

Unclassified

STD/CSSP/WPNA(2017)2

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

12-Oct-2017

English - Or. English

**STATISTICS DIRECTORATE
COMMITTEE ON STATISTICS AND STATISTICAL POLICY**

Working Party on National Accounts

**EXPANDING THE USE OF ADMINISTRATIVE DATA IN NEW ZEALAND'S TOURISM
SATELLITE ACCOUNT**

Joint meeting WPNA & WPFS

**To be held on 7-8 November 2017
OECD Conference Centre
Beginning at 11:15 am on the first day**

This document has been prepared by Garry Dunnet - Statistics New Zealand and will be presented under item 8 of the draft agenda

JT03420547

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

STD/CSSP/WPNA(2017)2
Unclassified

English - Or. English

EXPANDING THE USE OF ADMINISTRATIVE DATA IN NEW ZEALAND'S TOURISM SATELLITE ACCOUNT

1.1. Overview

This paper explains Stats NZ's use of electronic card transaction data to enhance the Tourism Satellite Account (TSA), and the most recent use of administrative data to provide estimates of expenditure by cruise ship passengers visiting New Zealand. It sets out the basic history of New Zealand's TSA, and the context in which it is evolving. It describes the sources and methods that are enabling better estimates through the use of administrative data. Finally, it briefly considers in general terms some of the pros and cons of using such data, with reference to aspects of the wider challenges facing Stats NZ in the emerging official statistics environment.

1.2. Background

To begin with, it is helpful to provide some brief history and context about New Zealand's TSA, and a short outline of the rationale for increasing the use of administrative data.

1.2.1. A brief history of New Zealand's TSA

Stats NZ's TSA has been around for over 20 years. The first pilot TSA was released in 1995, and at the time New Zealand was one of a handful of countries producing such an account. As the significance of tourism to the New Zealand economy has continued to grow in recent decades, the TSA has been developed to adapt to new data sources and its scope expanded to provide more detailed insight into New Zealand's tourism sector. It provides some of the benchmark measures that inform the sectors own growth targets, inform regional tourism estimates, and the overall policy debate about tourism's economic contribution. Key measures are now recognised as 'Tier 1 statistics', which have been agreed by the New Zealand government as a set of the most important statistics.

Since 2015, Stats NZ has used electronic card transaction (ECT) data as the basis of estimating households' tourism expenditure, replacing a method previously based on the Domestic Travel Survey.

1.2.2. Key stakeholders in the Tourism Satellite Account

Stats NZ works closely with other government agencies responsible for tourism and has established relationships with key sector groups.

New Zealand's Ministry of Building, Innovation and Employment (MBIE) now incorporates what was once the Ministry of Tourism, and has overall responsibility for tourism policy. MBIE are responsible for running the International Visitor Survey (IVS), a key input to the TSA, but also provide modelling and subnational estimates to augment the information available about the sector. It is MBIE who fund Stats NZ to produce the TSA, via a long-standing Memorandum of Understanding. MBIE are also partners in the use of the ECT data discussed in following sections.

Tourism New Zealand is a government entity responsible for marketing New Zealand to the world, through the '100% Pure New Zealand' campaign which has helped raise New Zealand's profile as a tourist destination.

Tourism Industry Aotearoa (TIA) is the group that shapes overall strategy for New Zealand's tourism sector, and advocates on behalf of the sector's interests. TIA are responsible for the 'Tourism 2025' strategy, for which some of the key targets are drawn from the TSA.

Tourism has a very strong regional dimension in New Zealand, and individual Regional Tourism Organisations (RTOs) specialise in representing and promoting their own regions. Regional Tourism Organisation New Zealand (RTONZ) is the overall national body.

There are many other stakeholders in New Zealand's tourism sector, and a range of specific issues like the continued emergence of cruise ship visits, help drive the development of the TSA. The strategic importance of tourism, coupled with strong growth in the sector, provides a high demand for, and engagement with, official statistics relating to tourism.

1.2.3. The official statistics environment

The uptake of more innovative uses of administrative data in the TSA is also shaped by trends in New Zealand's official statistics environment.

New Zealand has long made use of administrative data as an input to compiling official statistics. For social statistics, the Integrated Data Infrastructure allows researchers controlled access to an increasingly rich set of integrated data drawn primarily from government administrative data and surveys. This has become an important policy-making asset, supporting the New Zealand government's approach to social investment.

On the economic statistics side, we receive regular and comprehensive feeds of tax data from the Inland Revenue that have, among other things, allowed us to significantly decrease the survey samples for some of our key business surveys while extending the coverage of key variables. The Longitudinal Business Database contains integrated data from business surveys, the tax system, trade data, and more. It is increasingly used by the New Zealand productivity research community as a source of insight about firm-level dynamics.

The growing penetration of administrative data into statistical processes reflects Stats NZ's own goals to be an organisation that prioritises the use of administrative data. It also reflects the New Zealand government's wider objectives to promote the more effective use of data across the public sector, and to reduce the compliance burden on New Zealand firms. Stats NZ's custodial role in providing data integration and facilitating the use of administrative data is recognised in what has become referred to as the 'data ecosystem'.

The TSA represents one area where we are moving outside the boundaries of traditional use of government, or publically available private-sector, data. Other examples include efforts to obtain pricing data via supermarket scanner datasets and web-scraping, and property market valuations from private-sector providers. These initiatives provide promising opportunities but also throw up some challenges for official statistics, which are discussed in a later section.

1.2.4. Why use more admin data in New Zealand's TSA?

The changing official statistics environment described above provides some context for the increased uptake of administrative data in the TSA. More specifically, the use of administrative data has been propelled by both push- and pull-factors.

Push factors include:

- The discontinuation of the Domestic Travel Survey (DTS) in 2013. This was previously the main source for estimating domestic tourism expenditure by households.
- A high level of demand for better estimates of cruise passenger expenditure

- An increasing desire to reduce respondent burden on respondents.

Pull factors include:

- The availability of detailed, representative administrative datasets that align with relevant concepts
- Large administrative datasets provide more complete coverage than sample surveys, and enable more robust detailed analysis
- Strong and established relationships with stakeholders who are willing to supply data to Stats NZ
- An organisational environment in which data ‘broking’ is encouraged in order to maximise the sourcing and use of new administrative datasets.

The alignment of these factors has enabled Stats NZ to respond to changing demands by increasing our use of administrative data, rather than undertaking additional survey-based data collection.

The next sections will describe how administrative data has been used to develop better estimates of domestic tourism expenditure for New Zealand households, and visitor expenditure by cruise ship passengers in the TSA.

1.3. Estimating Household Tourism Expenditure

This section explains how ECT data is now used to calculate estimates of household tourism expenditure.

Since the 2015 edition of the New Zealand TSA, Stats NZ has used an administrative data source based on ECT data to collect and determine domestic household travel expenditure. The Household Tourism Expenditure Estimates (HTEE), developed by Stats NZ and funded by MBIE, cover the years ended March 2009 through until the latest year. Prior to the year ended March 2009, Stats NZ used data from the Domestic Travel Survey (DTS) undertaken by MBIE. The DTS collected the expenditure and behaviours of domestic travellers within New Zealand.

The DTS data collection began in 1999, with data available as both quarterly and annual series through to its cessation in 2013. The DTS data provided information on the nature of domestic travel activity, including the origin and destination of domestic travellers. MBIE categorised the data by purpose of travel, expenditure type, and length of trip (either day trip or overnight trip). The four travel purposes were: holiday, visiting friends and relatives, business, and other. The eight expenditure categories were: transport, accommodation, food, alcohol, gifts and souvenirs, recreation, other shopping, and gambling. DTS expenditure was available by purpose of travel, expenditure category, and length of trip.

The DTS was then supplemented with additional household tourism expenditure for outbound travel, off-trip purchases, and imputed rental on holiday homes – using a mix of sources and methods.

In the year ended March 2014, the DTS was replaced by a developmental version of the HTEE, which was further developed and fully integrated into Tourism Satellite Account: 2015. Additional refinements to these estimates are ongoing. The HTEE uses geographic information to determine tourism spending in New Zealand by New Zealanders and is available from the year ending March 2009. The DTS is used in determining prior year estimates.

1.3.1. HTEE source data

ECT data is provided to us by Marketview Ltd, who acquire this from two main sources:

- Paymark – the largest electronic card payment network in New Zealand

- Bank of New Zealand (BNZ) – spending by BNZ cardholders, which excludes any personal identifiers. We call this depersonalised spending.

1.3.2. Paymark data

Data is derived from all transactions made at merchants on the Paymark network. Approximately 70 percent of New Zealand retailers use the Paymark network. The dataset includes all eftpos and credit card transactions made at these retailers. There is no link to the person making the transaction, but transactions are linked to merchants. The Paymark dataset excludes ‘cash-out’ transactions.

From this data a complete valuation of New Zealanders’ spending can be generated, comprising:

- day of the week and time of the day
- where in New Zealand the transaction occurred
- ANZSIC06 (2006 Australian and New Zealand Standard Industry Classification) storetype
- domestic or internationally issued card.

1.3.3. Bank of New Zealand (BNZ) data

The BNZ dataset is based on the depersonalised eftpos (debit card) and credit card spending of approximately 600,000 BNZ cardholders (15 years +) in the New Zealand retail market. BNZ has approximately a 20 percent share of the cards market, meaning BNZ cardholders account for approximately 1 in 5 retail transactions. These cardholders are representative of the national population. The dataset includes spending at Paymark and non-Paymark retailers. It excludes ‘cash out’ transactions and bank transfers.

Through the BNZ dataset, Marketview receives a view of spending at virtually all merchants in New Zealand which receive electronic card spending, regardless of whether the merchant uses the Paymark network or not. They can identify where in New Zealand the transaction occurred and whether the transaction was conducted at a physical store or online.

1.3.4. Sample management

To ensure the BNZ cardholder base is both geographically and demographically distributed in line with the New Zealand population, a weighting is applied by Marketview.

While BNZ cardholders are distributed throughout New Zealand, small variations exist down to an area unit / customer age level. This weighting was calculated by determining the distribution of cardholders and comparing this to the distribution of the overall population.

Marketview use Stats NZ’s area unit population estimates as the basis for the national population. This enables the distribution to change over time, as each year of the data was compared with a different population estimate. For example, Marketview data from 2017 is weighted according to the 2016 population estimates. This ensures significant population changes – such as after the Canterbury earthquakes, or new subdivisions opening – are accounted for in the dataset.

The weighting factor is applied to the dataset by age (in five-year bands starting at 15–19), by census area unit, and by month. This weighting ensures the distribution of BNZ cardholders matches the distribution of the national population, by age, location, and over time. Weighting by age and location ensures management of any bias in the sample, as income and wealth typically increase with age, and wealth can correlate with where a person lives.

1.3.5. Combining data sources

By combining Paymark and BNZ data, Marketview produce a dataset that accurately quantifies:

- the value of spending of each transaction
- the source and origin of those payments eg business vs personal, domestic vs. international tourist
- where in New Zealand the cardholder lives (the area unit the card resides in)
- where each transaction took place eg physical store vs online, Auckland vs Invercargill
- the industry category of the merchants, as defined by 2006 ANZSIC codes
- the time and day of the purchase.

1.3.6. Defining household tourism expenditure

Household tourism expenditure is defined as expenditure that occurs outside a 40km radius of the meshblock in which the cardholder's address is located, and aligns with industries defined as tourism industries. The 40km reflects the New Zealand definition of travel outside one's usual environment. Tourism industries encompass both characteristic and related industry data along with selected non-tourism industries.

Marketview apply this 40km radius to the combined Paymark and BNZ dataset to determine the HTEE. Exceptions are made where regular behavioural spending patterns show a person's usual environment extends to an area outside the 40km radius, such as commuters. This is removed from the HTEE.

Additional data on internet transactions is collected specifically for selected tourism industries that require travel in order to consume a purchased good. For example, internet expenditure on accommodation and air passenger transport is collected.

1.3.7. Scaling household tourism expenditure data to total economy

As ECT data reflects only one aspect of household tourism expenditure across the New Zealand economy, Marketview upscale their dataset by adding in a factor for cash and other payment methods. This is calculated as the difference between electronic card spending and total economy spending based on the Australia and New Zealand Standard Industrial Classification (ANZSIC) industry information supplied from Stats NZ's Annual Enterprise Survey (AES).

For example, Marketview may record the total value of electronic card spending in ANZSIC industry G4110 at \$100 for the year, with 10 percent being tourism (\$10). The total industry value of G4110 as calculated from AES was \$120. The Marketview card value is thus upscaled by a multiple of 1.2, yielding a total market value of \$120, consistent with AES. The tourism component is still 10 percent, hence tourism spending for that year is calculated at \$12.

The assumption used is that consumer and business spending on cash versus card on tourism and non-tourism related trips are equal.

1.3.8. The HTEE dataset

The HTEE dataset provided by Marketview presently covers the years 2009–16. At the time of compilation, AES data was available to the 2015 financial year. To produce the HTEE through to 2016, Marketview have estimated the value of each industry in the 2016 provisional year by applying movements for each industry from additional Stats NZ data sources, including GST data, to the 2015 AES data.

For example, Marketview took annual movements in spending for ANZSIC industry G4110 from the Retail Trade Survey. They applied this to the 2015 AES data to determine a 2016 provisional estimate. They estimated other industries from data indicators sourced from Stats NZ.

Marketview will update the provisional year estimate as AES data becomes available and indicator data is revised a part of the annual publication cycle of the TSA.

1.3.9. Turning industry based HTEE into tourism products

The HTEE industry dataset is then broken down into tourism defined products using annual supply-use commodity proportions and retail industries sales data. For validation purposes it is then confronted against Household Consumption Expenditure commodity data net of overseas visitor expenditure and New Zealanders' travel expenditure abroad. This isolates New Zealanders' spending within New Zealand, allowing for a comparison on an equivalent expenditure basis with the HTEE.

1.3.10. Additional household tourism expenditure

While the HTEE dataset provided by Marketview captures most household tourism expenditure, the TSA supplements the HTEE product breakdowns with its own product expenditure estimates. These include some off-trip purchases of tourism-specific consumer durable goods and imputed rental on holiday homes.

Both the HTEE and additional Stats NZ tourism product data then provide the initial expenditure levels to feed into the balancing process. These levels can be subsequently modified where necessary.

1.4. Using administrative data to develop cruise expenditure estimates

Most recently, Stats NZ has expanded its use of a composite set of administrative data to provide estimates of visitor expenditure by cruise ship passengers visiting New Zealand.

Historically, New Zealand's international visitor expenditure measure and macro-economic outputs, including the TSA, have not captured the full value of expenditure undertaken by cruise ship visitors.

New Zealand's IVS surveys airport departures only therefore captures those cruise visitors who complete their cruise in New Zealand before flying out. This scope has not accounted for the significant and growing number of cruise visitors who cruise in and cruise out of New Zealand.

While this expenditure represents a small portion of total visitor expenditure, it has been recognised as a coverage gap and there is demand for a clearer picture of the contribution of the cruise industry to the tourism sector.

In order to address this under coverage in these statistics, Stats NZ and MBIE have worked alongside the New Zealand Cruise Association (NZCA) to source and utilise administrative data to inform expenditure estimates with the added benefit of a regional port dimension.

The key data insights that have enabled this development are primarily:

- Cruise ships schedules
- Cruise ship manifests of passenger and crew
- International card transaction data.

1.4.1. Overview of approach

For each cruise season, NZCA provides a ship schedule outlining the dates and port locations together with arrival and departure timings relating to each specific cruise ship visit. Cruise ship manifests containing the count and details of passengers and crew are sourced from the New Zealand Customs Service. International card transaction data is acquired from contracted consumer-spending analysis firm Marketview, providing transaction date, time, country of card issue, location (territorial authority), transaction value, and industry classification.

A cruise visitor can be identified as any international cardholder who makes a transaction on two dates and in two territorial authorities that align with a particular cruise ship voyage. Timing of expenditure must

also occur when the ship is in port (allowing for disembarking and embarking times). Specific exclusions are applied to expenditure on goods from certain industries, and card expenditure occurring outside a defined time before and after a cruise ship visit.

Local shore excursion operators and RTO's were involved in defining a realistic geographic distance around each port in which a cruise visitor could be expected to travel and spend.

Unique international passenger and actual crew counts by nationality (passport) are used to scale the determined card spend. Ratios of card-to-cash expenditure from the IVS are used to impute each nationality's cash expenditure in order to estimate total expenditure.

The resultant output is available by quarter, by country of origin and by port. It is subsequently added to provisioning, bunkering, shipping agent and shore excursion expenditure data all sourced directly from firms and operators. This enables a comprehensive cruise expenditure impact estimate by season to be derived and incorporated into macro-economic outputs.

For cruise passenger expenditure, the use of a variety of administrative sources together has enabled Stats NZ to fill a difficult-to-measure gap in coverage in the TSA. Working with the sector has allowed clear definitions to be set that can be operationalised through the detail available in the administrative data itself.

1.5. Opportunities and challenges ahead

The preceding sections have described two cases where administrative data has been used to improve New Zealand's TSA. Administrative data opens up new opportunities, but the use of increasingly diverse and composite administrative sources also brings challenges. This section offers some general thoughts about the pros and cons of incorporating more and more administrative data into the tourism estimates.

1.5.1. Where next?

Following the success of ECT data in the New Zealand TSA, Stats NZ continues to look for other ways to extend our use of administrative data to provide a better picture of tourism in New Zealand.

Conversations are underway about the potential to use data from peer-to-peer accommodation providers to increase the understanding of the contribution of this type of activity to the NZ economy.

To date providers have been generally willing to share their data with Stats NZ, under the appropriate conditions, at least to further exploratory work.

1.5.2. The opportunities: coverage, granularity, efficiency

Some of the advantages of using these novel datasets are obvious. The ECT data provides a representative sample of a significant portion of New Zealand's domestic tourism transactions, at no additional burden to respondents. The coverage allows more robust detailed estimates, and data is available in a timely way and at comparably low cost.

There are, however, some challenges to face as we extend our use of this kind of data.

1.5.3. The challenges: managing a distributed statistical supply chain

As administrative data plays an increasing role in compiling official statistics, the needs increase for us to reshape our traditional approach to statistical management. The section below touches on a few areas where new considerations and approaches are emerging, in response to control of the statistical supply chain becomes more distributed. Some of these are particularly relevant for emerging and novel uses of private-sector administrative data.

Continuity of supply

Data collection becomes the responsibility of the provider, many operating in competitive commercial environments. The provision to Stats NZ of data in this way is also likely not to be governed by the Statistics Act 1975, and is thus reliant on the willingness of the provider to continue to supply it.

Contracts and building relationships over time can help mitigate the loss of a key provider, and statistical production has always been subject to the discontinuation of data collections.

Managing data quality

Where data collection (and in some cases compilation) is no longer the direct responsibility of the NSO, it is important to develop a good understanding of the quality of data being supplied. In some cases, clear data supply agreements are in place that specify metadata that must be supplied along with the data itself as a quality assurance mechanism.

As we move to an environment where administrative data makes up a greater proportion of the raw statistical material, a system-wide approach to defining and managing data quality is also evolving. Stats NZ has been asked to play a lead role in this conversation across government in New Zealand, working to bring a common approach to this. The use of models like the 'Steady State' statistical architecture may provide part of a platform.

Trust and sharing of data

For many of the exploratory cases where private-sector providers have supplied data, they have to date done so on an ad hoc basis. Data is usually provided based on the protections to sensitive commercial data guaranteed under the Statistics Act. It is also usually subject to agreed restrictions on the wider sharing of data beyond the immediate narrow use for which the data was requested. Turning these one-off arrangements into an approach that balances the providers' trust and maximising the wider utility of these datasets is part of an ongoing conversation.

Maintaining independence

The mission of NSOs and other official statistics providers rests largely on the independence of estimates. Where commercial stakeholders in the statistics also become data providers for the compilation of those statistics, this raises interesting questions about the processes and controls necessary to ensure the ongoing independence of estimates.

The legislation governing statistics in New Zealand (the Statistics Act 1975) is under review, which may provide a more up-to-date framework for governing less linear data supply arrangements. Many of the emerging uses of this kind of data are exploratory and at an early stage. As they become more embedded in the compilation of official statistics, more formal governance arrangements will evolve alongside.

Infrastructure and data management

In many cases, the central management and storage of administrative data by Stats NZ is relatively mature. Infrastructure and processes have been established to manage the transfer, updating, storage and processing of the larger (mainly government) administrative data sets for which we have responsibility.

As we procure smaller, niche datasets to fulfil particular needs in a more agile way, we also need to evolve standard systems and protocols for transferring and storing data.

This is by no means an exhaustive set of the issues that Stats NZ faces as it begins to grapple with non-traditional data supply. They represent the subset that are currently at play in the context of developing the TSA, which itself is a good example of new ways of producing statistics. As in many parts of the world New Zealand's official statistics environment is changing, and the National Accounts is guided by the wider organisation to navigate these new directions.

1.6. Conclusion

The incorporation of administrative data into the New Zealand TSA has been a success story. It has replaced a previously burdensome survey, improved the quality of estimates, and allowed for better insight into the dimensions of a key part of the New Zealand economy. It has been an example of pushing the customary boundaries of using administrative data traditionally sourced from government.

Stats NZ is generally still at an early stage of partnering with non-government providers. Based on successes to date, we are exploring other areas where this kind of data may augment traditional data collection and help fill some otherwise difficult data gaps.

The TSA examples illustrate some of the wider challenges and opportunities facing NSOs in adapting traditional statistical collection and compilation process.