

Successfully using electronic card transaction data to produce regional estimates of tourism spend: The New Zealand experience

Michael Webster
Principal Analyst – Business and Economic Development
New Zealand Ministry of Business, Innovation and Employment

Introduction

New Zealand uses electronic card transaction (ECT) data to produce regional estimates of tourism spend. This paper will cover the benefits, limitations and success of this method, along with possibilities for deeper economic insights. After providing some background about why the Monthly Regional Tourism Estimates (MRTEs) were developed, the paper will reveal how they are generated and used, along with how they might be improved. Next, an innovative method for estimating onshore cruise expenditure from ECT data will be discussed. The findings presented will be relevant to researchers and policy-makers who are interested in gathering insights from administrative data sources, and in capturing more granular estimates of regional tourism spend.

Background

2011 Tourism Data Domain Plan

The Ministry of Economic Development (MED) produced New Zealand's first Tourism Data Domain Plan (TDDP) in 2011. The TDDP identified accurate estimates of regional tourism spend as a key gap in New Zealand's tourism statistics.

The primary regional measures of tourism spend in 2011 were the International Visitor Survey (IVS), a sample survey of around 5,000 international visitors at airports; and the Domestic Travel Survey (DTS), a survey of 15,000 New Zealand households per year. Neither survey had sufficient sample sizes to obtain appropriately robust measures of regional tourism spend. It was also noted in the TDDP that it was unlikely that the two surveys could be improved sufficiently to provide accurate regional information.

In mid-2012, MED was combined with a number of other government departments to form the Ministry of Business, Innovation and Employment (MBIE), which oversees the New Zealand government's tourism research programme

The Regional Tourism Estimates and the Regional Tourism Indicators

MBIE took an innovative approach to identify tourism spending by using Electronic Card Transaction (ECT) data as the basis for the Regional Tourism Indicators (RTIs), which was launched in 2012 as a monthly index. The aim of this approach was to provide tourism operators with a regional, regular, up-to-date view of tourism spend and growth.

Alongside this, the Regional Tourism Estimates (RTEs) were launched in 2013 to provide regional yearly dollar estimates of tourism spend using the RTIs, weighted using International Visitor Survey (IVS) and provisional data from the Tourism Satellite Account (TSA)¹. This functioned to produce yearly estimates of regional tourism spend.

Following extensive consultation, however, the RTI indexes were found to be difficult to understand for users as they weren't linked in a meaningful way to spend figures. There were also various methodological issues that reduced the usefulness of the statistics. Following a review of the indicators in 2015, a process began to completely redevelop the process from the ground up, to produce more meaningful monthly spend estimates.

Monthly Regional Tourism Estimates

The Monthly Regional Tourism Estimates (MRTEs) were developed in late 2015 to mid-2016. Combining the dollar value estimates from the RTEs with the monthly frequency of the RTIs, these monthly releases provide a dollar value estimate of international and domestic tourism spend in the regions.

Regional Tourism in New Zealand

New Zealand generally represents its tourism industry regionally through Regional Tourism Organisation Regions (RTO Regions) geographical units. These areas of "touristic interest" are generally individually marketed by a Regional Tourism Organisation or Economic Development Agency. There are 31 RTO regions in New Zealand, composed of varying geographical and population size. A more detailed regional breakdown is the Territorial Authorities (TAs), of which there are 67.

Definition of tourism goods and services

The MRTEs define tourism goods and services using a subset of the ANZSIC06² product categories from the TSA, which includes: accommodation; retail sales - alcohol, food and beverages (supermarkets); other retail sales; food and beverage services (cafés and restaurants); cultural, recreational, and gambling services; fuel; and land and sea passenger transportation.

Several TSA categories, however, are not included in the MRTEs:

- *Air passenger transport*: as most flights are purchased overseas, ECT data does not adequately cover this area.
- *Imputed rental on holiday homes*³: ECT data does not sufficiently cover this area.
- *Education services*: as this category includes students in New Zealand for up to 12 months, it is difficult to distinguish their spending from the domestic population, so this category is excluded.

¹ The TSA is described in more detail on page 4; *Summary of data sources and process*.

² Australia and New Zealand Standard Industrial Classification 2006.

http://archive.stats.govt.nz/browse_for_stats/industry_sectors/anzsic06-industry-classification.aspx

³ An estimation of the rent that an owner of a second home (used only, or partly by the owner) effectively pays to themselves. This calculates the real benefit/income received, despite there not being a financial transaction.

Methodology of the Monthly Regional Tourism Estimates

Having provided the background to why the MRTEs were developed, the next sections will explore their methodology, use and potential future improvements.

Summary of data sources and process

The MRTEs estimate spending for geographic areas, product categories and countries using:

- a base of ECT data,
- International Visitor Survey (IVS) data from MBIE, and
- The Tourism Satellite Account (TSA) dataset from Statistics New Zealand.

The International Visitor Survey (IVS)⁴ is a quarterly survey of departing visitors that provides information on the characteristics, behaviour and expenditure of international visitors. A summary of spending by country, year and quarter is used from the most recent IVS in calculating the MRTE dataset. IVS spending includes prepaid and package spending in the total spending.

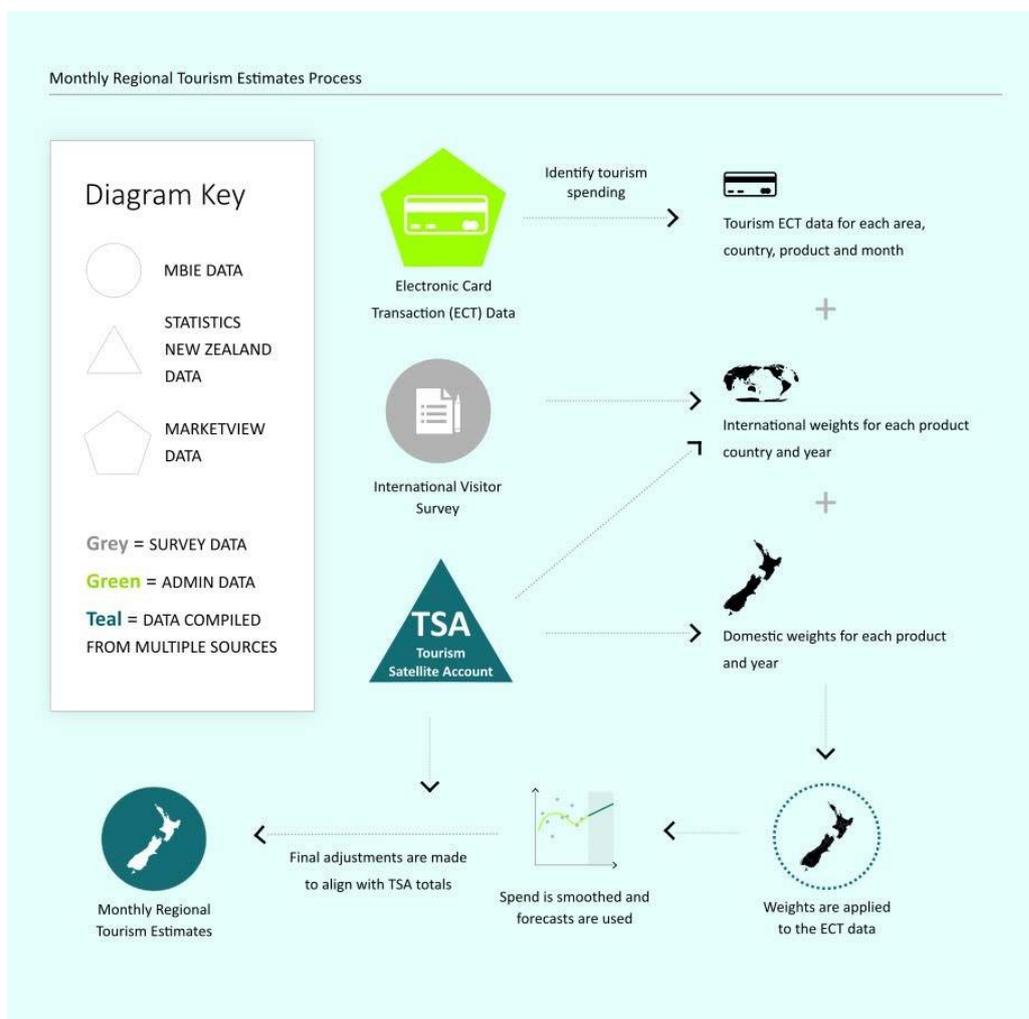
The TSA⁵ measures the contribution of tourism expenditure to the New Zealand economy and is categorised into 11 groups covering all tourism-related spending. The TSA is published each October for the preceding year ending March. The TSA is revised for the whole series each year and is compiled under the United Nations World Tourism Organization framework.

The final values calculated by the MRTE process represent total spend in New Zealand dollars excluding GST. See Figure 2: Monthly Regional Tourism Estimates process.

⁴ <http://www.mbie.govt.nz/info-services/sectors-industries/tourism/tourism-research-data/ivs>

⁵ <https://www.stats.govt.nz/information-releases/tourism-satellite-account-2017>

Figure 1: Monthly Regional Tourism Estimates process



Identifying tourism spend from ECT data

MBIE purchases ECT data from Marketview Ltd, a commercial data provider, specialising in the management and analysis of ECT data for consumer spending. MBIE receives anonymised aggregated spend values for total card spend and transaction counts which are broken down by Territorial Authorities (TAs), Tourism Satellite Account (TSA) product category, and country of origin (international) or region of origin (domestic). The full dataset is updated monthly to reflect new information about merchants.⁶ Data is sourced from two payment networks by Marketview; Paymark and Bank of New Zealand (BNZ).

The Paymark network is used by approximately 75% of New Zealand retailers and provides information on the overall card spend in all locations, including the nationality of cardholders. When a merchant is not on the Paymark network, the value of international expenditure is estimated based on spending at other similar merchants.

⁶ Sometimes, merchants are categorised incorrectly resulting in changes to the historical series. These changes are typically very small and do not make appreciable differences to the historical series.

The BNZ network provides the location of cardholders, merchants and spend values. The New Zealand market includes approximately 700,000 BNZ cards at any one time, representing around 20% of electronic cards and accepted as being representative of the broader New Zealand population for product and geographic spending patterns. This data is used to determine domestic tourism patterns. To ensure customers entering and leaving the BNZ system do not adversely affect results, only customers that have had at least one transaction per month for the previous 12 months are included in the dataset.

Marketview applies a number of rules to identify the tourism spend which is then used to calculate the MRTEs. Non-tourism card spending is part of the raw dataset, but is not used in MRTE calculations. The following sections describe the rules applied to identify tourism spending.

Defining domestic tourism

Marketview creates a “primary environment” for each domestic cardholder from the location linked to their card, deemed to be where the card holder lives, and this helps to identify day to day spending which is not included under tourism. In the case of commuter spending a “secondary environment” is created for a location where a cardholder regularly spends in 26 weeks out of the previous 52, and this is not counted as tourism.

An assumption is made that domestic card holders do the bulk of their everyday shopping from their “primary environment”. If spending occurs outside a 40 km radius of the cardholder’s “primary environment” (or “secondary environment” if this is created for the cardholder), the spending is categorised as tourism. The 40 km reflects the New Zealand definition of travel outside one’s usual environment.

Some domestic cardholders spend extended periods of time away from their “primary environment” (e.g. students living away from their home address). For these cases, a “secondary environment” is created in the location where a cardholder has at least 75% of their spending and at least six transactions over the last three months on supermarket and grocery coded spending, in a different Territorial Authority to that which is shown as their place of residence.

Centralised accounts and online spend

Some merchants process all transactions in a central location (e.g. rental cars). As it is not possible to determine the origin of all transactions for that merchant all spending for the merchant is attributed to the location with the most substantial proportion of the merchant’s spending.

Some spending is not able to be attributed to a specific product category and is attributed to online spend. Online spend is distributed proportionately across all regions through the weighting process.

Weighting process

The MRTEs apply what is happening at a national level to the regions. This application is done by weighting the Marketview data (which represents card spending only) to a national view of spending which includes cash spending and un-attributable online spending.

Domestic spend weights

A set of weights is created to match the domestic spend to TSA totals. A summary of ECT data is created for domestic spending by product and year ending March is then matched with the TSA domestic totals (defined as household demand and government demand) for each product and year using 'iterative proportional fitting' (otherwise known as 'raking'). In raking, many small adjustments are made to a set of weights to find the optimum fit while maintaining equal totals.

For each product and year:

$$\text{Weight} = \frac{\text{TSA household demand} + \text{TSA government demand}}{\text{Total card spend}}$$

International spend weights

A set of weights is also created to match the international spend to TSA totals. The process differs slightly to the domestic spend. The total year ended March IVS spend is first adjusted to match the TSA totals.

$$\text{IVS adjusted spend} = \text{IVS spend} \times \frac{\sum \text{TSA international demand}}{\sum \text{IVS spend}}$$

Summary ECT data is then created for international spending by product, country and year ending March and aligned with the TSA international demand and IVS adjusted spend totals using the raking process described above.

The result from this step is a weight for each year, product and country:

$$\text{Weight} = \frac{\text{TSA international demand}}{\text{Total card spend}}$$

Smoothing function

The set of weights is further refined to include a monthly component and is smoothed to reduce the distortive effect of outliers. Weights are applied to each of the international and domestic datasets, giving values for each month, product, area, and country (for international). A smoothing function is then applied to the weights. Smoothing functions often use a curve (or spline) with a number of fixed points called knots. The smoothing function used in this calculation is a fitted spline with 24 knots. The script iterates through applying the smoothing algorithm until an optimum fit is achieved with minimal difference from the annual year-end March totals in the weighted dataset from the TSA totals for each product, country and year.

National total forecasts

National totals from the TSA are forecasted forward to generate monthly estimates for the MRTEs. As the TSA is released in November for the year ending March, forecasts are produced to create national totals for the MRTEs to match the current month. Domestic and international datasets are still kept separate at this point, as they are matched again to the TSA totals later in the process.

For each year, product, and country (for the international dataset) combination a forecast is made of the next three years using the exponential smoothing state space model (ETS) algorithm. ETS is a general and flexible forecasting algorithm that models error, trend and seasonal elements together. This forecasting process requires an annual reweighting of all MRTE series as new data from the TSA is released, to revise the forecasted totals to match this data.

Total spend weighting

The total spends for each month, product, and country (for international) is calculated by multiplying the weightings by the total card spend given in the Marketview data:

$$Total\ MRTE\ Spend = Total\ Card\ Spend \times Weight$$

The international and domestic datasets are then bound together as one.

Final TSA-level weighting

For each year and product combination, small adjustments are made to ensure that the final results align with the TSA totals:

$$Adjustment\ to\ TSA = \frac{TSA\ Total}{MRTE\ Total\ Spend}$$

$$Final\ MRTE\ Spend = Total\ MRTE\ Spend \times Adjustment\ to\ TSA$$

The final dataset is saved and then processed into various formats.

Use of the MRTEs

The MRTEs have become a widely used resource for the tourism industry. RTOs (Regional Tourism Organisations), the tourism marketing arms of local councils, use the data for key performance indicators (KPIs). The data is used by the wider industry to support business plans and cases, as well as informing Government policy development related to regional and seasonal dispersal. The data was particularly successful in showing the recovery from the November 2016 Kaikoura earthquake and driving Government policy in this area; and showing the impact of the DHL British and Irish Lions' rugby tour in June and July 2017.

As it is used as a performance indicator, the MRTEs have received criticism due to its revision process. Because the revisions can potentially be substantial (dependant on changes in the TSA) this can have a substantive impact on historical series, and hence, performance measures. MBIE has clearly explained that the data is provisional and estimates only, however, this remains an issue for RTOs.

In some of the smaller RTO regions, such as the Ruapehu district, spending estimates have been criticised for not matching 'on the ground' data. Ruapehu tourism spending on cultural and recreational services is dominated by one large business, Ruapehu Alpine Lifts. The MRTE estimate for Ruapehu cultural and recreational services is substantively larger than the publically available annual report income for the company. This is partly due to the way that online spend is

proportionally attributed across the country, and the fact all spending is weighted up to meet TSA totals. This problem shows that the data should be used with caution for smaller regions.

Electronic card transaction data, along with other administrative sources, is seen widely as a highly accurate source of information. As a result, the MRTEs have been criticised for using survey data (from the IVS) as undermining the accuracy of the underlying ECT data, despite the partial nature of this data.

Potential future improvements

Forecasting and revisions approach

MBIE is exploring options to minimise the disruption to users caused by large annual revisions to the MRTEs. One viable option is to make more frequent use of the quarterly IVS when weighting and forecasting international spending in the MRTEs. By recalibrating to the most recent IVS data every three months instead of once a year, the MRTE international spend series would be revised more frequently, but it would minimise the impact of changes that come through when the new TSA results are used. Ultimately the IVS estimates are used in the annual TSA figures regardless. In other words, the growth patterns in international spending in the MRTEs will eventually be recalibrated to the IVS spending patterns when these annual revisions are made.

Quality measures

There are different levels of quality for each regional estimate in the series. What affects quality the most is the number of merchants for which data is collected – this can vary between regions. There are two main EFTPOS services in New Zealand- Paymark and Verifone. While Paymark covers around 75% of national transactions, this can vary significantly at a regional level. When a large number of tourism businesses, or one particularly significant business, is not included in the Marketview data, this can lead to less accurate estimates of spend, especially in smaller regions. MBIE intends to publish a quality measure (developed with Marketview) that will help provide users with a level of confidence for each specific series.

Improving centralised accounts

Currently, centralised accounts spend is attributed to the region with the most activity. Stats NZ administers a register of businesses that in theory could provide proportional estimates of business activity by geographical unit. One possible improvement could be to access this data and apply proportions to the spend to better attribute spending across the country.

Pre-purchased spending

The MRTEs currently do not include tourism spending occurring before tourists enter the country, such as pre-purchased package deals. Marketview has access to ECT data from other countries, such as the UK and US, however, so estimates for pre-purchased spending could potentially be incorporated into the MRTEs.

Onshore cruise expenditure estimates

Having described the MRTEs, this paper will now explore another innovative approach to using ECT data to produce tourism statistics, this time to develop estimates of international onshore cruise passenger and crew spend in New Zealand.

Background

In the 2011 Tourism Data Domain Plan, onshore cruise ship spending was identified as another gap in New Zealand tourism statistics.

To enable the provision of a comprehensive and official measurement, three specific areas were identified where key cruise related economic data were required:

1. Providoring – examples include berthage, refuelling and resupplying ships
2. Shore excursion – reflecting the likes of pre-booked tours and activities primarily through cruise lines
3. Onshore expenditure – independent spend undertaken by passengers and crew.

Of these, the most challenging area to measure related to onshore expenditure. Most cruise passengers are transit passengers, meaning they are excluded from the International Visitor Survey. Additionally, most data on cruise passengers are held by cruise lines, and difficult to access.

In early 2017, MBIE, along with Stats NZ, New Zealand's main statistical agency, began work to develop a new method to produce cruise passenger and crew spend estimates, using ECT data, again from Marketview. This approach involves the use of card transaction data together with cruise ship scheduling to aid the identification and derivation of expenditure by passengers and crew.

Estimates of cruise expenditure were first published in December 2017.

Method

Scheduling

For every 'cruise season' there is a schedule outlining the dates and port locations together with arrival and departure timings relating to each specific cruise ship visit. This has been obtained historically from the New Zealand Cruise Association. There were some limitations to this approach, however. The schedule was manually updated and required the producer to update schedules based on changes due to weather or medical emergencies. This did not always happen. Moreover, the times at which ships entered port were often inexact.

As a new development in the 2018 year, AIS (Automatic Identification System) geolocation data was used from a database administered by Maritime New Zealand (a Government department that regulates the use of New Zealand's waterways). This dataset contains GIS data for all ships in New Zealand waters, with a new record being generated for each ship every five minutes. Using this data, the time at which cruise ships entered port could be much more accurately estimated. The data shows the ships' speed at any given time, which meant that the precise time when the ship docks at the port could be calculated (speed approximately equal to zero).

Matching the schedule with ECT data

A cruise ship passenger was defined as any international cardholder who makes a transaction on two dates and at two territorial authorities that align with a particular cruise ship voyage.

Marketview takes the schedule and matches the list of dates, times and locations with their data on all international card transactions in the Paymark network. They then apply a number of rules to exclude cardholders who behave in ways that are not possible nor reflective of a cardholder being on a cruise voyage.

- The transactions must be made in the time window between one hour after the ship's arrival and one hour before the ship leaving port. Outside of this window, they are excluded.
- A 'catchment area' was defined for each port analysed, representing the longest distance a passenger could be expected to roam from a port. If a cardholder made a purchase outside of this area, they are excluded.
- If the cardholder makes a transaction with a merchant deemed to be an accommodation provider, they are excluded.

Marketview then internally scale the results to represent the market share on the Paymark network, using their BNZ dataset, which is deemed representative of all transactions and merchants in the New Zealand economy.

Once cardholders are identified as being on a cruise at some point during their visit, estimates can be made on their expenditure pre and post cruise visit.

Current exclusions from estimates

New Zealand cruise passengers are currently excluded from the figures. The domestic ECT dataset was likely to pick up too many false positives when estimating domestic cruise spend. There were cases where people travelling between popular commuting/travelling points, such as the cities Auckland and Tauranga, could coincidentally produce patterns of transactions similar to cruise ship schedules, and incorrectly be assigned cruise spend. The potential for false positives is much smaller for international spend.

Several cruise ships stop in one port in New Zealand, usually Auckland, as part of a wider Australasia cruise. Passengers from these ships currently are not included in the estimates, as they cannot be easily identified.

Cash spending remains a big gap. Anecdotal evidence suggests that cash makes up a relatively high proportion of overall spending, and the cruise ships often serve their passengers with automatic teller machines (ATMs) on board. There is currently no way of identifying this spend.

Use of the onshore cruise expenditure estimates

These estimates are now part of official statistics and incorporated into the Tourism Satellite Account, and hence the Balance of Payments and National Accounts. While there are limitations on what is produced, it is an improvement from the assumption-based model that was used previously. The industry mostly appreciates that the estimates are now official and publically available, though there were some concerns around the difference between the numbers and an independently-developed economic impact assessment (EIA) on the impact of the cruise sector. The official figures were significantly smaller than what was estimated in the EIA.

Conclusion and recommendations

Developing the MRTEs and cruise spending estimates using ECT data has shown that useful insight can be drawn from these administrative datasets, even if the information departs from traditional and statistically robust methods, such as surveys. The clear advantages are:

- The data does not require any additional burden on people (for example, asking them to fill in a survey).
- There are many millions of responses in these data, which means they can be looked at in more granularity.
- It is much cheaper than traditional sources of data, such as surveys.

However, these advantages must be mediated by the following:

- The data is collected for another purpose, which means that it could be subject to bias and/or error that is often difficult to determine.
- The proprietary nature of much of this data means that access to the information can be limited, and ownership may need to be negotiated.
- Privacy and confidentiality must be protected, as well as ensuring that the data is permitted to be used for the proposed analysis.

In New Zealand's experience, the challenge lies in understanding and explaining to the public the limitations associated in using these data.

Currently, the data should be seen as complementary to traditional data sources, not a replacement – at least until the limitations are adequately understood and resolved.