from decision-making to organisational structures.

A way forward for data regulation in New Zealand

The following are some specific suggestions that could be considered by government to promote DDI and help close the gap between New Zealand and other countries in realising benefits.

Champion the value of DDI

One way to have the benefits of DDI more fully emphasised and understood could be through a "data champion" of some description. There have been several proposals for new roles in the data policy system, from a Data Commissioner to a Government Chief Data Officer. The Data Futures Forum (2014a) proposed both an independent "Data Council" to "act as guardians of the system", and a "champion" to "drive innovation and data-sharing".

We agree that a data specialist could improve data policy and use across government. This could be a "Chief Data Advisor", similar in concept to the existing Chief Science Advisor. A Chief Data Advisor could provide independent advice on policy proposals, with a focus on enabling DDI. In our view this could help balance the policy-making system, and produce more effective data-related policy as a result, without necessarily adding another regulatory layer or hurdle for the application of DDI.

A Chief Data Advisor could be responsible for:

- promoting the effective use of data by governments, businesses and the general public;
- assessing data-related government policy proposals from an economy-wide perspective;
- raising public awareness of the benefits of effective data use; and
- promoting good practices and ethical use of data by providing guidelines and raising awareness.

Deepen the debate about personal data

An early step in effective policy-making is to understand the problem that is being solved. At present, the debate about protecting the interests of the subject of data is often framed as a debate about privacy. In part, this is accurate. However, in our view, re-framing the discussion away from simply "privacy" and towards preventing the misuse of personal data would enable New Zealand to establish a data regulation regime that would both reduce the harms suffered by data subjects through the mishandling of their personal data, and help enable DDI for personal data within bounds.

Misuse of personal data can lead to a range of harms that are broader than the "pure privacy harm" of personal information being shared without consent. For example, mishandled personal data could lead to the data subject being discriminated against by insurance providers or banks; becoming the victim of identity theft or fraud; or having his or her physical security threatened.[7]

Wittes (2011), a leading proponent of the view that privacy does not provide an adequate conceptual framework for data regulation, says:

> "Our lives are described by a mosaic of data - an ever-widening array of digital fingerprints reflecting nearly all of life's many aspects. Our mosaics record our transactions, our media consumption, our locations and travel, our communications and our relationships. They are, quite simply, a detailed portrait of our lives - vastly more revealing than the contents of our underwear drawers, yet protected by a weird and incoherent patchwork of laws that reflect no coherent value system"
>
> "[T]echnology's advance and the proliferation of personal data in the hands of third parties has left us with a conceptually outmoded debate, whose reliance on the concept of privacy does not usefully guide the public policy questions we face... the term privacy has become a crutch - a description of many different value of quite-different weights - that does not usefully describe the harms we fear".

Wittes suggests supplementing privacy rules with other concepts, so that society and policy-makers can have a more precise discussion about exactly what we are trying to protect. He offers the concept of "Databuse", "the unjustified deployment of user data in a fashion adverse to the user's interests", as a supplement to the concept of privacy.

Clearly identifying the specific harms that society wishes to prevent or sanction, may enable a better balance to be struck between protection and innovation.

[7] As an example of the breadth of issues currently grouped under the concept of privacy, the Office of the Privacy Commissioner (2014)reported on a survey that asked people how concerned they were about ten "privacy" issues. These issues covered a wide range of topics, including "the information that children put on the internet about themselves", "your credit card or banking details being stolen", and "surveillance of New Zealanders by agencies of overseas governments".

Consider new policy approaches to prevent harms and enable innovation

Once the specific harms we wish to avoid are identified, there are a range of general policy approaches to preventing them, from prevent actions that create a risk of harm (risk avoidance), to sanctioning those causing the harm after it has occurred to clarify the rules and deter future harm.

At present, New Zealand data regulation tends towards the risk avoidance approach, by limiting data collection, retention and re-purposing, in order to avoid the risk of pure privacy harms, and other potential harms such as fraud, stalking or blackmail.

Abrams (2014) comments that "[m]any national laws were enacted before analytic-based research was well understood. National laws that provide only one legal permission mechanism to process personal data, for example, explicit consent for research, are particularly problematic". New Zealand's current framework, which requires consent from the subject for collection and repurposing of personal data, falls into this problematic category. Abrams goes on to say that, "avoiding processing [or using data] because one finds resolving the conflict between risks too difficult, leads to the loss of meaningful benefits to individuals and society as a whole".

The Data Futures Forum (2014a) has also suggested that there are some cases where the collective interest in achieving a public good overrides the data subject's "right to privacy". In other words, there are circumstances where an agency could share or use personal data without specific permission and this would not constitute an abuse of personal data. An example might be sharing data in order to identify candidates for a health screening programme, or families that would be likely to benefit from extra support from the state. This type of situation offers significant scope for social benefit through, for example, better targeting of health and social interventions. It also carries a higher risk of abuse of the data. Cavoukian (2014) supports this kind of constructive approach rather than focusing on the risk of re-identification of anonymised data, and criticises the exaggeration of risks of anonymised data that tends to sway policy-makers into making rules that make it hard to share data.

Together these approaches suggest that New Zealand should analyse the costs and benefits of a new regulatory framework that enables more data sharing, re-purposing and re-use without explicit consent in some circumstances, but which also more heavily sanctions incidents of actual harm. For example, New Zealand could consider a legal framework that loosens the current restrictions on data collection and re-use, but makes it an explicit offence to use personal information in a way that causes a defined harm.

The balance between preventing risk or harm and enabling innovation is a decision made through a process of debate by policymakers, firms using data, and the public at large. We think that there is scope to deliver greater flexibility in the collection and use of data while continuing to minimise risks to individuals.

Data comes to the farm

When to sell stock. When to stop milking. When to breed. In farming, even the smallest decision can have a serious impact on your bottom line. A new data modelling tool called Farmax is helping to take the guesswork out of farming, predicting budgets and profits by modelling everything from grass cover to stock performance. Farmax is already boosting margins per hectare for farmers all over the country, contributing more than $450 million to their businesses.

5. Appendix – Review of existing studies of DDI

In this section we provide a brief overview of the existing literature on the value and effects of DDI.

We divide the literature we reviewed into three groups:

- Studies of the value of DDI in general – The recent papers estimating the value of DDI in Singapore, Japan and Australia are the most directly comparable with our work, and we compare our results with theirs in another part of this report. There are also some academic papers looking at the value of DDI in general to businesses.
- Studies of the value of particular datasets or approaches – Conceptually these are a sub-set of the value of DDI. They include studies of public sector information, "open" data and "big" data, as well as a study of the value of the Census to New Zealand.
- Studies of policy issues – Some authors have looked at the policy questions that DDI has raised. This includes the interesting study of Lerner (2012) that estimates the impact of the EU's e-Privacy Directive on investment by venture capital firms in online advertising companies. There are important questions about whether policy settings have kept pace with the extensive sharing of data that now takes place. We discuss some of these issues, and others, in the policy section of this report.

We exclude from our review studies of the economic impacts of the Internet or ICT more generally. There is extensive literature on these impacts that is reviewed in Glass et al (2014; a previous paper for Google and other members of the Innovation Partnership), ICTNET (2012) and Katz (2012).

We have also excluded from our review papers that just discuss the potential applications of DDI or its impacts in specific cases. It is clear that the possible uses of DDI are extremely numerous. We have included six case studies in this report that cover a range of examples, including modelling likely pasture and animal growth on farms, calculating carbon footprints, and analysing shopping behaviour through loyalty schemes. Just a few other examples we have come across in the literature include reducing hospital admissions and emergency visits with a tele-health programme in the UK[8], increasing student pass rates and reducing dropout rates with teaching systems that adapt to student performance in the US[9], analysing German federal government employment programme performance more accurately in order to improve or eliminate ineffective efforts[10], and the introduction of real-time bus information in Chicago (Tang, 2012).

The value of DDI in general

AnalysysMason (2014a) estimates the value of DDI to the Singaporean economy. The authors estimate that DDI and its associated services contributed SGD4.3 billion to the economy in 2013 across seven studied sectors, a figure equivalent to around 2% of total GVA in those sectors. The authors estimate much larger impacts in the future, with DDI and associated services expected to contribute SGD6 billion in growth in GVA by 2018.

AnalysysMason separates out different types of DDI services as in Table 2, and calculates the cost savings and revenue increases that could be enabled by each service type within each of seven sectors: health, education and social services, ICT, transport and logistics, financial services, trade, and manufacturing. These revenue and cost impacts are then scaled based on estimated adoption levels of each service in each sector.

[8] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215264/dh_131689.pdf
[9] http://www.forbes.com/sites/bruceupbin/2012/02/22/knewton-is-building-the-worlds-smartest-tutor/
[10] Cited in McKinsey (2011)

TABLE 2
DIFFERENT TYPES OF DDI SERVICE

DDI SERVICE	DESCRIPTION
The new bazaar	Services that allow firms to offer one-off, personalised products at one-off, personalised prices (if they are not free)
Services here and now	Services that help users complete their 'real world' tasks in real time by combining location data with other pertinent information
Attentive services	Services that seek to "understand" users' behaviour and circumstances to anticipate their needs and proactively change the service provided
Service for people like me	Services that target groups of consumers who share common characteristics
Intelligent planning	Services that extract patterns from the analysis of large, often disparate and/or anonymous datasets to help firms or governments prepare and respond to large-scale social trends

SOURCE : ANALYSYSMASON (2014A)

AnalysysMason (2014b) takes a similar approach in their study of DDI in Japan, estimating that DDI and associated services contributed over JPY7 trillion to Japan's economy in 2013 across the same seven sectors, equivalent to about 3.4% of Japan's overall GVA. The authors estimate that by 2020 the value of DDI in Japan could be more than twice as large, at more than JPY15 trillion.

The sectoral impacts vary from a modest 2% of GVA in health, education and social services (although in these areas the non-market values are potentially higher), to a much larger 10% of GVA in transport and logistics and manufacturing. AnalysysMason estimates that the biggest impacts in Japan will be in their categories of "services for here and now", like mapping services or fleet management systems, and "intelligent planning", like analysis of historic traffic congestion data to aid in urban planning.

This paper also has an extensive policy discussion, which we refer to later on in this review.

PwC (2014) studies the economic impacts of DDI in Australia. The authors estimate that DDI added A$67 billion in value to the Australian economy in 2013, equivalent to around 4.4 percent of GDP or roughly the same contribution as the retail sector. The authors suggest that Australia has substantial room to improve, with A$48 billion of additional value available but not captured from DDI in 2013.

Barua et al (2011) examine the impacts of effective data use for businesses based on a sample of 150 Fortune 1,000 companies. The headline result is that relatively small improvements in firms' effectiveness in the use of data led to significant improvements in financial returns.

The study is in three papers that look at:

- Financial impacts, including productivity of employees, return on equity, return on invested capital, and return on assets.
- Customer-focused impacts, like the ability to innovate to derive revenue from new products, and the ability to expand the existing customer base.
- Operational impacts, for example asset utilisation, the accuracy of planning and forecasting, and on-time delivery of products or services.

The study analyses performance metrics for the firms and associates that with the results of a survey that collected information from each firm on the quality, sophistication of use, usability, remote accessibility and the use of portable systems to exchange data with customers.

In a similar vein:

- Brynjolffson et al (2011) found that the use and application of "data-driven decision-making" makes firms 5-6% more productive based on survey data on the business practices and information technology investments of 179 large publicly traded firms.
- La Valle (2011) reports the results of a survey of 3,000 businesspeople globally that found that higher-performing organisations were more likely to apply analytics than intuition when making decisions than lower-performing organisations. The tendency to use analytics was greater in higher-performing firms across a wide range of business functions.

The value of particular datasets or approaches

McKinsey Global Institute (2011) discusses how the use of large datasets can create value for businesses and government agencies by:

- creating transparency, just by making data more accessible,
- enabling experimentation to discover needs, expose variability and improve performance,
- segmenting populations to customise offers, including in the public sector where these techniques are much less common,
- replacing or supporting human decision-making with automated algorithms based on data analysis, or
- innovating to generate new business models, products, and services.

The authors find that the potential value of big data is very variable across sectors because of the varying prevalence of four types of barriers: a lack of appropriately talented workers, insufficient IT intensity, the lack of a data-driven mind-set, and lack of data availability. The authors suggest that five sectors in particular are good candidates for large financial gains from the use of big data: healthcare, retail, public sector administration in Europe, manufacturing, and personal location data. The analysis is based on US and international case studies with reported impact numbers and expert interviews.

Other similar studies include the following:

- PwC (2012) estimates the potential value of big data for Australian retail and consumer businesses from additional revenues and reduced promotional spending. The authors estimate an earnings before interest and taxation (EBIT) value of $1.3 billion in increased revenues and $2.5 billion in avoided costs.
- Bakshi (2014) finds that businesses that make greater use of online customer and consumer data are 8-13% more productive as a result, although "amassing data has little or no effect on its own".
- Tambe (2014) analyses the profiles of 175 million user profiles on LinkedIn to identify those with skills in specific big data-related technologies. Firm investment in those technologies was associated with 3% faster productivity growth, but only for firms with access to big data sets, and only for those that were well-connected to labour networks with the appropriate expertise.
- OECD (2014, page 5) notes evidence from Japan that manufacturing companies using data analytics could generate maintenance costs savings worth almost 15% of shipment costs in 2010 and significant electricity savings. It also quotes an example from agriculture, with geo-coded maps of fields and real-time monitoring of every activity able to increase farmers' yields.

McKinsey Global Institute (2013) discusses the potential economic benefits that can be realised by providing greater access to open data in seven "domains" across the global economy: education, transportation, consumer products, electricity, oil and gas, healthcare, and consumer finance. The study estimates global value of as much as $3 trillion annually across these domains from open data. This is a large number - for comparison, the total value of all goods exports was $19 trillion in 2013, according to the World Trade Organisation.[11]

Deloitte (2011) is also a study of open data, focusing on the value of improving access to information held by public sector organisations. It is based on a review of practices in 21 public sector bodies across Europe, endeavouring to determine which charging model for publicly held information works best. The authors conclude that lower charges are generally better, estimating the value created from reducing charges for information in benefits for those who reuse the data, and in changes in cost for the public sector bodies themselves.

[11] http://www.wto.org/english/news_e/pres14_e/pr721_e.htm

The study shows very substantial increases in re-use of public data from reduced charges and more open availability, e.g. from moving to Internet distribution. Barriers to lower prices include dealing with the loss of revenue from lower prices (although in some cases lowering prices increased revenue because of the extra demand that it created), and how to manage the transition from existing arrangements that involve a commercial partnership with a single company that then has incentives to delay or prevent a move to more open data.

The study also includes a review of the mobile apps market, motivated by the relevance of more open public data for application developers to use, and an assessment of the state of play on open data portals, the standard method agencies use to make data more available.

Vickery (undated, but likely 2011) is a review of recent studies on the value of public sector data reuse. The author estimates the direct value of the reuse of public sector information at €32 billion for the EU27 in 2010, with the indirect benefits adding another €108 billion a year for a total economic impact of €140 a year. The study suggests that moving to a more open data stance, with easy access to data for free or at marginal cost could add another 40% to this total, making open data worth around €200 billion in 2010, around 1.7% of EU27 GDP. The study cites substantial cost savings through efficiency gains for public bodies from open data, but notes that revenue from data users is likely to be only modest because the greatest benefits come from the widest possible distribution of data.

Fioretti (2010) also focuses on open data, reviewing the role of public sector information in society, and considering the potential of open data. The author cites estimates of the value created from other countries: a £6 billion boost to UK GDP from that government's open data plans, around €15 million in direct financial benefits to Denmark from open publication of that country's official address database, or €1.6 billion in market revenues for geo-spatial information in Germany in 2006, with more than half of that value created by a navigation market based on services provided for free to consumers.

Bakker (2014) is a New Zealand study that assesses the economic costs and benefits of the use of Census and associated population statistics information. The study estimated the potential value based on the areas where Census information is primarily used (e.g. funding of education services). Local and central government organisations, as well as the private sector were examined, however very limited attention had been given to the specific types of applications of data innovation by consumers and businesses.

Policy issues

AnalysysMason (2014a, 2014b) separate out policies relevant to DDI into positive, proactive policies (like making government data more open), and restrictions and consent requirements (like limits on collection of personal data, or transfer of data overseas). Both studies emphasise the particular importance of allowing combination and repurposing of data, allowing international data transfers, and not imposing restrictive consent requirements in order for DDI to flourish. They suggest some alternatives to these types of controls that might achieve the same goal but with less impact on worthwhile DDI activity.

The authors of the Singapore study (AnalysysMason, 2014b) go slightly further; they suggest minimising restrictions on DDI and carefully considering the potential value that may be lost from any controls in this period when the potential applications of DDI are so uncertain.

PwC (2014) says that improved DDI performance by the government is crucial to increasing the contribution of DDI to the Australian economy. It recommends that the government prioritise the provision of open data and provide senior political leadership to "get on with it" in order to support wider innovation by other players. PwC argues that the health industry offers the biggest opportunity for Australia to boost data-driven innovation.

Use of DDI by firms can raise tricky privacy issues. As noted in the body of our report, there is literature on the difficulties of effective anonymisation. Sweeney (2000), using 1990 US Census data, finds that 87% of the United States population were uniquely based just on five-digit ZIP code, gender and date of birth, and about half could be identified based just on gender, date of birth and city/town of residence.

De Montjoye (2013) studies the uniqueness of user movements in anonymised cellphone call data that includes the cellsite that was used for a call. The authors conclude that user movements are quite unique, ie, that very few location reads are required to isolate a particular user within even quite a large dataset of 1.5 million customers over 15 months. Narayanan (2008) demonstrates the difficulties of ensuring that individuals cannot be re-identified within anonymous datasets by combining an anonymised dataset with a public dataset that includes similar data.[12]

Cavoukian and Castro (2014) argue that reporting of these studies substantially exaggerates the risks of re-identification, and that actually the problem is that datasets were not properly de-identified in the first place (page 1):

> "Contrary to what misleading headlines and pronouncements in the media almost regularly suggest, datasets containing personal information may be de-identified in a manner that minimises the risk of re-identification, often while maintaining a high level of data quality .... While these articles raise important issues concerning the use of proper de-identification techniques, reported findings do not suggest that de-identification is impossible or that de-identified data should be classified as personally identifiable information."

[12] Techliberty provides an example of something similar closer to home. http://techliberty.org.nz/1-day-finds-that-anonymity-is-hard/

Abrams (2014) is an attempt to define a common ethical framework for use by those embarking on and those governing big data analytics projects. The five basic values in the framework involve defining the benefits that would come from the project; ensuring that big data analysis will add additional value compared with other approaches to counter the additional risks of bringing together large and diverse datasets to look for hidden insights; ensuring that analysts understand how long the algorithms they create are valid for; ensuring the project respects the interests of all of the affected parties; and that the analysis is fair. Assessment against these five principles would be combined with ensuring that any project has a clearly defined purpose that explains why the analysis is being undertaken.

The Conference Board of Canada (2012) emphasises the need to "uncover" the economic impact of privacy policy, laws and regulations, since most of the costs and benefits of commercial privacy policies are uncertain or hidden from view. It argues that privacy risks are managed through a combination of individual, company and government actions, but most of the attention is paid just to the third lever. It also says that privacy regulators and advocates largely come from a legal perspective, with little attention paid to the economic or commercial impacts of their regulatory efforts.

The paper says the annual direct fiscal cost of the Canadian privacy regulators is CAD$40 million, but the total annual economic costs is around 900 times higher at CAD$3.6 billion, with negative flow-on impacts on company behaviour and business investment. Economic costs include process and administration efforts, changes to business processes, and impacts on innovation and market efficiency, and hidden benefits include improvements in market functioning due to privacy regulation that increase levels of trust between consumers and their service providers. The authors do not see compensating benefits that justify these large costs. Sector analysis in the paper reveals the retail trade and financial services sectors are worst affected.

The authors also point out the limitations of the concept of consent. In many cases, (such as retail loyalty cards) consumers consent to share their private information in return for goods and services that are less expensive. But in other cases there is no meaningful consent possible because provision of personal information is a condition of availability of the service, or because the conditions associated with consent are very confusing or very time-consuming to understand.

Lerner (2012) argues that decisions about the scope of privacy regulation can have significant impacts on investment and innovation. The paper examines the effect of privacy policy changes on venture capital investment in online advertising companies analysing the effects of the European Union Privacy and Electronic Communications Directive ("EU e-Privacy Directive"), which regulates the electronic collection and use of personal data more tightly than in other countries.

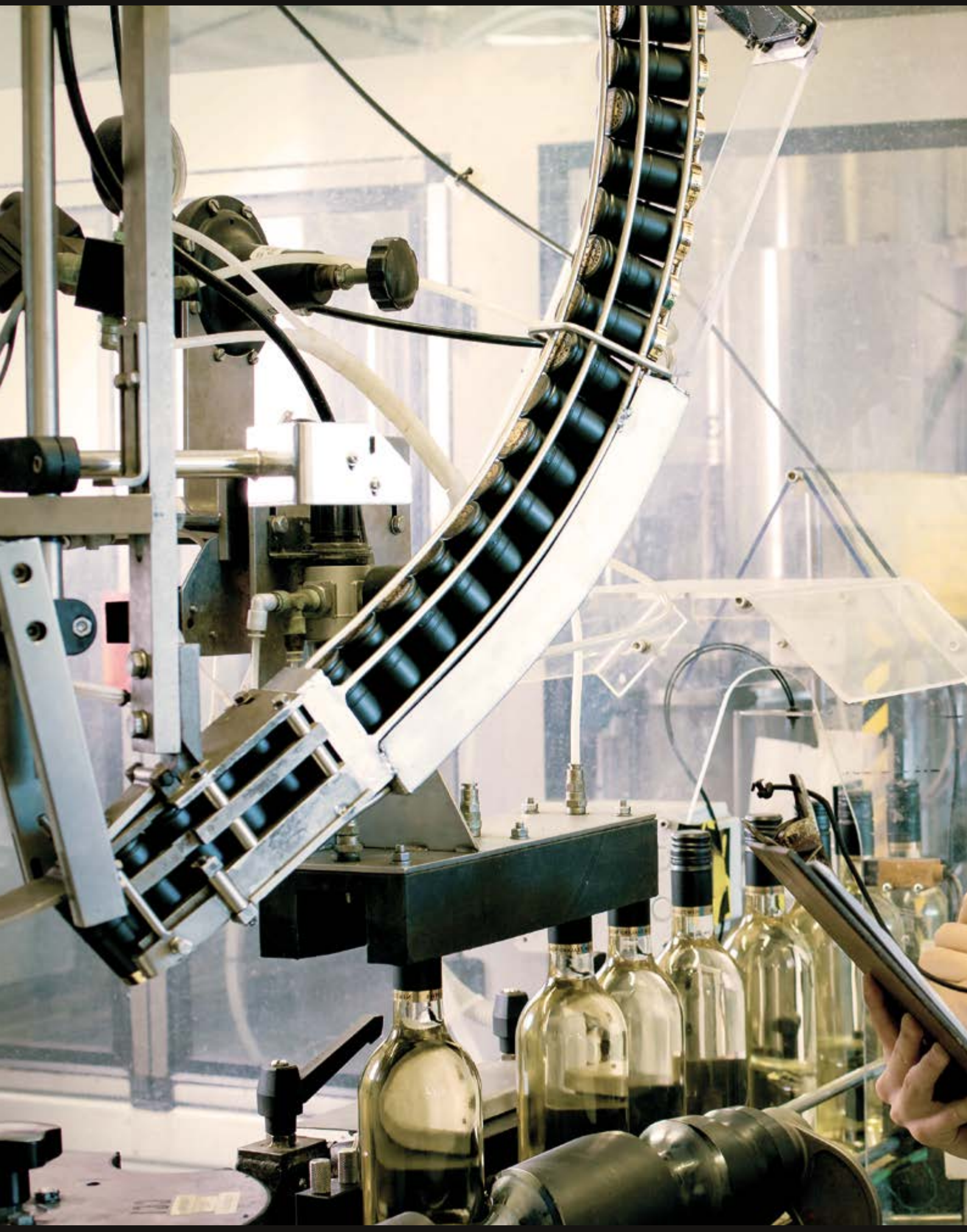
The paper finds that VC investment in online advertising companies decreased significantly in the EU relative to the United States after passage of the EU e-Privacy Directive. The EU e-Privacy Directive has led to an incremental decrease in investment in EU-based online advertising companies of approximately €249 million over eight-and-a-half years from passage through the end of 2010. When paired with the findings of the enhanced effects of VC investment relative to corporate investment, the author says this may be the equivalent of approximately $750 million to $1 billion in traditional R&D investment.

Wittes (2011) explores the possibility that the advance of technology and the proliferation of personal data in the hands of third parties has resulted in a conceptually outmoded debate, where the prevailing concept of privacy does not usefully guide the public policy questions now faced. Privacy covers such a wide multitude of potential harms that we may be better off with a more focused privacy concept, with other harms dealt with in other ways.

Wittes argues that a delicate balance needs to be struck in privacy policy. In reality people regularly and happily trade confidentiality and anonymity for convenience. They positively expect credit card companies to keep an eye on their transactions to protect them against fraud, and they do not mind credit reporting when the details the agencies report are favourable, accurate, and enable them to obtain credit.

OECD (2014) is a wide-ranging paper summarising the main interim findings of an OECD project looking at the role of data and analytics in driving economic growth through its impact on countries' accumulation of knowledge-based capital. It defines DDI, talks about which industries stand to gain the most, and outlines policy challenges on the supply side (the provision of data and analytics) and the demand side (making use of DDI), as well as broader societal challenges. The authors claim that the economics of data favour market concentration and dominance by a few firms. Lerner (2014) presents a contrary view.

EC (2014a) sets out some steps to accelerate the transition towards the "data-driven economy of the future". It notes that Europe has been slow to embrace the data revolution compared to the USA and lacks comparable capabilities. It suggests a wide range of actions including to build data infrastructure, to make more data available openly, to increase skills and cooperation amongst those doing data analytics, and to study policy issues related to privacy, data-mining, security and cross-border transfer. An accompanying paper (EC, 2014b) updates on efforts to speed the adoption and effectiveness of cloud computing in Europe.



6. Appendix – Technical matters

Types of data

This appendix explains our methodology in more detail, and outlines the various elements of our study. It begins with a short discussion on defining types of data.

In general terms, data is structured information, where the structure is imposed through data definitions that specify what is being measured and qualitative characteristics of the measurements. We discuss below some types of data, which were illustrated conceptually in Figure 2.

There is much discussion about so-called "big data", a term that has various definitions but by which we generally refer to the processing of large quantities of often unstructured data to find insights for action. We do not find this a particularly helpful concept in our context, and its conceptual uncertainty obfuscates more than it elucidates, so we discuss it no further here.

Certainly datasets do not need to be large to be useful, a fact that is particularly relevant given the size of many of the datasets available to New Zealand organisations. Depending on the context, a few dozen data points may be sufficient to reveal general trends and patterns, although larger datasets may be needed for more sophisticated analysis. The real issue is not the size of the dataset, but smart innovative use of the data that is available, applied to a particular problem.

Metadata is data that describes some other data. This may include information about the structure of the data, such as descriptions of the particular variables that are recorded and the possible values that these may take. Metadata may also include ancillary data that provides additional information about the records in a primary dataset. For example, in a library content of books are the primary data, and the library's catalogue database provides metadata about the books such as the authors' names, publishers, year of publication, subject matter, etc.

Open data is data that has been made publicly available for use in an accessible format, along with sufficient metadata to understand and analyse the data, and the legal freedom to use it for most or any purpose without asking for further permission. Government agencies are increasingly realising the value of making their data open and accessible to increase the transparency of government decision-making, and ensure the widest possible distribution of useful data that others can use and build on.

Internal enterprise data is information collected and generated by an organisation as part of (or as a by-product of) its own operations. Examples include customer transaction information, employment records, and operational cost information. Some organisations are choosing to make some of their data openly available but firms may feel there are commercial sensitivity or privacy issues to manage. Some argue that there is substantial societal value in making this type of data more open, because it has value for other purposes that do not compete with the firm itself. For example, aggregated cellphone calling data is useful in responding to disasters and other emergencies (Bengtsson 2011), or aggregated data from accountancy services can give insights into the operation of the economy.

Personal data is data that identifies an individual, for example as generated from transactions or other interactions with organisations.[13] Personal information (or data) is defined in and special rules that apply to personal information in New Zealand are created by the Privacy Act. If personal information is published, steps are typically

[13] The term "personal information" is used in New Zealand's Privacy Act; here we have called it "personal data" for consistency with the other types of data discussed. In some other jurisdictions, "personal information" is known as "personally identifiable information".

taken to protect the privacy of individuals within the dataset, for example by removing any identifying information (ie anonymising the dataset) and/or by aggregating the data of multiple individuals.

By definition, data that is analysed on a computer must be **machine-readable**. However, computer analysis is not restricted to quantitative information. Other types of digital data such as photos, video, and audio can be analysed using appropriate algorithms. This may include producing quantitative data as a first step in the analysis, for example algorithms are available that can count customers and analyse their movements through a retail store using CCTV video data as the input.

Focus on seven sectors of the economy

We examine the impacts of DDI on seven broad sectors of the New Zealand economy. These sectors are based on aggregation of similar industries in the standard ANZSIC 1-digit classification, shown in Table 3. The seven sectors combined represent 95% of gross value added (GVA) in New Zealand (Figure 14).[14]

TABLE 3

ECONOMIC SECTORS USED IN OUR ANALYSIS, AND CORRESPONDENCE WITH ANZSIC INDUSTRIES

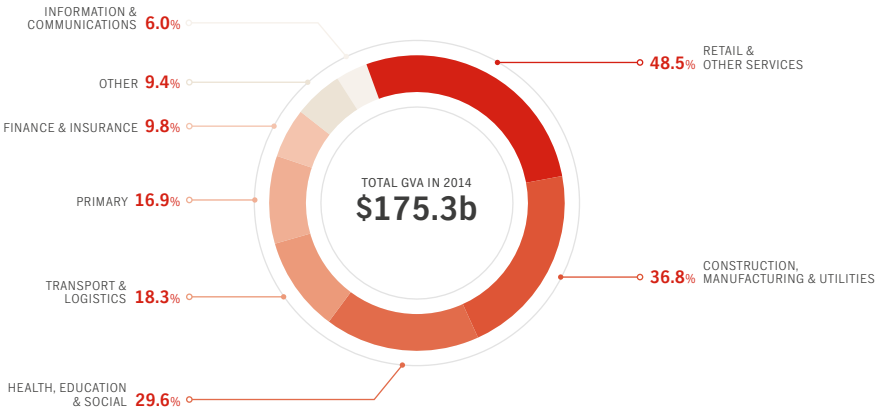
SECTOR FOR ANALYSIS	ANZSIC INDUSTRIES
Primary	Agriculture, forestry & fishing Mining
Construction, manufacturing & utilities	Construction Manufacturing Electricity, gas, water & waste services
Transport & logistics	Wholesale trade Transport, postal & warehousing
Retail & other services	Retail trade Accommodation & food services Rental, hiring & real estate services Professional, scientific & technical services Administrative & support services Arts & recreation services Other services
Information & communications	Information media & telecommunications
Finance & insurance	Financial & insurance services
Health, education & social	Public administration & safety Education & training Health care & social assistance

[14] GVA is closely related to GDP; the difference relates to the effects of government taxes and subsidies on goods and services. In order to simplify our analysis we focus on GVA effects, which are expected to be highly correlated with GDP impacts. The seven sectors in our analysis account for 94% of New Zealand's GDP.

The size of firms (measured by their number of employees) varies considerably across sectors (Figure 15). The primary sector is dominated by small and medium enterprises; around half of employees work for a firm with fewer than 100 employees. However in the information & communications, finance & insurance, and health, education & social sectors, large organisations dominate.

FIGURE 14

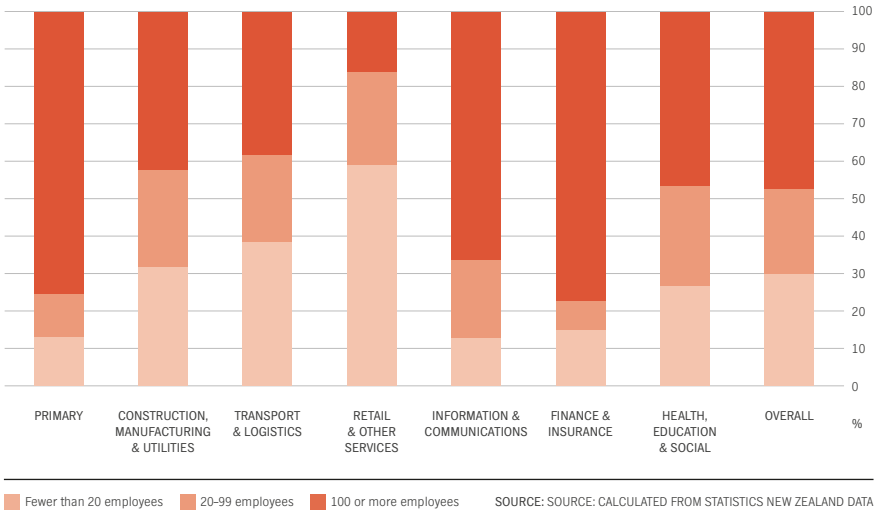
GROSS VALUE ADDED FOR THE SECTORS USED IN OUR ECONOMIC ANALYSIS



SOURCE: ANALYSIS OF STATISTICS NEW ZEALAND DATA

FIGURE 15

EMPLOYMENT BY SECTOR AND SIZE OF FIRM IN 2014



SOURCE: SOURCE: CALCULATED FROM STATISTICS NEW ZEALAND DATA

Overview of our methodology

Our estimation of the impacts of DDI on the New Zealand economy is based on the methodology used by AnalysysMason (2014a, 2014b) to estimate the value of DDI in Singapore and Japan. We have made adjustments to the methodology to reflect the structure of the New Zealand economy, and the information available about DDI in New Zealand.

We estimate the impacts of DDI on GVA in each of the seven sectors through two main channels:

- **Cost reductions** through operational and other efficiencies created by data and analytics, which enable organisations to increase productivity, and release scarce resources that can be put to alternative uses in the economy; and
- **Revenue increases**, for example due to more efficient marketing and pricing, and the development of new goods and services based on data and analysis.

Figure 16 illustrates the main components of our model for estimating the impact of DDI on consumers and producers in a single sector. The circles in the figure reflect the three key inputs for each sector – the DDI adoption rate, and the proportionate impacts on revenues and costs of those who adopt. Consumer and producer impacts for each sector are calculated and the results were added across sectors to generate an overall value of DDI in New Zealand.

The same methodology was replicated, with different inputs, for each of the seven sectors, for 2014 and 2020. The difference between the 2014 and 2020 valuation is an assumption of higher uptake of DDI in 2020.

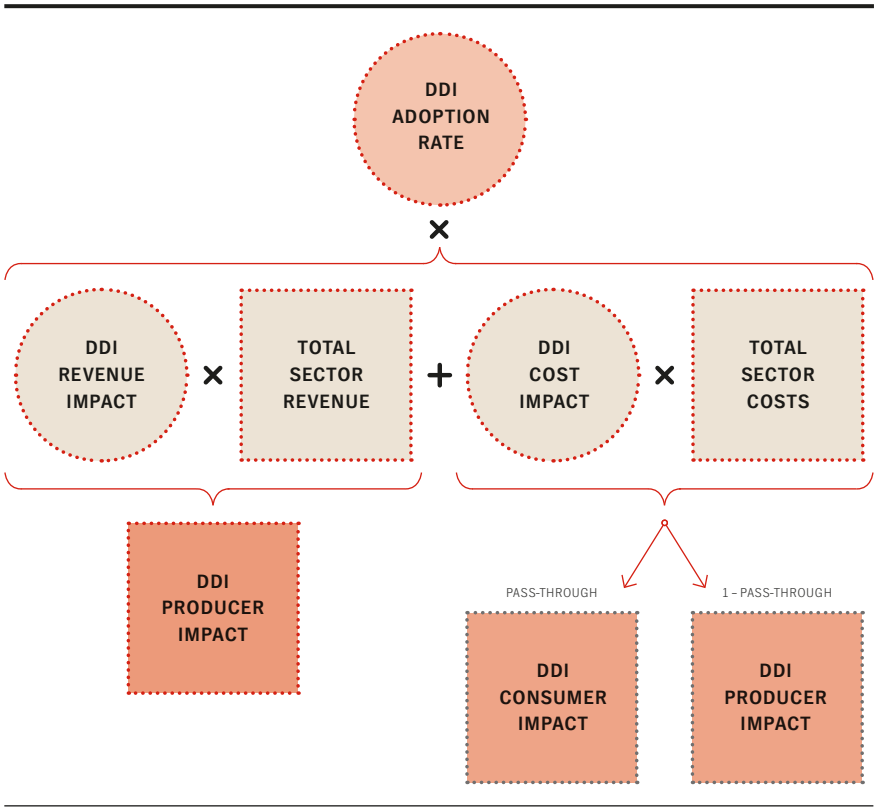
Information sources

Our analysis was informed by the following:

- Reviews of international academic literature and studies of the value of DDI in other countries (summarised in this report).
- A series of interviews with a range of businesses and government organisations in New Zealand to determine how they are using data and analytics and how this is affecting their performance. We also developed a smaller set of detailed case studies.
- Data on firm performance and indicators of innovation rates by industry in New Zealand, including Statistics New Zealand's Annual Enterprise Survey and the Business Operations Survey.
- Our own experience and judgement regarding the structure of the New Zealand economy, and the ways that DDI may affect organisations in different sectors.

FIGURE 16

ILLUSTRATION OF OUR METHODOLOGY FOR ESTIMATING THE IMPACTS OF DDI ON GVA IN A SECTOR



SOURCE: SAPERE AND COVEC; ADAPTED FROM ANALYSISMASON (2014A,B)

Calculating the impacts

As illustrated in Figure 17, the overall impact of DDI in a sector is estimated as:

DDI impact = Adoption rate × [Cost impact × Total costs + Revenue impact × Total revenue × Profit margin]

Where:

- **Adoption rate** reflects the overall average uptake of DDI in the sector
- **Cost impact** reflects the estimated percentage reduction in total costs achievable by organisations as a result of DDI
- **Revenue impact** reflects the estimated percentage increase in total revenues achievable by organisations as a result of DDI
- **Profit margin** reflects the increase in revenues that is not offset by increased costs

The analysis was repeated across sectors (with different inputs) and then aggregated to estimate the total economic value of DDI.

Information about how we calibrated each of the model parameters for each sector is given below along with our other data sources. We also estimate the split of the impact of DDI between consumers and producers in each sector. This was modelled by assuming that some proportion of DDI-related cost savings are passed through to consumers, and the remainder is retained by firms.[15]

DDI adoption rates

Adoption of DDI is expected to vary by sector depending on a number of characteristics such as the nature of economic activity in the sector, the types of relevant data that are available, and whether the sector has historically been quick to adopt new technologies. In estimating adoption rates by sector for New Zealand we have been guided by international literature (especially Analysys Mason (2014a, 2014b), Schroeck et al (2012), Lavalle et al (2011), and Talend (2012), combined with insights from our interviews and our previous experience.

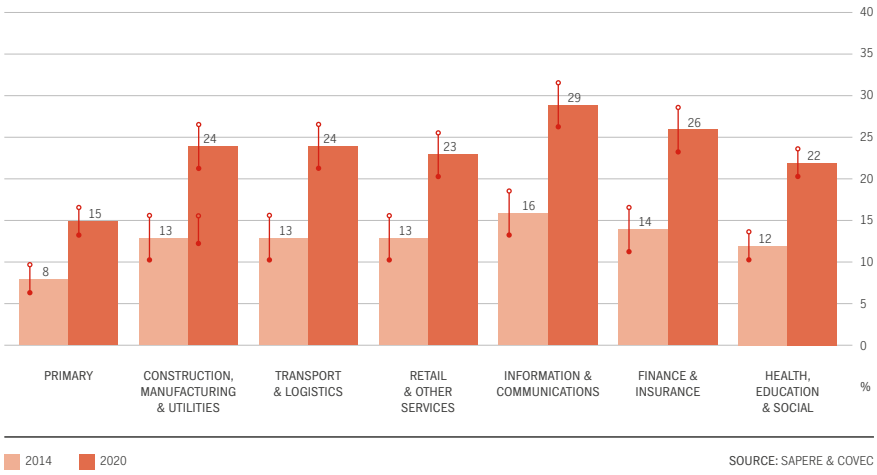
These overall averages were translated into adoption rates for each industry (at the 1-digit ANZSIC level) using information from Statistics New Zealand's Business Operations Survey (BOS). The BOS reports the overall innovation rate for each industries. We assume that industries with higher innovation rates are more likely to adopt DDI, and we used the BOS results to generate a DDI adoption rate for each industry that reflects the cross-industry variation in innovation while the weighted average across all industries (using GVA for each industry as the weight) equals the total adoption assumption.

We also made a small number of manual adjustments to the resulting adoption rates by industry, to reflect our findings from interviews with organisations in different sectors. The resulting assumed DDI adoption rates for each sector are shown in Figure 17 along with the possible range of adoption levels.

[15] It is also expected that the introduction of new products and services based on DDI will confer benefits on consumers that are not captured as revenue by firms. These consumer surplus benefits are very difficult to quantify and we have not attempted to do so. However it is likely that our results under-estimate the impacts on consumer surplus by excluding such benefits.

FIGURE 17

ASSUMED ADOPTION RATES OF DDI BY SECTOR



SOURCE: SAPERE & COVEC

Revenue and cost impacts

To estimate the effect of DDI on gross revenues and total costs of firms that adopt DDI we assumed an overall average impact of DDI on adopters of:[16]

- An increase in gross revenues of between 5% and 10%, with between 2% and 30% of this being retained as increased producer surplus and the remainder being offset by costs of production, depending on the industry profit margin.
- A reduction in total costs of between 1% and 2%, with between 30% and 80% of this passed through to consumers in the form of lower prices and the remainder retained as increased producer surplus, depending on the assumed rate of pass through for each industry, which in turn depends on the competitive intensity in each industry.

As with DDI adoption, we used the results of the BOS to disaggregate these overall averages across industries. Specifically, the BOS reports the proportion of businesses in each industry that undertook innovation in four categories: "goods or services", "operational processes", "organisational or managerial processes", and "marketing methods". We classified "goods or services" and "marketing methods" as revenue-related innovation and the remaining categories as cost-related innovation, and assumed that the innovation rates reflect the overall opportunity for revenue and cost innovation respectively in each industry.

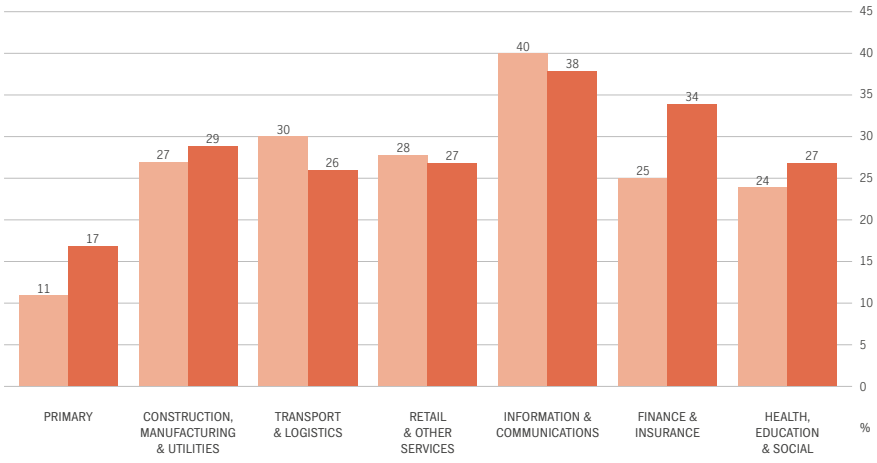
The GVA-weighted averages of these innovation rates by sector are shown in Figure 18. These proportions were then used together with data on GVA per industry to calculate revenue and cost impacts of DDI such that the GVA-weighted average across industries equalled the overall averages above. We also made a small number of manual adjustments

[16] These assumptions are primarily based on AnalysysMason (2014a, b), findings from our interviews of New Zealand businesses, and our own experience and judgment.

to the results by sector to reflect information gathered in our interviews of New Zealand organisations. Our resulting assumptions about the effects of DDI on revenues and costs of adopters by sector are shown in Figure 19, along with the likely ranges of these effects.

FIGURE 18

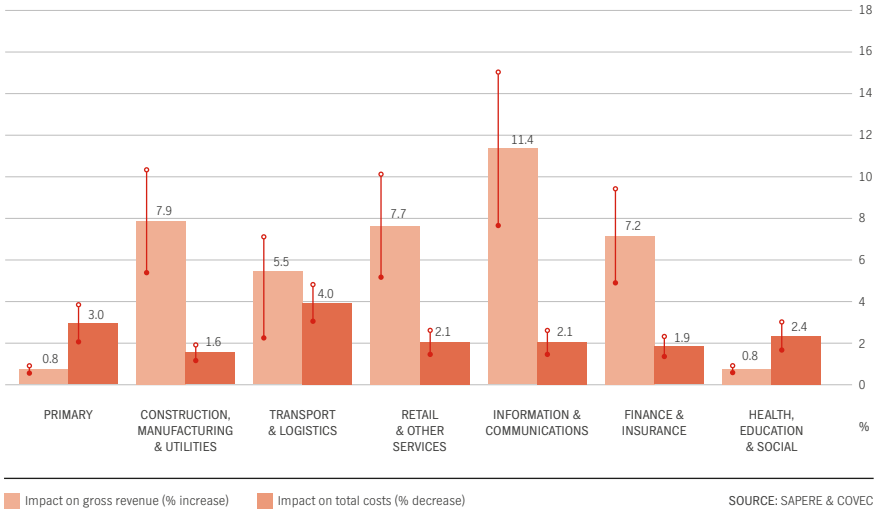
RATES OF "REVENUE" AND "COST" INNOVATION BY SECTOR (GVA-WEIGHTED AVERAGES)



SOURCE: SAPERE AND COVEC ANALYSIS OF STATISTICS NEW ZEALAND DATA

FIGURE 19

ESTIMATED IMPACTS OF DDI ON REVENUES AND COSTS OF ADOPTERS IN NEW ZEALAND



SOURCE: SAPERE & COVEC

Industry revenues, costs, and margins

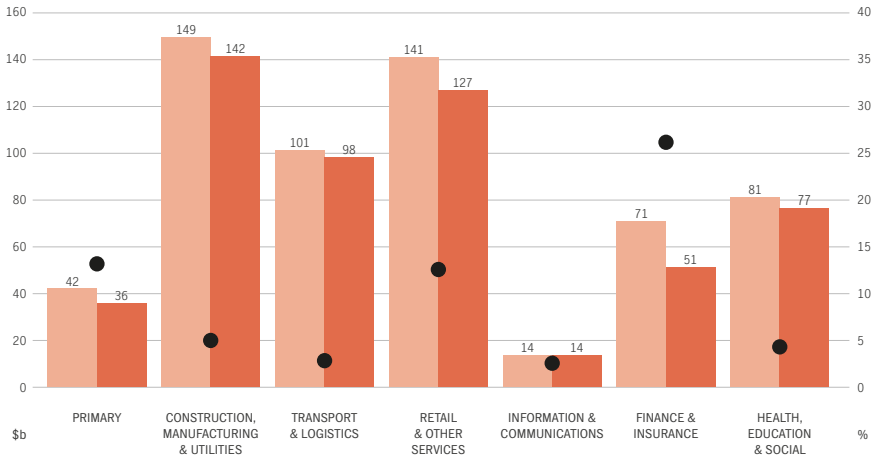
Data on total revenues and costs by industry was obtained from Statistics New Zealand's Annual Enterprise Survey for 2011 to 2013 (Figure 20). Average revenues and costs for each industry over the three years were calculated and used as the revenue and cost estimates in our modelling for 2014. An average over three years was used so that temporary fluctuations in industry revenues and costs, for example due to short-term changes in international commodity prices, are less likely to affect our analysis.

Revenues and costs in 2020 were estimated by increasing revenues at an average real annual rate of 1.5%, while also increasing costs so that a constant profit margin was obtained.

Profit margins were calculated as the ratio of the difference between revenues and costs to total revenues (averaged from 2011 to 2013). As can be seen in Figure 20, profit margins vary significantly across sectors, while total revenues and costs are related to the overall size of each sector.

FIGURE 20

REVENUES, COSTS, AND AVERAGE MARGIN BY SECTOR (ANNUAL AVERAGE FOR 2011-13)



SOURCE: CALCULATED FROM STATISTICS NEW ZEALAND ANNUAL ENTERPRISE SURVEY DATA

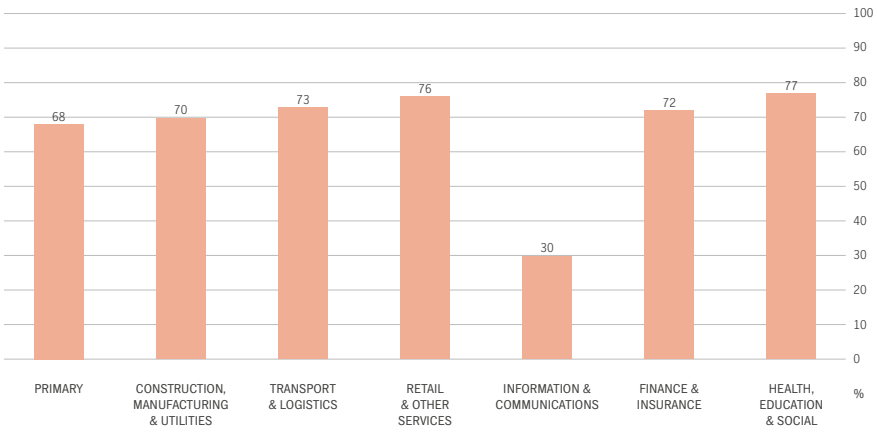
Pass-through of cost reductions

Economic theory predicts that firms in competition will pass through some but probably not all cost reductions to their customers in the form of lower prices. The rate of pass-through is expected to be higher in industries where competitive intensity is greater. An indirect indicator for an industry is provided by the Hirschman Herfindahl Index (HHI).[17] HHIs for New Zealand industries have been estimated by Stevens (2011) at the 1-digit ANZSIC level.

We assumed that pass-through ranges from 80% for the most competitive industry (indicated by lowest HHI) to 30% for the least competitive industry (highest HHI). Rates of pass-through for industries with profit margins between these two extremes were determined on a linear scale between 30% and 80%. The resulting pass-through assumptions by sector are shown in Figure 21.

FIGURE 21

ASSUMED RATE OF PASS-THROUGH OF DDI COST REDUCTIONS BY SECTOR



SOURCE: SAPERE & COVEC

[17] The HHI is calculated as the sum of squared market shares. A monopoly has an HHI of 10,000, and lower values generally reflect more intense competition.

Interviews and case studies

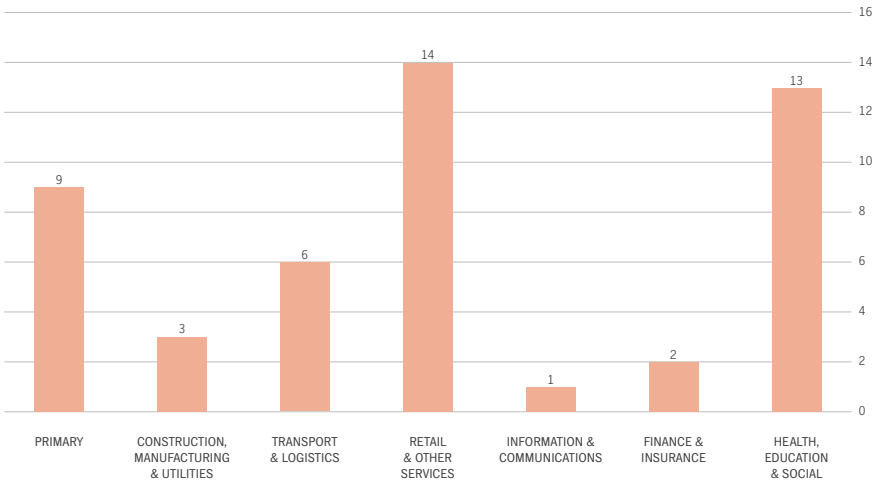
As well as the quantitative work, we conducted 48 structured interviews with New Zealand organisations across a range of sectors, including private businesses and government departments (Figure 22). Some of these were providers of data and analytics services to other organisations, but most were users of data and analytics services that were either supplied internally or purchased from other organisations.

While not a formal survey and thus not suitable for statistical analysis or hypothesis testing, our interviews were structured around a set of questions and topics designed to gather information about how organisations are using data and how this is impacting on their performance. We also asked about the overall level of uptake of DDI within each sector and in comparison with leading international organisations in the same sector.

In addition to the interviews, we generated six more detailed case studies of DDI in different types of organisations. These are included with this report and were also used to inform our analysis, in particular as an information base to adapt international estimates of the impacts of DDI to local conditions in New Zealand.

FIGURE 22

SUMMARY OF THE NUMBER OF INTERVIEWS CONDUCTED BY SECTOR



SOURCE: SAPERE AND COVEC

Limitations of our analysis

The scope and accuracy of our analysis is limited by available information. In addition to not measuring spill-overs as mentioned above, other important limitations are:

- We ignore the consumer surplus impacts from new and improved goods and services as a result of DDI, to the extent that these consumer benefits are not captured as increased revenues by firms.[18] Our results therefore likely under-estimate the positive impacts of DDI on consumers.
- There is uncertainty about the key parameters in our analysis. For this reason we present a range of results that is intended to capture the possible range of impacts of DDI, given these uncertainties.
- Somewhat ironically, there is limited data available on firms' use of data in decision-making. Our assumptions on the proportion of firms that today use DDI are based on international evidence and on what we learned from interviews. They are uncertain therefore.
- We have sought to estimate the value and effects of DDI that are associated with recent technological advancements, and that we believe can be quantified with reasonable certainty.

[18] Such impacts depend on the value that consumers place on new goods and services, which by nature is very difficult to quantify, so we have not attempted to do so.

7. Appendix – Case studies

LICENCED TO MILK

HOW FONTERRA'S GLOBAL INTELLIGENCE OPERATION DELIVERS REAL VALUE FOR NEW ZEALAND

On the face of it, dairy is a simple business. Grass goes in, milk comes out, and someone (mainly, these days, someone overseas) drinks it.

For Terry Childs, Fonterra's General Manager of Global Dairy Intelligence, that couldn't be further from the truth.

As one of the world's major food producers, Fonterra is constantly making decisions about production and distribution that don't just affect the company – they have the potential to affect the entire New Zealand economy.

And since mid 2009 those decisions have been largely based on the information and insights provided by Terry and his team of seven.

Initially the focus was on global supply and demand of dairy products, to try to help the company forecast price. Even gathering that data, though, has its challenges. "Markets like Europe and the United States are relatively straightforward thanks to reports available from the USDA and Eurostat," says Terry. "But China is a different matter, and needed a whole lot more work."

For Terry's researchers and data specialists that meant firstly identifying the accessible sites that looked like they had the right information, then translating them as they were mostly in Mandarin. Once the data was found it was cross-checked with other sources to validate it, then automatic processes were developed to regularly "scrape" the sites, extract data from PDF reports in many cases, translate it, then present it in a standard format.

The Marketing Analytic System (MAS) the team has developed now forms part of everyday management decisions at Fonterra and the data collected has expanded well beyond supply and demand to include over 1200 sources worldwide. This now includes on-farm economics, trade data, tariffs and agreements, global production and consumption and retail numbers, as well as softer measures around brand health and customer sentiment.

Social media has been especially useful in this last area and Terry says it's a handy reality check for management. "When something does go wrong it's easy to focus on the odd negative comment in the media ... what we can do is show the executive team 400 tweets and say here are some negative comments, here are some positive ones, and here are some neutral ones. The result is a more objective view of things."

While the Marketing Analytic System is fiendishly complicated, at its heart its role is straightforward: automate the collection of data, then turn that into actionable insights. It's this second part that is a current focus for the team, says Terry:

"Spreadsheets have been around forever but they're not necessarily the best way to show insights. So we're working on using simple, interactive visual techniques to turn all those facts and figures into insights people can use to help make sound decisions."

A good example of data driven insights in action was a recent development in the Chinese dairy market. Despite the country's GDP growth declining, milk prices were trending upwards – not what other indicators would have predicted. When Terry's team dug deeper they identified a supply issue, and predicted that Chinese demand for whole milk powder would increase.

This insight helped Fonterra leadership adjust the company's product mix – including the balance between whole milk powder, skim milk powder and butter – to meet the forecast Chinese demand and take advantage of the sales opportunity.

Sometimes research tasks can take the team in unexpected directions. "If we want to know if competitor company X is planning to make more milk powder in the future, then real estate stories from news websites might tell us that they've bought the land they'll need to build a plant on ... so that all goes into the mix.

At the moment, the insights are used by around 100 people across the business but Terry says in the near future that will increase to around 1000, as the power of data-driven insights becomes more widely accepted and integrated into the forecasting and decision-making processes at New Zealand's biggest exporter.

SEEING THE CARBON FOR THE TREES

HOW FOUNDATIONFOOTPRINT IS SIMPLIFYING SUSTAINABILITY

Sustainability reporting isn't a new idea. Businesses and government organisations have been gathering data about their energy efficiency and environmental impact for decades.

And there's never been a shortage of data. Energy use, vehicle mileage, waste tonnages, paper consumption, air conditioning temperature ... any decent sized organisation can easily find enough figures every month to fill any number of spreadsheets (and as buildings, vehicles and business equipment becomes smarter, this number keeps growing).

And that, says FoundationFootprint founder Chris Lindley, is exactly the problem his business set out to solve when he set it up in 2007.

"The old way of doing this was just terrible," he says. "In a large business that wanted to report on, say, 12 metrics, that would often mean 12 spreadsheets being emailed around from one department to the next, sitting in in-boxes for days or weeks, versions getting confused ... then when the auditors came it was a recipe for disaster."

Worst of all, he says, the resulting reports were sometimes so inaccurate that they had little business value beyond something for the marketing department to talk about in their advertising.

FoundationFootprint set out to change all that, and create a simple, powerful system that combined data from a wide range of sources and formatted them in a standard way to create reports tailored to users in different parts of a business.

Essentially, FoundationFootprint does this by connecting with a business's internal and external data suppliers and gathering information on a regular cycle, either automatically (where APIs and other interfaces permit) or manually, via email or secure upload.

The platform then converts each set of data it receives – say, a power bill for a business location, or a monthly air travel report for a division or company – into a standard form to base its reports on.

The outputs depend on what a company wants to measure, but the key measures are the same in every case: dollars spent (financial efficiency) and carbon produced (environmental impact).

Different parts of an organisation will see different dashboards. While an exec-level dashboard would show overall efficiency and impact, different divisions can receive a more tailored view.

"So if air travel is a focus, we can show a company not just how much they're spending and what the environmental impact of that is, but also drill down and see how efficient the spend is ... so if one department or even one individual typically books late and pays through the nose, we can identify that so that cost can be managed."

At an organisational level, some companies are seeing carbon footprint reductions of up to 10%, through accurate measurement, setting targets, then tracking progress towards them – all via the FoundationFootprint platform.

As well as comparing their performance to targets and their own historical measures, a growing number of companies are also benchmarking themselves against similarly sized organisations, or others in similar sectors. Of FoundationFootprint's 180 clients (80% of which are outside New Zealand), 100 are choosing to be anonymously benchmarked.

For Chris, the business combines two life-long passions: computing and the environment. "I've been involved with computers and programming since I was a kid, and grew up watching wildlife documentaries on the BBC. And they always seemed to end with a message about how human activity is threatening species or destroying habitats. FoundationFootprint means I can use cloud computing and data analysis to give businesses the insights they need to lessen their impact, while at the same time becoming more efficient operators."

The scale and variety of the data FoundationFootprint deals with is huge. At client BNZ, for example, information is captured from over 2700 "locations." These can be anything from a bank branch, to a floor in an office building, to a fleet vehicle. With an average of four data sets recorded for each location, that's more than 10,000 data points going into each report (6-monthly, in BNZ's case). Under the old "spreadsheet shuffling" model, creating the reports would have required an immense time commitment from the bank and led to a huge lag between data collection and report delivery. Now, once all the data has been captured, the reports and dashboards are delivered to users instantaneously.

While business efficiency is a key focus for FoundationFootprint's 180+ clients, data efficiency is a big focus for Chris. "As part of the process we help clients identify which of their internal and external data suppliers are providing high quality information in a timely manner – which in turn helps everyone improve their business processes."

NO ACCIDENT

HOW SMART USE OF DATA IS HELPING ACC DELIVER RESULTS AND WIN TRUST

By any standards, the Accident Compensation Corporation is a huge organisation with a lot on its plate. Every year, 1.8 million claims are lodged for everything from rolling an ankle at rugby practice to serious workplace accidents.

And its ACC's job to make sure every one of those claims is processed promptly, and that all cases – especially the most urgent – receive the attention they deserve.

For ACC Manager Strategic Intelligence and Analytics, Andrew Rae, smart use of data is playing an increasing part in delivering on ACC's mission.

"We've got over 30 years' worth of historical data to work with... that's over 40 million claim forms, with all the information they hold."

It's by taking a scientific approach to that data that ACC is working to deliver better outcomes for both its claimants and the New Zealand businesses and taxpayers that fund it. The client service optimisation project is focussed on prioritising the 1.8 million claims the business receives each year, with an end goal of producing the right rehabilitation outcome so people are well enough to return to work, or build independence and capability, after an injury.

Previously, cases were escalated and given extra attention based on a "one size fits all" set of rules, including how long someone had been off work with an injury. That could mean that a relatively simple injury that didn't heal properly for some reason (say a broken leg) might not receive the specialist support it needed as quickly as it could – meaning that something that could be put right in weeks could potentially stretch into months.

Under the new system, rolling out from 2014, information relating to the client and their claim receives the appropriate cover decision and gets to the right case owner quickly. This means the right entitlements are provided early on. This, along with an appropriate rehabilitation plan, sets a target for producing an effective client outcome.

While the client service optimisation project is aimed at helping individuals, ACC is also using data at an organisational level, by identifying injury patterns and targeting prevention strategies.

"Our historical data can give us a pretty clear view of what's causing accidents and where they're happening," says Andrew. "So if we know there's an activity or business sector that's costing us $26 million a year, you could make a strong case for spending money on prevention in that area." By targeting prevention strategies based on data analysis, ACC hopes to deliver better value for all its stakeholders. "So if we invest wisely in prevention in an area like that and reduce the accident rate by 5 percent, that's a very good outcome and a huge saving," says Andrew.

While the organisation from the CEO down supports the new approach, it's had its challenges, including collecting data in a standard usable way from claim forms that allow people to describe the way they were injured in their own words. Section has been added.

One important benefit of this new approach that can't be measured in days or dollars – but is at the heart of ACC's new data-driven approach – is improved trust. Trust and confidence are key to running ACC ''. If data can help lead to a better rehabilitation outcome for our customer, then we believe that will lead to greater trust and better relationships for everyone''.

And if this seems like a bit of a soft measure for a data analyst, Andrew and his team are onto it. ''We're working on developing client trust measures right now, and they'll be a key metric for the business in the coming years.''

FROM INTUITION TO INSIGHT

HOW SMART DATA IS TRANSFORMING FARMING

"There's a well-known saying in farming circles," says Farmax Senior Technical Specialist Steven Howarth: "The difference between a good farmer and an average one is two weeks."

What Steve's getting at, is that the timing of the decisions a farmer makes can have a huge impact on the farm's profitability. When to sell stock. When to breed, or buy stock in. How long into the season to keep milking. When to cut surplus grass for silage, or keep grazing it.

Do it at the right time and yields will be optimised, you'll buy less feed in, you'll get the best prices for the stock you sell and pay less for the stock you buy – all of which make the difference between a successful farm and an also-ran.

For most of New Zealand's farming history, those decisions were based on intuition and experience – which is fine if you happen to be intuitive or experienced. Looking forward, farmers are increasingly finding themselves in uncharted territory. With increasing market volatility, larger more complex farming operations and increasing pressures to be environmentally compliant, the need for tools to support factual based decision making is increasing.

Farmax, Steve says, doesn't replace traditional farm decision-making; it just gives the farmer (and often the farm consultant) a powerful and accurate tool to model the likely effects of different plans before they are implemented.

"Farmax gives the farmer a computer model of their business," says Steve. "That includes farm area, grass cover, stock numbers and stock performance. It then predicts what the likely feed budget and profit is going to be going forward."

But where Farmax really comes into its own is when its modelling power is combined with the farmer's own insights and decision making.

"So you can look at what would happen if you milked for 15 days longer. Or if you left stock to graze instead of making silage. Or what the impact would be on a sheep and beef farm of altering the ratio between the two."

For any option the farmer puts in, Farmax will forecast the costs required and the profit or loss the change will lead to – giving the farmer a solid basis for making what are often costly changes.

Farmax can be especially helpful in tough farming conditions too. Prolonged droughts can mean farmers have to make difficult decisions around when to sell stock. With Farmax,

the decisions become more objective – if a specific amount of rain doesn't fall by a certain date, the stock is sold ... reducing the chance of selling poor-condition stock at equally poor prices.

As New Zealand farms increasingly shift from family-owned properties to company-owned and managed operations, Farmax makes it easy to gather and retain knowledge about a farm too. "Previously, if a manager left, his or her knowledge about the farm, what works and what doesn't, left with them. With Farmax, a new manager can walk in, look at the plan, see how it's working and pick things up with no dip in productivity."

While Farmax started out as a modelling tool for sheep, beef and deer farmers, these days it also offers versions for dairy farms too. The platform has over 1000 users, with an estimated 5,000 farms being modelled, and it's especially popular with the influential farming consultant sector, with around 80% of sheep and beef farming consultants using it, and a growing number of dairy consultants.

And it works. While Steve admits it's hard to isolate the actual financial benefit Farmax delivers (after all, the farmers using it may well be better farmers in the first place), the difference between farms with Farmax and without are staggering. "For sheep and beef farms using Farmax, looking over eight seasons at their average gross margin per hectare, our users' farms yielded $600/ha, compared with non-Farmax places at just $350."

Of course a number of other factors are at play, but across the entire user base, the better farming decisions that Farmax has helped them make have contributed an estimated $450 million to the collective bottom line.

It's accurate too, with dairy users reporting Farmax's predictions coming within 5-10kg of milk solids per cow over a season (against an average annual production per cow of 350kg).

While the system gives farmers complete ownership and control of their data, opt-in benchmarking trials have shown even greater productivity lifts are possible when farmers can benchmark their own production against others and the industry. Additional benchmarking services may be offered to all users in the future.

And while Farmax will never replace the need for skilled farmers making decisions based on their experience, it is rapidly becoming a vital tool for professional farmers. As one Northland user said after a couple of years' using the system, "Farmax is just as important to me as my tractor." And that's pretty high praise.

UNLOCKING HEALTH

HOW SMART DATA IS PUTTING PATIENTS FIRST IN CANTERBURY

Seven years ago, Carolyn Gullery says, Christchurch Hospital was not a great place to be in midwinter. "Every year, sometimes more than once, we'd hit what we call gridlock. The wards were full, which meant people turning up to the emergency department had nowhere to go if we needed to admit them. We had to cancel elective surgery too, because if people needed post-operative care we just didn't have the beds. It wasn't good."

As General Manager of Planning and Funding, Carolyn realised big changes were needed. While simply building more wards and filling them with beds might offer a short term solution (if the money could somehow be found), it did nothing to address the alarming rise in both hospital admissions and people going into aged care facilities.

"So we started with the data," says Carolyn.

After the 2011 earthquakes, 106 hospital beds were lost so Canterbury needed to explore new ways to reduce reliance on hospital beds to look after a vulnerable population.

"We pretty quickly saw that a big proportion of our winter admissions were being driven by one condition: Chronic Obstructive Pulmonary Disease (COPD)."

Digging deeper, it turned out there were multiple opportunities to manage patients better and in their own homes and communities when safe. While processes were in place for GPs to manage milder cases without patients going to hospital, most of the admissions followed 111 calls – and ambulance protocols resulted in COPD patients being transported to hospital.

Once there, they were almost always admitted and treated, Carolyn says, "as if they were on a conveyer belt – as it was quite hard to get an unwell anxious patient back home even when they wouldn't benefit from a hospital admission."

At peak winter times, up to 60 beds could be occupied by COPD patients, many of whom could be happier and just as well cared-for in their own homes.

Based on this insight, the DHB decided to change the way it managed patients with COPD, by having ambulance staff decide whether a GP referral, home care by a nurse, or transport to a local emergency clinic was better for the patient. The results were dramatic and immediate, with a 30% reduction in COPD patients transported to hospital.

While the change in COPD management led to a significant improvement, it was just one example of a wider shift towards using data in a patient-centred way.

When the DHB began the process in 2007, many patient records were still paper-based, and those stored electronically were on separate systems that weren't accessible to clinicians in different parts of the system.

To solve this, they developed HealthOne – a system that uses and manages a single electronic patient record, wherever in the system they are being cared for. Built on New Zealand company Orion Health's Concerto platform, HealthOne integrates secondary care (hospital) records with information from GPs, pharmacies, laboratories and other providers.

The result for healthcare workers is that it's now far easier to get a complete picture of the person in their care. So if someone turns up at the emergency department, the team will know their pre-existing conditions, what medication they're taking and a whole lot more, without having to rely on the patient telling them, or waiting for GP records to arrive. As Carolyn describes it, "It's the difference between dealing with an endless stream of strangers, and actually knowing something useful about the patients and their histories."

For a patient, this means faster, more appropriate care. It also means less frustration from having to tell every new person they meet in the healthcare system exactly what's wrong with them.

So far the system holds records for 480,000 Cantabrians, as well as the region's hospitals. It is used by almost all of Canterbury's general practices and pharmacists. And it's working. As well as eliminating winter gridlock – and the subsequent effects on patient care, waiting times and staff frustration – the new approach has helped the Canterbury DHB reduce patient acute medical admission rates to 30% below the national average. This means better use of resources, more personalised patient care and better outcomes all round.

EFTPLUS

HOW SMART DATA IS TURNING TRANSACTIONS INTO INSIGHTS

New Zealanders love electronic payments. We were one of the first markets to widely adopt EFTPOS and these days each of us swipes, dips or taps our cards 190 times a year.

That all adds up to big business. Payment processing business Paymark (with 75% market share) reports that its 75,000 merchants process over 900 million transactions annually – worth over $48 billion.

For New Zealand business EFTPlus, though, the real value isn't in the dollars and cents; it's in the shopper behaviours contained in all that data and what you can do with it to offer a better service.

EFTPlus helps retailers succeed by operating a virtual loyalty programme – using a customer's EFTPOS and credit cards to track spending in order to reward their loyalty, and give retailers powerful insights into who's really buying their stuff.

The idea for the business came about when its founder Julian Cox became fed up with having to cart around a wallet full of coffee cards just to get a free cup every ten visits.

"With so many New Zealanders paying electronically, he knew there had to be a better way," says CEO Marcus Hoefliger. "So we started looking at the transaction data that merchants could already access to see what could be done."

Before long Paymark asked what the company was up to, and soon a relationship was formed that works for retailers, shoppers and EFTPlus.

While every merchant (often a bar, restaurant or some other business that relies on repeat visits) operates the white-labelled system its own way, the common elements are:

- Sign up: a customer sends a registration text, typically after seeing a point of sale poster or table talker in store. Most retailers then reward this with an immediate gift (such as a free drink).
- Profile completion: customers are then invited to complete their profile, including registering credit and debit cards. Completion rates at retailers such as specialty grocers are typically 85%, but lower in the hospitality sector.
- Insights: via an online dashboard, the merchant sees who's spending and when. (When a retailer's point of sale system is integrated with EFTPlus, they can also see exactly what's being bought.)
- Loyalty: based on transaction data, customers receive offers based on what they've spent or on special occasions (e.g. their birthday).
- Emarketing: as shopper data is accumulated, the merchant can use it to power more sophisticated customer relationship management (CRM) programmes and communications.

Smart retailers, Marcus says, are using the power in their shopper data to both provide a better shopping experience for their customers and add to their own bottom lines.

"The beauty of EFTPlus sitting on top of a long trail of Paymark transaction data is that when Joe Bloggs signs up to my loyalty programme I don't just see his behaviour from now on, I can see it for the last three years and tailor my offers accordingly. That's really powerful."

The other advantage is that it lets merchants see the impact of data-driven insights on individual customer behaviour. "By comparing a customer's behaviour at, say, a bar before and after they join a well-run programme," Marcus says, "we'll typically see revenue from that customer more than double."

That's a huge uplift in what can be a fiercely competitive sector. Interestingly, combining data from competing businesses can provide powerful insights too.

"Comparing your performance today to your own performance last week, or last year, is a useful measure but it isn't everything. What we can offer is a benchmark against other local businesses in the same sector – and that's really valuable business intelligence."

Of course, EFTPlus is a lot more than a data-driven version of the humble coffee card. By looking through the "what" (the data) to understand the "who" (the customer), Marcus says retailers can point their offers exactly where they're likely to be most effective.

"By looking at the transaction data it's easy to identify previous big spenders who for some reason have stopped visiting. Sending them a targeted message with an attractive offer can work wonders."

An upcoming feature that lets customers send feedback on their experiences through the platform is another plus, giving disgruntled customers the opportunity to help a retailer put things right, rather than just walk away unhappy and never return.

While EFTPlus mainly uses Paymark transaction data (sometimes combined with information a retailer's own POS system), the company is also looking at combining with other sources, including weather forecasting data. "The great thing about weather is that it's relatively predictable," says Marcus. "So if you're a waterfront bar and you can see that there's a great sunny weekend coming up, you can push an offer to the kind of high value customers whom you know from past behaviour like to enjoy a beer in the sun with their mates."

While this might seem a little too personalised for some, Marcus reports that most customers are easily reassured that their personal information is being well looked after. EFTPlus stores anonymous tokens rather than cardholder identities, and data is only provided to merchants in aggregate, not in an identifiable form.

If the system seems smart today, the future sounds even more impressive. EFTPlus is already working with ecommerce sites to apply the same principles to online shoppers as they do in the real world. And as payment platforms move from cards to phones the opportunities are almost limitless.

"If your card is just an app on your phone, that means your card has now got location data built into it," says Marcus. Combine that with technologies like (indoor positioning system) iBeacon and the future for shoppers, retailers and data-driven businesses like EFTPlus is looking pretty bright.

The numbers (since launch)

- 15.5 million transactions processed
- $575 billion transaction value
- 160,000 members (incl. Indonesia)
- 862,000 rewards provided by merchants

8. References

Abrams (2014), "A Unified Ethical Frame for Big Data Analysis".

Analysys Mason (2014a), "Data-driven innovation in Singapore", 28 January 2014

Analysys Mason (2014b), "Data-driven innovation in Japan", 31 October 2014

Bakker (2013), "Valuing the Census", Report for Statistics New Zealand

Bakhshi, Bravo-Biosca, Mateos-Garcia (2014), "Inside the Datavores: Estimating the Effect of Data and Online Analytics on Firm Performance", Nesta, March 2014

Barua, Mani and Mukerjee (2011), "Measuring the Business Impacts of Effective Data"

Bengtsson et al (2011), "Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti"

Brynjolfsson et al (2011), "Strength in Numbers: How Does Data-Driven Decision-making Affect Firm Performance"

Cavoukian and Castro (2014), "Big Data and Innovation, Setting the Record Straight: De-identification does work", IPC and ITIF, June 2014

Conference Board of Canada (2012), "Exploring the Iceberg: The Economic Impact of Privacy Policy, Laws, and Regulations on Commercial Activity"

Data Futures Forum (2014a), "New Zealand's Data Future: Full Discussion Paper"

Data Futures Forum (2014b), "Harnessing the Economic and Social Power of Data", Final report

Data Futures Forum (2014c), "Summary of engagement", A summary of stakeholder engagement during the project

de Montjoye et al (2013), "Unique in the Crowd: The privacy bounds of human mobility", Scientific reports, 3

de Serres et al (2014), "An international perspective on the New Zealand Productivity Paradox", New Zealand Productivity Commission Working Paper 2014/01

Deloitte (2011), "Pricing of Public Sector Information Study", October 2011

Department of Internal Affairs (2014), "Cloud Computing, Information Security and Privacy Concerns"

European Commission (2014a), "Towards a thriving data-driven economy"

European Commission (2014b), "Report on the Implementation of the Communication: Unleashing the Potential of Cloud Computing in Europe"

Fioretti (2010), "Open Data, Open Society"

Glass et al (2014), "The Value of Internet Services to New Zealand Businesses", Report for Google and Internet New Zealand

ICTNET (2012), "ICTNET Final Report - Economic impact of ICT: economic evidence and policy drivers"

Katz (2012), "Impact of broadband on the economy", Report for the ITU

KPMG (2012), "Modelling the Economic Impact of Cloud Computing" for the Australian Information Industry Association

LaValle et al (2011), "Big Data, Analytics and the Path From Insights to Value", MITSloan Management Review, 52:2

Law Commission (2011), "Review of the Privacy Act 1993"

Lerner, J (2012), "The Impact of Privacy Policy Changeson Venture Capital Investment in Online Advertising Companies"

Lerner, J. (2014), "The Role of 'Big Data' in Online Platform Competition", August 2014.

McKinsey Global Institute (2011), "Big Data: The next frontier for innovation, competition, and productivity"

McKinsey report (2011), "Open data: Unlocking innovation and performance with liquid information"

Narayanan et al (2008), "Robust de-anonymisation of large sparse datasets", in Security and Privacy, IEEE Symposium, IEEE

New Zealand Government (2011), "Declaration on Open and Transparent Government", August 2011

New Zealand Government (2012), "New Zealand Government Cloud Programme: Programme Business Case", 18 May 2012, released under the Official Information Act.

New Zealand Government (2014), "2014 Report On Agency Adoption of the New Zealand Declaration on Open and Transparent Government", June 2014

OECD (2013), "Exploring data-driven innovation as a new source of growth: Mapping the policy issues raised by "big data", in Supporting Investment in Knowledge Capital, Growth and Innovation", OECD Publishing.

OECD (2014), "Data-driven Innovation for Growth and Well-being", Interim Synthesis Report, October 2014

Office of the Privacy Commission (2014), "Individual Privacy and Personal Information"

Open Knowledge (2014), "Global Open Data Index"

Productivity Commission (2014), "Boosting Productivity in the Services Sector: Final Report", May 2014

PWC (2012), "Big Data - The next frontier for innovation"

PWC (2014), "Deciding with data: How data-driven innovation is fuelling Australia's economic growth"

Schroeck et al (2012), "Analytics: The real-world use of big data", IBM Institute for Business Value

Stevens (2011), "Competition in New Zealand: An analysis using micro data", MED working paper

Sweeney (2000), "Simple Demographics Often Identify People Uniquely", Carnegie Mellon University, Data Privacy Working Paper 3.

Talend (2012), "How big is big data adoption"

Tambe (2014), "Big Data Investment, Skills, and Firm Value", Management Science

Tang et al (2011), "Ridership effects of real-time bus information system: A case study in the City of Chicago"

Unisys (2014), "Unisys Security Index: New Zealand", May 2014

Varian et al (2002), "The Net Impact Study - the projected economic benefits of the Internet in the United States, United Kingdom, France and Germany"

Vickery (2011), "Review of Recent Studies on PSI Re-use and Related Market Developments"

World Wide Web Foundation (2013), "Open Data Barometer"