Redeveloping the confidentiality method for Statistics New Zealand business demography data

Alistair Ramsden and Caleb Moses
Statistical Methods
Statistics New Zealand
New Zealand
alistair.ramsden@stats.govt.nz and caleb.moses@stats.govt.nz

Abstract

Context: Statistics NZ produces official business demography statistics, including counts of businesses and employees, disaggregated by categorical variables such as geography, industry, and employee count size group.

Objectives: Statistics NZ has recently redeveloped how it confidentialises business demography statistics, to meet identified customer priorities, and also statutory responsibilities. This aligns with Statistics NZ's generic statistical business process model, other Statistics NZ strategic priorities, the NZ open government information and data programme, and the NZ data and information management principles.

Key Messages: We are analysing prospective confidentiality methods, determining which method is best, and implementing it. The proposed method uses a series of retained random numbers aligned with the series of raw interior cell data, modulo calculations on these retained random numbers to generate consistent exterior cell random numbers, random rounding, and bounded random noise, to automate producing confidentialised counts and count magnitudes. This method arguably produces limited data risk and data utility loss.

Discussion: The proposed method relies on a corpus of previous methodological work. The proposed method is a test case for other work in progress developing a general automated confidentiality service (ACS), which seeks to automate key aspects of confidentiality within the Integrated Data Infrastructure (IDI) and NZ Census.

1 Overview

1.1 What are Statistics NZ business demography statistics?

Statistics NZ business demography statistics (BD) provide an annual snapshot of the structure and characteristics of NZ businesses.

In 2015, 13 NZ.Stat BD tables with associated machine readable files were released by Statistics NZ. These tables and files reported BD measures such as counts of business enterprises, counts of business geographic units, and count magnitudes of businesses' employees, and several categorical variables. The largest BD machine readable file was 122mb, and included 38.8 million rows of data (Statistics NZ, 2016c).

1.2 What is statistical confidentiality?

Statistical confidentiality is the stewardship of data for statistical purposes, and includes responsibility for both protecting data and ensuring its beneficial use. It requires proper practices for both providing and restricting access to data products (Duncan et al, 2011).

Statistics NZ is legally required to protect the information of individuals and businesses, and we apply different methods to do this, depending on the type of data (Statistics NZ, 2015b). Key goals of statistical confidentiality include utility, safety, simplicity, and consistency. The first two goals together – utility and safety – lead to rules that aim to release as much detail as possible, as well as to protect entities that need to be protected (Statistics NZ, 2015a).

In other words, we must balance the competing objectives, of allowing statistical analysis of confidential or private data, while maintaining standards of privacy and confidentiality (O'Keefe et al, 2013a). This creative tension is explicitly reflected in Statistics NZ's most recent corporate strategy, and in the New Zealand Government's open government information and data programme, and data and information management principles (Statistics NZ, 2016d) (Open Data NZ, 2016a and 2016b).

Ensuring we protect data, ensuring its beneficial use and non-misuse, and ensuring its outputs are not be considered *'creepy'* by the public or the Privacy Commissioner, are all relevant objectives of the legislative and regulatory frameworks which underpin statistical confidentiality. This is the case, even where terms like beneficial and *'creepy'* are highly subjective, and even where this subjectivity and the relevant legislative and regulatory frameworks change over time.

1.3 What were the confidentiality problems with business demography data?

Before 2015, BD statistics were published as unrounded data, and relied on a Statistics NZ confidentiality rule called the *'drive-by'* rule, which maintained that this data was already in the public domain, and therefore did not need protection. In 2015, a review by Statistics NZ and the Crown Law Office determined that this rule was flawed and should not be relied on, and that a new confidentiality method should be put in place.

The confidentiality method for the 2015 BD release was therefore changed, and based on the Statistics NZ confidentiality standard for business collections (Statistics NZ, 2013). The 2015 method perturbed and suppressed BD data. In particular:

- counts of businesses were perturbed using random rounding to base 3 (RR3)
- counts of businesses of 1 or 2 were suppressed as '(c)'
- count magnitudes of businesses' employees were perturbed using graduated random rounding (GRR)
- count magnitudes of businesses' employees based on counts of businesses of 1 or 2 were suppressed as '(c)'.

These changes to the 2015 BD dataset were a concern to many customers. Our legal boundaries constrained what data we could release, but the confidentiality method we used in 2015 rendered our NZ.Stat tables unusable for many. Statistics NZ is committed to making more data available, not less. Balancing confidentiality and our current legislative environment, with the desire to release more data, has driven work over the last two years to develop a better solution for BD data in 2016 (Statistics NZ, 2016b).

The 2015 changes – particularly suppression – failed to meet key customer needs. Investigation into how to better meet customer needs identified, among other things, that the suppression implemented in 2015 is in many cases breakable and ineffective.

Further doubts were raised about the effectiveness of GRR as sufficient confidentiality protection for count magnitudes of businesses' employees. These doubts suggested that the stronger p percent (P%) rule should be used instead of GRR. P% was therefore also investigated, and a significant number of breakages of this rule were identified in tested expected high risk data samples.

1.4 What is the P% rule?

The P% rule is used by Statistics NZ to determine values which need protection, particularly dollar value magnitudes, but also any other magnitudes where the level of disclosure risk does not allow for a lower level of protection (Statistics NZ, 2013). For the P% rule, a value needs protection if the second largest contributor to that value can estimate the largest contributor's value to P% of its real value. (Statistics NZ, 2015b).

For each output value, we can sort the contributors – for BD data, contributors are businesses – as largest, second largest, and so on, to the smallest one. *P* is calculated as: *P = ((total - second largest) - largest) / largest.* (Where there are weights, the total is the weighted total for the cell.)

Where *P* is below the threshold value P%, then *P*'s largest contributor's value can be estimated with unacceptably disclosive precision by the second largest contributor. The total output value, which is the sum of all the contributor's values, thus needs to be protected (Statistics NZ, 2013).

1.5 What was the proposed solution?

We considered a range of potential confidentiality methods to solve these problems. In doing so, we identified a connection between this BD confidentialisation work and Statistics NZ's automated confidentiality service (ACS) project. In particular, we identified the potential effectiveness of two ACS methods, noised counts and magnitudes (NCM), and fixed random rounding to base 3 (FRR3).

Further investigation of NCM demonstrated NCM eliminates the need for counts suppression, and sufficiently protects magnitudes against potential P% rule breakage.

2 What were our initial research questions?

2.1 Initial investigation

We wanted to better understand the potential results of an attack on the largest 2015 BD release file. This file contained NZ business counts and businesses' employee count magnitudes by region, industry, and employee size group. We tested the protection provided in 2015 by RR3, GRR, and the suppression of count 1s and 2s and their associated count magnitudes with '(c)'s. This involved mining the data to identify combinations of variables where the raw counts might be derived with certainty.

We then modelled the distribution of the randomly rounded counts. Given this model, we were able to compare the disclosure risk with and without suppression, by inferring the distribution of the unrounded raw counts, using the rounded counts.

We were able to identify clear cases where the raw value of cells containing small values could be arithmetically derived. This was done with a brute force attack on a subset of about one million rows of data. Based on this, we estimated that an attacker could break the 2015 confidentiality method, by arithmetically comparing cells with cell totals, in approximately 7% of the rows in the file. We were able to demonstrate that suppressing 1s and 2s as described gives away more information about value distribution to a potential attacker than not suppressing 1s and 2s.

Suppressing cells containing 1s and 2s in this way creates clear opportunities for worse disclosure of sensitive data. RR3 and GRR without suppression gives superior protection. The 2015 method was neither optimal from a customer point of view, nor from a statistical confidentiality point of view.

2.2 Initial results – range of potential confidentiality methods

Method description	*Confidentialisation method*	*Confidentiality rules needs met?*	*Business demography customer needs net?*	*Risks, assumptions, questions*
(1) Method used before 2015	No protection applied (neither perturbation, suppression, nor limits to data access)	No 1s and 2s published	Yes 1s and 2s published	Relied on the *'drive-by'* rule which is inconsistent with the crown law interpretation of Statistics Act S37, and the Confidentiality Standard for Business Collections (2013)
(2) Method used in 2015	RR3 of counts, GRR of count magnitudes, suppression as '(c)' of count 1s and 2s, suppression as '(c)' of count magnitudes based on count 1s and 2s	No count 1s and 2s and magnitudes based on count 1s and 2s are suppressed as '(c)'; on investigation ~7% of these are breakable cases	No count 1s and 2s and magnitudes based on count 1s and 2s are suppressed as '(c)'	The breakable cases are in hindsight inconsistent with Statistics Act S37, and the Confidentiality Standard for Business Collections (2013)
(3) Method proposed by Statistics NZ's confidentiality network	RR3 of counts, GRR of count magnitudes, suppression as 0s of count 1s and 2s, suppression as 0s of count magnitudes based on count 1s and 2s	Yes count 1s and 2s and magnitudes based on count 1s and 2s are suppressed as 0s; however this will explicitly under report	Maybe count 1s and 2s and magnitudes based on count 1s and 2s are suppressed as 0s; and this will explicitly under report	Lower risk and lower utility outcomes than (4), but generally consistent with the Confidentiality Standard for Business Collections (2013)
(4) Method proposed by the BD team statistical methods analyst	RR3 of counts, GRR of count magnitudes	Yes count 1s and 2s are RR3, magnitudes based on count 1s and 2s are GRR; however some rounded count 0s align with rounded magnitude non 0s and potentially lead to breakable cases	Maybe count 1s and 2s are RR3, magnitudes based on count 1s and 2s are GRR	Higher risk and higher utility outcomes than (3), but generally consistent with the Confidentiality Standard for Business Collections (2013)
(5) Method outlined by the chief methodologist	Either (3) or (4) plus P% rule suppression or aggregation	Yes same as (3) or (4) plus P% rule suppression or aggregation would further lower risk and utility	Maybe same as (3) or (4) plus P% rule suppression or aggregation would further lower risk and utility	P% rule required where outputs would 'adversely affect the contributing businesses'

Table 1. Business demography – summary of proposed confidentialisation methods – October 2016

Note 1: The issue of RR3 sparsity (where attackers assuming 0s and 3s are 1s and 2s are likely to be correct) has also been investigated, and this is apparently not a significant issue for the dataset as a whole. For the dataset as a whole, we estimate that rounded 0s were raw 1s and 2s in only 34% of cases, which is good protection; and rounded 3s were raw 1s and 2s in 65% of cases, which is adequate protection. (Arbitrarily, we consider poor protection to be where 80% or 95% or more of rounded values are raw 1s and 2s.)

Note 2: While the term 'sensitivity' is not mentioned in the Statistics Act in relation to confidentiality – e.g. in S37 of the Act – it is a commonly used term in Statistics NZ's current confidentiality related office rules – e.g. those rule developed in accordance with S37(5) of the Act. More information about these office rules is available in Statistics NZ's Confidentiality standards (New Zealand Government, 2013) (Statistics NZ, 2013).

3 How did our understanding of the problem change?

3.1 How did we identify an alternative, better potential solution?

Based on our first phase of investigation, and being that we were simultaneously working on starting up an automated confidentiality service (ACS) work programme, we proposed to put a sixth option on the table, which was the noised counts and magnitudes (NCM) method, where:

- each raw data value set is associated with a fixed, uniformly distributed random number set of identical length
- counts are randomly rounded to base 3 using these fixed random numbers (FRR3)
- each count magnitude contributing value within each count magnitude output value is randomly perturbed with added noise within the following range {-xz%..-x%|+x%..+xz%}, using these fixed random numbers, and the sum output value is rounded with graduated standard rounding
- in brief, this sixth option substantively lies in-between methods 4 and 5.

3.2 What were this method's additional benefits?

The additional benefits were that this method:

- protects both against differencing attacks and monte carlo attacks
- protects count 1s and 2s without suppression
- potentially protects magnitudes against P% breakage attacks in a way that GRR does not do – although we still needed to test this further
- allowed us to test NCM in a real world situation, both in terms of convincing our colleagues it was a good method, and to find out whether its results would better meet our end customers' needs.

3.3 How did we begin turning this potential solution into an actual solution?

The next methodological steps were to test whether this option would protect against P% breakage, and if so, to work out how to implement this method in a timely manner.

We also identified that inter-team communication would be time consuming, but critically important, and essential to get correct (Brooks, 1975). We had a very positive meeting about the proposed NCM method with the BD production team, and we agreed that part of our methodological team would simultaneously move cities for a few days, to work face to face with the BD production team, during method implementation.

4 What were our second phase confidentiality research questions?

4.1 Second phase investigation

The second phase had two objectives. Firstly, we wanted to test select small sets of expected high risk employers, both large and small employers, for count magnitude values breaking the P% rule. Secondly, we wanted to test whether NCM could provide sufficient protection for these breakages, and the level of noise required.

We designed and extracted two expected high risk test sets:

- test set one: {year = 2015, count of employers = 3, count of employees >= 1000, employee count group = 7 [100+ employees]}; n1 = n(test set one) ~ 120
- test set two: {year = 2015, count of employers = 3, employee count group = 2 [1-5 employees]}; n2 = n(test set two) ~ 11100.

4.2 Second phase results

P% perturbation noise level investigated	*count of safe n1*	*count of P% broken n1*	*P% broken percent*	*notes*
control (no perturbation)	40	80	66%	baseline
all round up + x%	100	20	15%	arbitrary tested value 1
all round up + ((y+yz)/2)%	120	0	0%	threshold tested value 2

Table 2. Test set one results.

P% perturbation noise level investigated	*count of safe n2*	*count of P% broken n2*	*P% broken percent*	*Notes*
control (no perturbation)	9400	1700	15%	baseline
all round up + x%	11100	0	0%	arbitrary tested value 1

Table 3. Test set two results.

We tested the P% rule for these two test data sets, and dynamically noised the data to find out what level of noise was required to ensure P% rule breakages were no longer potentially broken, based on the introduction of that noise. We concluded that:

- 66% of the tested selected expected high risk large employers' BD employee count magnitudes initially broke the P% rule
- higher than standard NCM noise {-yz%..-y%|+y%..+yz%} would potentially be effective perturbation, providing P% protection were required
- 15% of the tested selected expected high risk small employers' BD employee count magnitudes initially broke the P% rule
- standard NCM noise {-xz%..-x%|+x%..+xz%} would potentially be effective perturbation, providing P% protection were required
- we could therefore potentially scale the noise from ~+/-x% for small count magnitudes, to ~+/-y% for large count magnitudes.

4.3 What was the decision tree underpinning our final proposed solution?

Question one: Can count of employees data adversely affect the contributing businesses? Answer: Based on feedback from Statistics NZ's legal counsel, the Statistics Act's requirements do not change according to the known or expected level of harm or sensitivity. We can't argue that giving away the *'gist'* or substance of a characteristic of a business regarding its employee numbers is not a problem, but giving away the *'gist'* of a person's income is a problem. We are obliged to assume this answer is yes, and therefore GRR alone is now considered insufficient protection – we must P% protect these count magnitudes as well.

Question two: Can NCM perturbation P% protect these count magnitudes? Answer: Demonstrably yes.

Question three: How should we apply NCM perturbation to these count magnitudes? Answer: Further analysis by the BD production team and part of our methodological team during NCM method implementation identified that noise of ~+/-x% would in fact safely and sufficiently P% protect all 2016 BD count magnitude output values, and that no suppression would be required (Krsinich, 2016).

5 What is our proposed solution?

Statistics NZ is implementing an ‘input perturbation’ approach to confidentialise BD tables. Input perturbation involves adding a small amount of ‘noise’ to the data at the individual business or person level. This is done in such a way that the tables derived from this perturbed data are unbiased, contain as much information as possible, and provide targeted protection to sensitive cells.

We have developed an approach which perturbs both count and magnitude tables. This is the automated confidentiality service (ACS) noised counts and magnitudes (NCM) including fixed random rounding to base 3 (FRR3) method. NCM is being considered for use more widely across Statistics NZ as part of the development of ACS.

In the context of BD data, the respondents whose confidentiality is being protected are businesses. This means that tables of businesses’ employee counts are considered magnitude tables, as the number of employees is a magnitude with respect to the business.

Each business value is assigned a floating point random number uniformly distributed between 0 and 1. These random numbers are fixed to ensure the same degree of perturbation is applied to the business over different outputs.

For business count tables, the business-level random numbers are used to generate a new random number for businesses grouped together in a cell, and this is the basis for FRR3. FRR3 ensures that the same group of businesses will always be rounded the same way in related tables.

These random numbers are also used to generate a ‘noise multiplier’ for the generation of employee count magnitude tables. This noise aggregates to the table level in such a way that it is targeted towards sensitive cells where there is a disclosure risk.

Individual values are protected by at least +/-x%. For the most vulnerable cells with only one business, we guarantee this +/-x% level of uncertainty about the employee count of that business. For cells composed of many businesses, the noise will tend to cancel out. We flag cells with high levels of noise, so analysts treat these with caution.

Suppression of small counts is not required. Application of NCM in some other contexts – for example, census outputs – would require some suppression of small cells in the count tables. This is because there is a small chance of FRR3 being ‘breakable’, for certain combinations of rounded cell values in the interior and margins of the table (Ramsden, 2014), and then for attribute disclosure to be possible from the tabulation variables.

For BD tables, however, there is no risk of disclosing new information from the tabulation variables used to define the business count tables, because this information is already in the public domain. In addition, the original counts cannot be used to help derive the businesses’ employees count magnitude tables, which are already protected by at least +/-x%, even when the exact corresponding counts are known.

Therefore, no suppression is proposed for the BD output tables.

The benefits of this NCM method, compared with the previous confidentiality method, are that:

- more data will be released
- previously breakable cases are no longer released
- related tables will be consistent – each cell value appearing in more than one table will have the same output value each time (Statistics NZ, 2016b).

5.1 What statistical confidentiality literature supports this solution?

The proposed NCM method is supported by a wide corpus of previous statistical confidentiality literature.

Perturbation of one form or another is a preferred confidentiality method for tables of counts when the degree of perturbation is more important than the number of modified cells. In other words, it is recommended for confidentiality protection in tables of counts when smaller changes in many cells are better than bigger changes in fewer cells (Slyuzberg et al, 2007).

Fixed random rounding methods use modulo arithmetic over contributing interior table cells' associated random numbers for rounding, to determine random numbers for rounding for table exterior or margin cells. This method is substantively effective protection both against differencing attacks and monte carlo attacks, and ensures raw counts of 1 or 2 are never output as 1 or 2 (Fraser et al, 2005).

The noise method used to confidentialise count magnitudes is also a very simple method, easy to apply and understand, both for Statistics NZ staff and datalab users who need to confidentialise their outputs to Statistics NZ's standards (Krsinich et al, 2002) (Statistics NZ, 2015a). Potentially, considerable time and effort is required, to develop a magnitude aggregation and suppression strategy for each specific output (Slyuzberg et al, 2007). The noise method avoids this.

Input perturbation of this kind is being used in production by a number of other official statistical agencies. For example, the US Census Bureau uses a 'noise infusion' method to protect longitudinal employment data (Abowd et al, 2012), and the Australian Bureau of Statistics uses noise in the protection of frequency tables accessed via their remote server TableBuilder (Chipperfield et al, 2016) (Statistics NZ, 2016b).

6 Conclusion

The proposed BD confidentiality method is a test case for ongoing work to develop a general automated confidentiality service (ACS). ACS is crucial to ensuring that Statistics NZ improves how it supports both its internal and external customers to get the data they want safely, quickly, flexibly, and cheaply.

ACS should be capable of generating confidentialised tables directly from a dataset, and is crucial to improving how Statistics NZ and potentially other organisations provide customers' data. Its expected components include supported, mandated, and tested core software packages and applications which implement one or more automated confidentiality methods and processes.

ACS should be implementable via different software packages and applications for different groups of users with different use characteristics, however these will typically be either program code packages, which implement automated confidentiality via program code function calls, or web browser based applications, which implement automated confidentiality via a graphic user interface, and which guide customers through a sequence of well-defined steps and options (Statistics NZ, 2016a).

In recent years, statistical authorities have considered that, eventually, automated systems for output confidentialisation may become available (Duncan, et al 2011) (O'Keefe et al, 2013b). In part based on Statistics NZ's substantial and well-regarded expertise at developing and publishing guidance for *manual* systems for output confidentialisation, Statistics NZ is now genuinely seeking to incrementally craft this guidance into *automated* systems for output confidentialisation (Statistics NZ, 2015a) (Statistics NZ, 2016a).

Currently planned implementations include the BD product discussed in this paper to be delivered around November 2016, an IDI test product to be delivered around December 2016, a NZ Census test product to be delivered around March-June 2017, and a NZ Census final product to be delivered around November 2017-February 2018 (Statistics NZ, 2016a).

Statistics NZ also plans to review its Confidentiality Standard for Business Collections, based on the methodological analysis and confidentiality standards thinking underpinning this paper.

Acknowledgements

The authors are grateful for and acknowledge the assistance of our peers and colleagues, especially Aaron Beck, Antony Gomez, Frances Krsinich, and Mary Reid, among others, in the development of this paper, and the business demography confidentiality methodology, and official business demography statistics, which underpin this paper.

7 References

Abowd, J.M., Gittings K.K., McKinney, K., Stephens, B., Vilhuber, L., & Woodcock, S. 2012. *Dynamically consistent noise infusion and partially synthetic data as confidentiality protection measures for related time series*. http://dx.doi.org/10.2139/ssrn.2159800.

Brooks, F. 1975. *The Mythical Man-Month*. https://archive.org/stream/mythicalmanmonth00fred/mythicalmanmonth00fred_djvu.txt.

Chipperfield, J. 2014. *Disclosure-protected inference with linked microdata using a remote analysis server*. http://ro.uow.edu.au/cgi/viewcontent.cgi?article=4468&context=eispapers.

Chipperfield, J., Gow, D., & Loong, B. 2016. *The Australian Bureau of Statistics and releasing frequency tables via a remote server*. http://content.iospress.com/download/statistical-journal-of-the-iaos/sji969?id=statistical-journal-of-the-iaos%2Fsji969.

Duncan, G.T., Elliot, M. & Salazar-González, J.-J. 2011. *Statistical confidentiality. Principles and practice*. http://link.springer.com/book/10.1007%2F978-1-4419-7802-8.

Fraser, B., & Wooten, J. 2005. *A proposed method for confidentialising tabular output to protect against differencing*. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2005/wp.35.e.pdf.

Krsinich, F.,& Piesse, A. 2002. *Multiplicative microdata noise for confidentialising tables of business data*.http://www.stats.govt.nz/~/media/Statistics/browse-categories/business/business-character/multiplicative-microdata-noise-bus-data/mmnconbusdata.pdf.

Krsinich, F. 2016. *Confidentialising Business Demography tables using the noise for counts and magnitudes (NCM) method*.

New Zealand Government. 2013. *Statistics Act 1975*. http://www.legislation.govt.nz/act/public/1975/0001/latest/DLM430705.html.

O'Keefe, C.M., & Chipperfield, J. 2013a. *A summary of attack methods and confidentiality protection measures for fully automated remote analysis systems*. http://onlinelibrary.wiley.com/doi/10.1111/insr.12021/abstract & https://publications.csiro.au/rpr/pub?pid=csiro:EP123152.

O'Keefe, C.M., Westcott, M., Ickowicz, A., O'Sullivan, M., & Churches, T. 2013b. *Protecting confidentiality in statistical analysis outputs from a virtual data centre*. https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_1_Okeefe_Rev.pdf.

Open Data NZ. 2016a. *New Zealand Data and Information Management Principles*. https://www.ict.govt.nz/guidance-and-resources/open-government/new-zealand-data-and-information-management-principles/.

Open Data NZ. 2016b. *Open Government Information and Data Programme*. https://www.ict.govt.nz/programmes-and-initiatives/open-and-transparent-government/open-government-information-and-data-work-programm/.

Ramsden, A. 2014. *Developing, confidentialising, and publishing more granular NZ serious injury outcome indicator data*. https://secure.orsnz.org.nz/conf48/program/Papers/nzsaorsnz2014_paper_5.pdf.

Slyuzberg, M., Irving, J., Buckley, G., Camden, M., & Sullivan, T. 2007. *Optimising confidentiality in administrative data tabular data: a comparison of methods and techniques*. http://www.statisphere.govt.nz/~/media/Statistics/about-us/statisphere/Files/official-statistics-research-series/osr-series-v2-2007-optimising-confidentiality-in-admin-tab-data.pdf.

Statistics NZ. 2013. *Methodological standard for Confidentiality in Business Collections (Version 1.0.0)*.

Statistics NZ. 2015a. *Microdata output guide (Third edition)*. www.stats.govt.nz/~/media/Statistics/services/microdata-access/data-lab/Microdata-output-guide-2015.pdf.

Statistics NZ. 2015b. *Privacy, security, and confidentiality of information supplied to Statistics NZ*. http://www.stats.govt.nz/about_us/legisln-policies-protocols/confidentiality-of-info-supplied-to-snz/safeguarding-confidentiality.aspx.

Statistics NZ. 2016a. *Automated Confidentiality Service (ACS) roadmap version 1.4*.

Statistics NZ. 2016b. *Introducing new method for confidentialising business demography tables*. http://www.stats.govt.nz/browse_for_stats/businesses/business_characteristics/new-method-for-confidentialising-tables.aspx.

Statistics NZ. 2016c. *NZ.Stat DataHub table viewer*. http://nzdotstat.stats.govt.nz/wbos/Index.aspx.

Statistics NZ. 2016d. Statistics New Zealand's Strategic intentions for the period 2016/17–19/20 Annual report for the year ended 30 June 2016. http://www.stats.govt.nz/about_us/what-we-do/our-publications/annual-reports/annual-report-2016.aspx.

8 Copyright and Citation Information

8.1 Crown Copyright

8.2 Disclaimer

This research paper represents the views of the authors. It does not necessarily represent the views of Statistics NZ and does not imply commitment by Statistics NZ to adopt any findings, methodologies, or recommendations. Any data analysis was carried out under the security and confidentiality provisions of the Statistics Act 1975.

8.3 Liability statement

While all care and diligence has been used in processing, analysing, and extracting data and information in this publication, Statistics New Zealand gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

8.4 Citation

Ramsden, A., & Moses, C. 2016. *Redeveloping the confidentiality method for Statistics New Zealand business demography data*.