sciendo

Journal of Official Statistics, Vol. 35, No. 2, 2019, pp. 319–336, http://dx.doi.org/10.2478/JOS-2019-0015

Prospects for Protecting Business Microdata when Releasing Population Totals via a Remote Server

James Chipperfield[1], *John Newman*[1], *Gwenda Thompson*[1], *Yue Ma*[2], *and Yan-Xia Lin*[2]

Many statistical agencies face the challenge of maintaining the confidentiality of respondents while providing as much analytical value as possible from their data. Datasets relating to businesses present particular difficulties because they are likely to contain information about large enterprises that dominate industries and may be more easily identified. Agencies therefore tend to take a cautious approach to releasing business data (e.g., trusted access, remote access and synthetic data). The Australian Bureau of Statistics has developed a remote server, called TableBuilder, which has the capability to allow users to specify and request tables created from business microdata. The tables are confidentialised automatically by perturbing cell values, and the results are returned quickly to the users. The perturbation method is designed to protect against attacks, which are attempts to undo the confidentialisation, such as the well-known differencing attack. This paper considers the risk and utility trade-off when releasing three Australian Bureau of Statistics business collections via its TableBuilder product.

Key words: Business data; online access; perturbation; remote server; statistical disclosure control.

1. Introduction

While carrying out their role of collecting and disseminating data, statistical agencies generally need to determine effective ways of meeting two key objectives: to maintain the confidentiality of respondents and to provide its society with as much analytical value from the data as possible. The two most common types of data that require confidentialisation are data about people and data about businesses. Person-level data and business-level data have many aspects in common. However, there are some characteristics commonly associated with business data that may make confidentialising a more challenging problem.

Typically, some industries will be dominated by large businesses whose information is difficult to conceal without suppressing or altering the data to a large extent. For many business collections, continuous data items, such as turnover or profit, are of key interest to users. Some of these continuous data items, such as capital expenditure, may have many zero values and a few large values. Certain aspects of a business's operations can become

[1] Australian Bureau of Statistics, P.O. Box 10, Belconnen, Australian Capital Territory 2616, Australia. Emails: james.chipperfield@abs.gov.au, john.newman@abs.gov.au, and gwenda.thompson@abs.gov.au

[2] University of Wollongong, Wollongong, New South Wales 2522, Australia. Emails: mayue3588@gmail.com and yanxia@uow.edu.au

Acknowledgments: The views in this article do not necessarily represent those of the Australian Bureau of Statistics.

public knowledge, for example through the release of annual reports. Some users may also have access to administrative data related to the businesses. There are potentially high incentives for attackers to try to discover confidential information about businesses because this may lead to a competitive advantage. These issues can become even more problematic in countries with smaller economies, because of the limited number of businesses that operate in those countries.

For some statistical agencies, there are legislative differences between the treatment of person-level data and business-level data. For example, there can be opportunities to gain consent to publish business data. This may allow the release of more data, but can also make the process of applying confidentiality protection more complex. This is because there is a need to monitor which businesses provide consent and because confidentialisation is complicated in cases when consenting and nonconsenting businesses appear in the same cell of a table. Another example is where confidentialisation is required at multiple levels of business structure.

For the reasons listed above, release of detailed business micro-data by statistical agencies may allow attackers to discover confidential information about a business. This is why, at least in the case of the Australian Bureau of Statistics (ABS), the vast majority of its business data is still released in the form of broad-level tables.

A common approach for confidentialising tabular business data is suppression (González 2005; Tambay and Fillion 2013). This approach is easiest to apply when the full set of tables to be published is known in advance. If additional tables are requested, then any further suppression will need to take into account which cells were previously suppressed. In practice the full set of tables is rarely known in advance. This means that suppression is unlikely to be optimal and that the amount of information released with each set of additional tables will be increasingly suppressed. Some statistical agencies consider alternative approaches including accredited or “trusted” access (Abrahams and Mahony 2008), replacing sensitive data with synthetic data (Miranda and Vilhuber 2013) and various ways of perturbing micro-data.

The ABS has developed an approach that could be used to release confidentialised totals from business data through its remote server, called TableBuilder. A simple model for a remote server (Chipperfield and O’Keefe 2014; Chipperfield 2014; O’Keefe and Chipperfield 2013; Thompson et al. 2013) is: (1) an analyst submits a query (i.e., request for a table) to the remote server; (2) the remote server automatically modifies or restricts the query’s output; (3) the server sends the modified output to the analyst. Tambay (2017) use the ideas of a remote server (specifically TableBuilder) while also perturbing the underlying micro-data. For reviews of remote servers in use or in development in national statistical agencies, see Lucero et al. (2011), Reuter and Museux (2010).

There are some key advantages of a remote server. First, the degree to which an estimate is modified depends upon the output itself. For example, modification of an estimate may be relatively high if a cell is dominated by a single business and relatively low if a table cell has many small businesses of roughly equal size. Second, because an analyst is restricted from viewing the micro-data, less modification is needed than would otherwise be the case. Third, it allows users to gain rapid access to estimates they request. Fourth, the modification algorithm assures a specified level of protection is guaranteed.

This article evaluates the prospect of allowing access to business survey data via TableBuilder. For a full description, see Part 1 of Thompson et al. (2013). Section 3 defines disclosure and utility and discusses how TableBuilder's perturbation settings could be chosen to optimise the trade-off between disclosure and utility. Sections 4, 5 and 6 evaluate utility of TableBuilder outputs, conditional on a certain level of disclosure risk, for two surveys and one administrative collection of the ABS. Section 7 makes some concluding remarks, including a discussion of the prospects of releasing business data in TableBuilder.

2. TableBuilder

2.1. Totals

Here we describe the essential perturbation algorithm, but for a more complete description see Part 1 of Thompson et al. (2013). Consider any particular cell in a table and let there be n sample units contributing to the cell, where the units are indexed by $i = 1, 2, \ldots, n$. Define a continuous valued characteristic (e.g., income or turnover) for the ith unit (e.g., business) by y_i so that $|y_1| \geq |y_2| \geq |y_3| \ldots \geq |y_n|$. The absolute values are taken because it is the magnitude of y, not whether it is positive or negative, that has bearing on considerations of risk and utility. (Changing all y values from positive to negative in a cell would not affect the perturbation distribution P^*- this is as it should be because a large negative y value is just as sensitive as a large positive y value.) If we define the estimation weight for the ith unit in the cell by w_i the survey estimate of the total is $\hat{Y} = \sum_i w_i y_i$. Instead of releasing $\hat{Y}$, TableBuilder releases $\hat{Y}^* = \hat{Y} + P^*$, where $P^* = \sum_{i=1}^{K} \left(m_i d_i^* h_i^*\right) y_i w_i$ is the perturbation amount and:

- m_i is a positive constant parameter. This parameter moderates the magnitude of the perturbation relative to the value y_i. In particular, the parameter $m_1 (i = 1)$ is the most important of the parameters as it plays an important role in protecting the largest contributor's value, y_1. The optimal value of m_i depends upon the distribution of y within the cell, the risk measure and the utility measures. Given the complexity of these dependencies, the optimal values are calculated in simulation (see Subsection 3.4).
- d_i^* is a random variable taking the value -1 and 1 with equal probability and so determines the direction of the perturbation.
- K is the number of top contributors in the cell that are used in calculation of P^*. We found that there was little value in allowing $K > 4$ since the main aim here is to protect the largest contributor's value (see Subsection 3.1).
- h_i^*, for purposes of this evaluation, was a random value drawn from a symmetric triangular distribution with lower limit $1 - b = 0.7$ to $1 + b = 1.3$ and the mode of 1. h_i^* has mean $E\left(h_i^*\right) = 1$ and variance $Var\left(h_i^*\right) = b^2/12$ which, with $b = 0.3$, is equal to 0.075. The bimodal distribution generated by $d_i^* h_i^*$ is symmetric round zero, $\sigma_*^2 = Var\left(d_i^* h_i^*\right) = 1 + b^2/12$ and has little mass around 0. This avoids unacceptable small values while also ensuring that the perturbation has mean zero. Exploring other distributions would likely be a fruitful line of research (Krsinich and Piesse 2002; Evans et al. 1998; Tambay 2017).

The form of P^* is intuitive in the sense that its magnitude is in proportion to the size of the K largest, and most at risk, contributors. Allowing $K > 1$ allows more degrees of freedom to specify the perturbation distribution, P^*, and so will allow it to better approximate the optimum distribution.

There is no constraint in this procedure to ensure consistency between the perturbed estimates. This means a perturbed total for Australia will not exactly equal the perturbed totals for each state summed over all states.

The current functionality of TableBuilder is such that K, the m_is and the distributions of d_i^* and h_i^* are essentially fixed for a given business collection. This means, for instance, that it is not possible to allow P^* to depend upon whether or not a cell is known to be sensitive and that it is not possible to allow the value of K to vary across cells. In Section 7 we discuss the benefits of relaxing this constraint.

We can see that $E^*(P^*) = 0$, where '$*$' represents the perturbation process. We did consider generating P^* from the Laplace distribution so as to achieve ε-differential privacy (Dwork et al. 2006), but the utility loss was far too great.

Table 1 gives an example of the perturbation of a cell total. We set $K = 4$, $\mathbf{m} = (m_1, m_2, m_3, m_4) = (0.6, 0.4, 0.3, 0.2)$ and there are $n = 8$ businesses in this cell. The estimator of total $\hat{Y}$ = USD 263,719 is perturbed by $P^* = -$ USD 18,278 so that the released estimate is $\hat{Y}^*$ = USD 245,441. The particular choice of values for K and $\mathbf{m}$ in Table 1 are for illustration only.

A Unit Key is a positive integer less than 2^{32} that is permanently and randomly assigned to each unit. The Unit Key is the random seed used to generate the value of d_i^* for $i = 1, \ldots, n$. This means, once generated, all d_i^*s are fixed in all calculations of P^*. It also means that a unit's contribution to P^*, when applicable, is either always positive or always negative – this was to reduce the perturbation variance of differences between cell totals, where the cells had some units in common (for more discussion on this see Subsection 3.4 and in Section 7).

A Cell Key is calculated by summing the Unit Keys for all the units contributing to the cell and then dividing by a large prime number. This essentially means that the Cell Key depends upon the exact set of n records that belong to the cell. The random seed for h_i^*

Table 1. Example of perturbation (K = 4).

	Turnover	Estimation				Weighted	Perturbation
	(USD)	weight	Magnitude	Direction	Noise	turnover	amount
ID	y_i	w_i	m_i	d_i^*	h_i^*	y_iw_i	$P_i^* = m_id_i^*h_i^*y_iw_i$
1	72.1	458.2	0.6	1	0.95	33,036	18,831
2	65.3	185.7	0.4	− 1	1.02	12,126	− 4,947
3	65.3	752.7	0.3	− 1	1.54	49,151	− 22,708
4	50.1	612.6	0.2	− 1	1.54	30,691	− 9,453
5	49.2	977.5				48,093	
6	45.4	458.7				20,825	
7	36.9	896.3				33,073	
8	36.9	995.2				36,723	
Sum						$\hat{Y}$ = 263,719	$P^* = -$ 18,278

depends upon the Unit Key for unit *i and* its associated Cell Key – this means adding a unit to a cell will generate a new and independent value of h_i^* for each unit in the cell.

It follows that P^* and so $\hat{Y}^*$ will take the same value for any cell containing the same set of n units,- that is, TableBuilder will release the same estimate for logically equivalent cells because they will have the same set of contributors. This means that it is not possible to average over the effect of perturbation by requesting the same logically defined cell count in different tables.

Cells with a small number, say fewer than H, of contributing businesses are typically suppressed in the publications of many statistical agencies. TableBuilder effectively does the same thing. If $n \leq H$ then $\hat{Y}^* = 0$. If $H = 2$ then this provides protection against attacks on cells with sample counts of '1' or '2'.

The estimate of the count of units in the population belonging to a cell is $\hat{N} = \sum_{i=1}^{n} w_i$. Instead of releasing the ratio, $\hat{T} = \hat{Y}/\hat{N}$, TableBuilder releases $\hat{T}^* = \hat{Y}^*/\hat{N}$. (We do not discuss the perturbation of $\hat{N}$ here. For details see Chipperfield et al. 2016.)

2.2. *Confidence Intervals*

It is straightforward to derive 95% confidence intervals around each cell estimate, $\hat{Y}^*$. It would be slightly more difficult to derive confidence intervals for a linear combination (e.g., the difference) of perturbed cell estimates. TableBuilder does not release confidence intervals. Instead TableBuilder releases the variance of $\hat{Y}^*$, given by

$$\sigma^2 = Var_{s^*}(\hat{Y}^*) = Var_s[E_*(\hat{Y}^* \mid s)] + E_s[Var_*(\hat{Y}^* \mid s)],$$

where the first term represents the variation due to the sampling process, denoted by s, and the second term is the variation due to perturbation process, denoted by '$*$'. TableBuilder estimates the first term using the standard Jackknife. TableBuilder calculates the second term by $\sum_{i=1}^{K} m_i^2 \sigma_*^2 w_i^2 y_i^2$, where σ_*^2 is defined earlier. As mentioned, the variance cannot be used to construct 95% confidence intervals confidence intervals using ($\pm 1.96\sigma^2$) because the perturbations are not approximately normally distributed. As the ratio of the perturbation variance to the sampling variance increases, the more inaccurate the coverage rates based on the normality assumption would become. One option would be to suppress a cell estimate if this ratio is above some threshold value. This is a topic for further research.

3. Measuring Disclosure Risk and Utility

3.1. *Attack Scenarios*

We measure the disclosure risk with respect to three 'attack scenarios'. In each scenario, the target is the largest contributor to the cell ($i = 1$), the target value is therefore y_1, and the attacker knows that the weight of the largest and second largest contributors is equal to one ($w_1 = w_2 = 1$). The largest contributor is chosen to be the target because it, of all the units in the cell, has the highest associated risk of disclosure. In Scenario 1 and 2, the attacker does not know the value of y for any of the contributors to the cell. However, in Scenario 3, the attacker is the second largest contributor to the cell ($i = 2$) and so is able to use its known contribution, y_2, to improve upon the accuracy of Attack 1. This means that the disclosure risk of Scenario 3 will always be at least as high as Scenario 1.

Attack Scenario 1: The value of $\hat{Y}^*$ is used as an estimate of y_1. The attacker's estimate of y_1 under this scenario is $\hat{y}_1^{(1)} = \hat{Y}^*$

Attack Scenario 2: The attacker uses the difference between two cell estimates, $\hat{y}_1^{(2)} = \hat{Y}^* - \hat{Y}^*(i = 1)$ as an estimate of y_1, where $\hat{Y}^*(i = 1)$ is the same as $\hat{Y}^*$ except that the largest contributor, $i = 1$, is dropped from the cell.

Attack Scenario 3: This is the same as Attack Scenario 1 except that the attacker is also the second highest contributor to the cell $(i = 2)$. The attacker can use its known contribution, y_2, to improve its estimate of y_1. The estimate of y_1 under this scenario is $\hat{y}_1^{(3)} = \hat{Y}^* - y_2$.

Scenario 2 is an example of a differencing attack. Differencing attacks can be effective because any two tables on their own may have low disclosure risk but, when differenced, may have a high disclosure risk. They can be particularly effective when used via a remote server since, at least in the case of TableBuilder, the attacker is relatively free to request tables of their choice (Thompson et al. 2013). More detailed discussions about differencing attacks using remote servers can be found in O'Keefe and Chipperfield (2013).

In order for an attack to succeed the attacker needs:

1. To know that the target is in the sample. It is well known that statistical agencies typically select large businesses with a higher probability than smaller businesses. For smaller businesses, sampling may provide some protection since an attacker will not know if a particular business is selected in the sample. Since the underlying micro-data are not observed, it would be necessary to conduct a series of attacks in order to confirm whether or not a small business is actually in the sample (Chipperfield and O'Keefe 2014).
2. In the case of Attack Scenario 2, to know how to uniquely identify the target in terms of a set of quasi-identifiers. This allows the attacker to "drop" the target business from the cell in a table. To conduct Attack Scenario 1 and 3, the business does not have to be uniquely identified, often referred to as *identification*, only that the target business dominates the cell.
3. To circumvent TableBuilder's confidentiality protections and disclosing an attribute of the business.

TableBuilder gives users a high degree of flexibility in choosing a table's dimensions and scope. There is often considerable information about large businesses in the public domain which may in turn make identification likely (e.g., there may only be one private hospital in a small area). Accordingly, we conservatively assume that 1. and 2. occur with certainty. So for large businesses at least, the only protection available in TableBuilder is perturbation. Consequently, perturbation is the focus of how disclosure risk is measured (see Subsection 3.2).

3.2. *What is Disclosure*

In many organisations, disclosure is considered to occur for a business if published estimates can be used to accurately infer an attribute (e.g., total turnover) of a business.

It is not necessary for the attribute to be inferred exactly – the degree of (or threshold for) accuracy required for disclosure must be determined by the Statistical Agency.

We say that the disclosure risk from Attack Scenario s, τ_s, is acceptable if the probability that the estimate $\hat{y}_1^{(s)}$ is within $V_s\%$ of the true value y_1 is less than R_s. This is a stochastic generalisation of the P% Rule (Tambay and Fillion, 2013). Therefore the disclosure risk from Attack s, τ_s, is acceptable if,

$$\tau_s = \Pr\left(\frac{\left|\hat{y}_1^{(s)} - y_1\right|}{y_1} \leq V_s\%\right) \leq R_s \tag{1}$$

We can say that for attack s, V_s is the threshold value that draws the line between what does and what does not constitute a disclosure and R_s is the *acceptable disclosure risk*. Different values of (R_s,V_s) in different scenarios could be justified on the basis of whether the attack scenario is likely to occur in the first place (e.g., level of sophistication and prior knowledge required to carry out the attack) and the level of the business structure (e.g., business, enterprise, employee) that is attacked.

To illustrate the rule, consider the following example. Consider three businesses in a cell that have weights of 1 and Income USD 98, USD 1 and USD 1. Following Attack 1, a user could guess that the Income of the largest contributor is equal to the cell estimate of USD 100. This guess would be wrong by only 2% (USD 100–USD 98)/USD 98. TableBuilder would not release the unperturbed estimate of USD 100; it would instead release a perturbed estimate. Consider if the possible perturbed estimates (each equally likely) were 60, 70, 100, 130, 140, 150, 160. Again following Attack 1, if a user now guessed that the Income of the largest respondent (USD 98) is equal to the cell's perturbed estimate, the guesses would be wrong by -39, -29%, -2%, 33%, 43%, 53%, and 63%. The guesses using perturbed estimates would be within 18% about 15% of the time. The risk associated with Attack 1 would be acceptable if disclosure and the disclosure risk were $V_1 = 18$ and $R_1 = 0.15$, respectively.

3.3. *Defining Utility Loss*

We measure utility loss associated with the perturbed estimate $\hat{Y}^*$ by

$$L = |P^*|/\hat{Y}. \tag{2}$$

The magnitude of the perturbation, $|P^*|$, depends upon K and the 'magnitude values' m_i for $i = 1, \ldots, K$. The utility loss measure is also used by Yancey et al. (2002) in assessing utility loss of a sample mean. There are other useful measures of utility loss, including the mean square error and the mean absolute error (Domingo-Ferrer and Torra 2001).

3.4. *Optimal Magnitude Values*

The optimal value of $\mathbf{m}$ minimises L, given by (2), subject to the constraint given by (1) for $s = 1$, 2, and 3, where $(R_1,V_1) = (0.15,18)$, $(R_2,V_2) = (0.15, 11)$ and $(R_3,V_3) = (0.15,11)$. That is, the optimal value of $\mathbf{m}$ minimises utility loss subject to having an acceptable disclosure risk from Attacks 1, 2 and 3. Below we describe how the optimal values of $\mathbf{m}$ were obtained.

It is important to note that the *scale* of the distribution of y in the cell does not affect the optimal value of $\mathbf{m}$ – what is important is the relative size of y for the contributors to the cell. Table 2 shows a variety of distributions for y. The distributions are made up of between two and four units (other units could well belong to cell but, if they do, we assume they make a negligible contribution). For each of these distributions, we had to choose a value of K. We found limited additional benefit from allowing $K = 4$ and so we decided to set $K = 3$ for all distributions. This means where a distribution was made up of four units, only the top three contributing units were used in calculating the perturbation, P^*. The exception to this was Distribution 5, which was made up of only two contributors (of relative size 60 and 40), and so we set $K = 2$.

For each distribution of y, Table 2 gives the optimal value of $\mathbf{m}$ for $K = 3$. For a given distribution of y, the optimal value was found by:

(i) measuring the average value of L and measuring the disclosure risks, τ_s for $s = 1, 2$ and 3, for a range of different values of $\mathbf{m}$. For a given value of $\mathbf{m}$, these measures were calculated by simulating the perturbation distribution, P^*, 500 times and conducting Attacks 1, 2 and 3.

(ii) identifying the value of $\mathbf{m}$ from (i) that minimised L while also meeting the constraint on the risk from Attacks 1, 2 and 3 as described in the first sentence of Subection 3.4.

Table 2 shows that, as the distribution of y becomes more uniform, the optimal values in the vector $\mathbf{m}$ increase in size.

Figure 1 illustrates the risk-utility trade-off with respect to only Attack 2. *Utility Loss* is the average value of L and *Risk* $= \tau_2$ is the proportion of times condition (1) was met from Attack 2, over 500 independently generated values of P^*. Figure 1 plots *Utility Loss* by *Risk* for Attack Scenario 2 for a range of values of $\mathbf{m}$ and for two disclosure thresholds ($V_2 = 11$, 18). Recall that $V_2 = 11$ means that disclosure occurs when $\hat{y}_1^{(2)}$ is within 11% of y_1.

Table 2. Magnitude values that meet constraints on the disclosure risk and maximise utility for different contributor values.*

Distribution number	Distribution of y (relative size of contributors)				Optimal values (K = 3)		
	1st	2nd	3rd	4th	m_1	m_2	m_3
1	90	5	5	0	0.15	0.1	0.1
2	80	10	5	5	0.15	0.1	0.1
3	70	20	10	0	0.15	0.1	0.1
4	60	20	10	10	0.2	0.1	0.1
5	60	40	0	0	0.25	0.15	n/a
6	50	20	20	10	0.25	0.15	0.1
7	40	30	30	0	0.3	0.2	0.1
8	30	30	30	10	0.4	0.3	0.2
9	25	25	25	25	0.5	0.4	0.3

*The constraints on the disclosure risk that are imposed by Attacks 1, 2 and 3 are $(R_1, V_1) = (0.15, 18)$, $(R_2, V_2) = (0.15, 11)$ and $(R_3, V_3) = (0.15, 11)$.

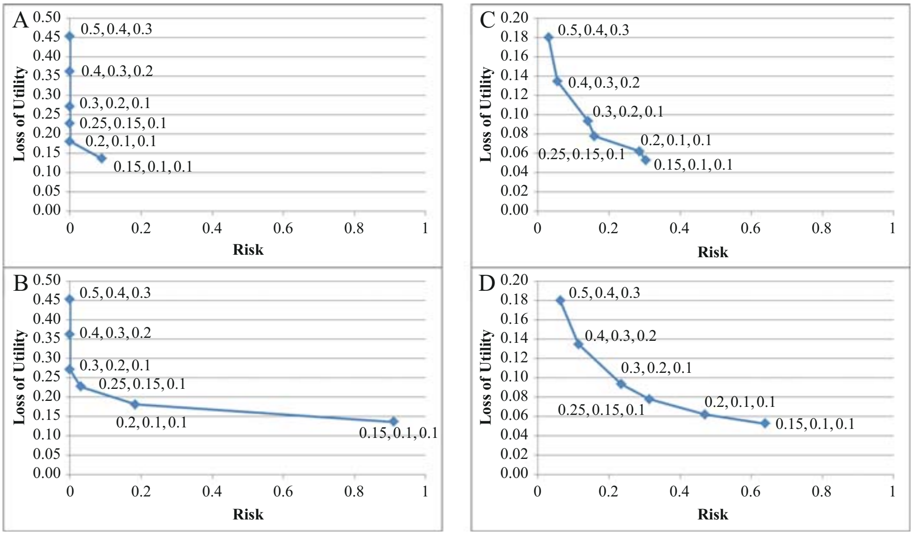

Fig. 1. Risk vs utility loss under attack scenario 2. Figures A and B have y values of (90, 5, 5, 0), Figures C and D have y values of (30, 30, 30, 10). Figures A and C define disclosure by $V_2 = 11$, Figures B and D define disclosure by $V_2 = 18$.

Figure 1A shows if the (relative) y values were (90, 5, 5), $\mathbf{m} = (0.15, 0.1, 0.1)$ and the disclosure threshold was $V_2 = 11$, that *Utility Loss* = 13% and *Risk* = 10%. Figure 1B shows that if disclosure was instead defined by $V_2 = 18$, *Risk* would ris dramatically to 90%.

Figure 1C shows that if the relative y values were (30, 30, 30, 10), $\mathbf{m} = (0.15, 0.1, 0.1)$ and the disclosure threshold $V_2 = 11$ that *Utility Loss* = 5% and *Risk* = 30%. Figure 1D shows that if the disclosure threshold was instead $V_2 = 18$ that *Risk* = 64%.

Ideally, TableBuilder would allow the choice of $\mathbf{m}$ to depend upon on the *actual* distribution of y in each cell (as per Table 2). As TableBuilder does not have this capability, we must choose a single value of $\mathbf{m}$ that guarantees an acceptable disclosure risk for *all* distributions of y in Table 2. The resulting optimal value would be $\mathbf{m} = (0.5, 0.4, 0.3)$. However, we did not use $\mathbf{m} = (0.5, 0.4, 0.3)$ because the utility loss was too high. The compromise value of $\mathbf{m} = (0.4, 0.3, 0.2)$, that we used in all empirical studies below, does not strictly have an acceptable disclosure risk for Distribution 9 in Table 2. (Note: because the disclosure risk is somewhat contextually free in the way it is defined here, we focus on measuring utility loss in the empirical studies).

Work on the optimal distribution for d_i^* is currently being investigated by some of the authors of this article. Consider using $q^* d_i^*$ instead of d_i^*, where $q^* = 1$ if n is odd and is equal to -1 if n is even. This change would reduce the disclosure risk of a differencing attack (Attack Scenario 2) while having no effect on the disclosure risk for other attacks. We see that, for any two cells that differ by a single target unit, the direction of the perturbation, $q^* d_i^*$, will be positive for one of the cells and will be negative for the other cell. This change would increase the perturbation variance of the difference, $\hat{Y}^* - \hat{Y}^*(i = 1)$, while having no impact on the perturbation variance of the individual totals, $\hat{Y}^*(i = 1)$ or $\hat{Y}^*$.

4. Evaluation of Employees Earnings and Hours

Employee Earnings and Hours (EEH) is a two-yearly survey of employing organisations in Australia. EEH uses a two-stage sample selection approach. The first stage involves selecting a probability sample of employer units, from the ABS Business Register. The statistical unit for the first stage comprises all activities of an employer in a particular state or territory based on the Type of Activity Unit (TAU). The sampling unit for the second stage is employee. Employees are in scope of the second stage selection if they earned pay during the reference period. Data collected in the survey are used to estimate the composition and distribution of average weekly earnings, hours worked, and the methods of setting pay (e.g., award only, collective agreement, and individual agreement). EEH currently applies suppression to protect respondents against disclosure where a 'respondent' can be an employee, TAU, or at the highest level of Enterprise Group.

4.1. Utility at Employee Level

Tables 3 and 4 summarise the utility loss resulting from perturbing estimates with $\mathbf{m} = (0.4, 0.3, 0.2)$. Here we measure the utility of typical EEH estimates.

Table 3 shows, in most cases, that perturbation changes the estimates by less than 1%. When perturbation makes larger changes (6–7%) to estimates, the associated sampling errors (not provided in Table 3) are also high, due to small sample sizes. For example, the estimate for Community and Personal Service Workers in Owner Manager of Incorporated Enterprises was perturbed by 7.3% and has a Relative Standard Error (RSE) in the range 25–50%.

Table 4 shows, again, that the percentage impact of perturbation is often less than 1%. As in Table 3, the larger differences are for estimates with RSEs between 25% and 50% (RSEs not provided in Table 4). For example, on the one hand, the estimate for Mining and Award Only is perturbed by−9.7% and has a standard error of 10–25%, whereas the estimate for Manufacturing and Award Only is perturbed by only 0.1% and has a standard error of 5%. Since these changes are significantly less than the RSEs the loss of utility would be minimal. Feedback from users of the EEH is that this level of utility loss is acceptable.

4.2. Protection of TAUs

The three attack scenarios are also possible at the TAU level. TableBuilder does not recognise the TAU hierarchy in any way and so its perturbation settings cannot manage disclosure risk at the TAU level. For example, TableBuilder does not recognise if all employees in a cell belong to one TAU. The question is whether, nevertheless, there is acceptable disclosure risk at the TAU level given perturbation is only designed to have acceptable disclosure risk at the employee level.

To illustrate, Table 5 summarises the data collected from a realistic but hypothetical sample of 25 employees who were themselves selected from three different TAUs and who all belong to a single cell of a table. We assume the attacker knows that the cell contains only the three selected TAUs and that the TAUs were selected with certainty. The inverse of the within-TAU employee sampling fraction is used to weight its sample of employees, thus giving the TAU contribution to the cell estimate. Table 5 shows the

Table 3. Estimates of average weekly total cash earnings (USD) by occupation and agreement type and percentage impact of perturbation (%).

Occupation	Award only (%)	Collective agreement (%)	Individual arrangement (%)	Owner incorporated enterprise (%)	All (%)
Managers	1127.1 (1.1)	1982.7 (−0.1)	2083.2 (0.2)	1304.2 (−2.0)	1928.9 (0.1)
Professionals	1147.7 (−0.9)	1383.9 (−0.3)	1507.5 (−0.5)	2036.8 (−0.3)	1436 (−0.1)
Technicians and trades workers	695.8 (−0.6)	1544.4 (−0.2)	1272.5 (0.6)	1219.6 (3.4)	1250.2 (0.3)
Community and personal service	546.5 (0.9)	841.2 (0.5)	587.8 (0.3)	932.1 (7.3)	709.1 (0.3)
Clerical and adminstration	708.4 (0.3)	1063.3 (−0.2)	962 (−0.2)	831.9 (−3.9)	970.2 (−0.1)
Sales workers	419 (−0.7)	485.1 (−0.1)	935.9 (−0.1)	949.7 (0.7)	606.7 (0.0)
Machinery operators	863.8 (0.2)	1509.7 (0.1)	1154.9 (−0.2)	1095.9 (6.7)	1284.7 (0.1)
Labourers	496.5 (0.8)	958.5 (1.3)	780.9 (−1.2)	1321.4 (6.5)	784.3 (0.7)
All	632.7 (−0.2)	1150.7 (0.0)	1277.9 (0.1)	1325.6 (−0.9)	1122.8 (0.0)

Table 4. Estimates of average weekly total cash earnings (USD) by industry and agreement type and percentage Impact of perturbation (%).

Occupation	Award only (%)	Collective agreement (%)	Individual arrangement (%)	Owner incorporated enterprise (%)	All (%)
Mining	1285.3 (−9.7)	2237 (0.7)	2538.6 (0.2)	1639 (−13.3)	2388.8 (0.0)
Manufacturing	614.6 (0.1)	1292.8 (−0.9)	1307.7 (1.1)	1270 (4.8)	1221.7 (0.3)
Electricity, gas, water and waste services	917.9 (0)	1745.4 (−0.3)	1851.5 (−0.5)	966.9 (8.8)	1735.3 (−0.3)
Construction	811.6 (−2.4)	2110.4 (−0.2)	1333.2 (−0.1)	1086.8 (1.6)	1440.2 (0.0)
Wholesale trade	706.2 (2.7)	1110.3 (−0.4)	1352.4 (0.1)	1152.3 (−1.9)	1258.5 (0.0)
Retail trade	479.4 (0.8)	489 (−0.3)	965 (−0.8)	992.6 (−2.5)	640.2 (−0.5)
Accommodation and food services	477.6 (0.5)	398.9 (0.2)	739.8 (0.5)	700.6 (−6.1)	539.3 (0.1)
All industries	633.2 (−0.1)	1150.8 (0.0)	1278.3 (0.1)	1352.6 (1.1)	1122.7 (0.0)

Table 5. Perturbation of weekly earnings.

	Total number of employees	Number of sampled employees	Contribution to unperturbed estimate	Sample RSE (%)
TAU 1	1700	11	2,639,240	13.1
TAU 2	630	8	996,000	4.4
TAU 3	140	6	198,660	14.3
Total unperturbed estimate	2480	25	3,834,000	9.1
Total perturbed estimate	2700	23	3,925,000	9.1

unperturbed estimates of Total Earnings of USD 3,834,000. Given that TAU 1 and 2 dominate the cell, it is likely that such a cell would be suppressed.

Consider if the unperturbed estimate of the cell was released. Under Attack Scenario 3, TAU 2 subtracts their contribution to the unperturbed estimate in order to estimate TAU 1's contribution to Total Earnings. The estimate would be $\hat{y}_1^{(3)} = 3{,}834{,}000 - 996{,}000 =$ 2,838,000 or 7% higher than the actual contribution of TAU 1.

By way of an aside, it is important to note that disclosing TAU 1's contribution to the cell estimate is not by itself disclosure. TAU 1's contribution to the cell estimate, which is based on only 11 out of 1700 of its employees, will differ from TAU 1's true Total Earnings (obtained by summing the Total Earnings of each of its 1700 employees) due to sample error. Based on the sample of eleven of its employees, the RSEs of TAU 1's Total Earnings is 13%. Within the framework of Section 3 we can say, assuming a normal distribution, that there is a 95% chance that TAU 1's contribution to the cell estimate is within 26% of TAU 1's true Total Earnings; or equally we could say that TAU 1's contribution to the estimate is within 18% of its true Total Earnings about 83% of the time. So even if the attacker was able to exactly calculate TAU 1's contribution to the estimate, the sampling of employees provides some protection against disclosing TAU 1's true Total Earnings. Here we conservatively assume that a TAU's contribution to Total Earnings and its true Total Earnings are the same.

If we repeat Attack Scenario 3 using perturbed, rather than unperturbed, estimates we see that the estimate of TAU 1's contribution would be $\hat{y}_1^{(3)} = 3{,}925{,}000 - 996{,}000 = 2{,}929{,}000$ and would be 11.5% larger than the actual contribution of TAU 1. Over repeated perturbations, we showed (details not given) for this example that TableBuilder would not provide sufficient protection (using $R_3 = 11$, $V_3 = 18\%$) of 'TAU 1's contribution' from Attack 3. Furthermore, we could equally have constructed an alternative example whereby a cell only contains the eleven employees from TAU 1. In this alternative example, the risk from disclosing TAU 1's contribution to the cell estimate would be higher still.

In conclusion, while the TableBuilder perturbation settings guarantee a minimum level of disclosure risk at the employee level, they have little control over the disclosure risk for TAUs that are selected with certainty (typically TAUs with more than 50 employees). However, TAUs that are sampled without certainty may well have sufficient protection if the protections of sample error were to be taken into account.

5. Utility of Releasing International Trade in Goods via TableBuilder

International Trade in Goods is a monthly administrative by-product collection of all in-scope imports and exports to/from Australia. ABS policy is that these commodity values must be protected only if that business officially requests (i.e., 'self-select') to be protected against disclosure. When such a 'self-selected' business contributes to a cell, it is determined whether or not the value of the commodity associated with that business breaks confidentiality rules – if it does then the cell is suppressed. Staff who work on this collection describe the current process of suppression as "involved and time-consuming".

Possible alternative approaches to managing disclosure risk:

i. Perturb the commodity values for only self-selected businesses prior to releasing the data as a public use file. In the Australian context, there is a certain level of public sensitivity to releasing even a perturbed commodity amount for a self-selected business. For this reason, this option was not considered further.
ii. Perturb all cell estimates as described in Section 3. This assumes that all businesses self-select and so will result in more perturbation than is strictly required.
iii. Perturb commodity values for self-selected businesses so that, even if they belonged to a cell on their own, the disclosure risk from Attack 1 is acceptable (using the criteria $(R_1,V_1) = (0.15,18)$). The values for businesses that do not 'self-select' are not perturbed. Users can access the micro-data via TableBuilder with all its perturbation routines turned off. In theory, this would give the same estimates as Approach I, but avoids releasing business-level micro-data.

Tables 6 gives published estimates for merchandise exports by state and from the International Trade in Goods and Services, Australia (ABS cat. No 5368.0). We see that under approach II, some estimates are significantly changed by perturbation. A large perturbation is always due to a small number of dominant businesses in a cell. Table 6

Table 6. Merchandise exports (USD M) by state/territory.

	Published estimate	Perturbed estimates-protect all businesses under approach II (percentage impact of perturbation %)	Perturbed estimates-protect only self-selected businesses under approach I and III (percentage impact of perturbation %)
NSW	2822	2910 (3.1)	2837 (0.5)
VIC	1406	1409 (0.2)	1400 (−0.4)
QLD	2927	2971 (1.5)	2927 (0.0)
SA	728	765 (5.1)	724 (−0.6)
WA	9205	8877 (−3.6)	9253 (0.5)
TAS	207	192 (−7.4)	207 (0.0)
NT	444	512 (15.4)	444 (0.0)
ACT	4	6 (35.4)	4 (0.0)
AUS	17743	17642 (−0.6)	17796 (0.3)

shows that the utility loss under the approach III is quite small by comparison. This is because, at least in the estimates of Table 6, large dominant businesses often do not self-select. Feedback from users is that the loss of utility under approach III is acceptable at a high level and further work is planned to consider whether this would also be the case for estimates at fine levels. For cells in which dominating businesses self-select, the perturbation applied by TableBuilder may be unacceptably high. In the next section we see a situation where the perturbation is, in fact, unacceptably high.

6. Utility of Releasing Land Management Practices via Tablebuilder

Land Management Practices Survey (LaMPS) estimates are released every financial year. LaMPS selects a sample of agricultural businesses in Australia above a minimum cut-off size. LaMPS aggregates are released in the form of tables. Suppression is then applied to table cells that are considered to have an unacceptable disclosure risk. Often, estimates in the cell of a LaMPS table will contain a small number of dominant contributors. Next we briefly show the utility of key estimates after they have been perturbed via TableBuilder.

For eight Australian states and territories, Table 7 shows the (RSE) of the published estimate of Total Nitrogen Fertiliser and the impact of perturbation. Using the magnitude values (0.4, 0.3, 0.2), the impact of perturbation is under 5%, with the exception of the Australian Capital Territory (ACT). In a few cases, the impact of perturbation is comparable to the RSE associated with the estimate (e.g., in NT, the RSE was 3.7% and the impact of perturbation was 3.2%).

For the ACT, the impact of perturbation was 42%. This estimate has low utility after perturbation. The impact of perturbation is high because the ACT has a comparatively low number of contributing businesses and some dominant contributors. While LaMPS would not appear to be suitable for release via TableBuilder, the next section discusses some ways forward.

7. Final Remarks

The three case studies in this article discuss the challenges of allowing access to business data via TableBuilder. For some estimates, TableBuilder can provide an effective level

Table 7. RSE and percentage impact of perturbation: Total nitrogen fertiliser applied by state/territory.

State	RSE (%)	Impact of perturbation (%)
NSW	2.7	0.6
Vic	5.6	0.6
Qld	2.9	0.3
SA	4.1	1.8
WA	2.1	0.3
Tas	4.8	3.4
NT	3.7	3.2
ACT*		42.4
Australia	1.4	−0.1

*Published value for ACT was incorporated into NSW.

of protection against disclosure without noticeably affecting utility of the estimates. However, there are certain cell estimates that would not seem to be suitable for release using TableBuilder – these include cells containing a small numbers of businesses that are dominant contributors. More work would be required before the ABS would consider allowing access to its business data via TableBuilder.

However, we believe that the work and findings here will be applicable to other statistical agencies. This is because the features of the business data considered in this paper are common across the world: dominating businesses (e.g., monopoly and duopolies); the need to protect against disclosure at multiple levels of a business hierarchy; and data collected from samples and from administrative sources. Below, we discuss possible applications our work and future work that will improve the disclosure risk-utility trade-off of a remote server approach.

A practical application of our work would be to release as much data as possible through TableBuilder, but to exclude certain subsets of businesses (large businesses). Other methods could be explored for releasing these data subsets – for example, users with a particular research need for the excluded data could apply for access through a special user request, and other methods (such as suppression) could be applied to protect the data. This approach would allow the release of a wide range of business data in a cost-effective way, while still retaining the flexibility to release specific estimates via means other than TableBuilder.

There are some interesting areas for further work:

1. The attacker does not know the target's estimation weight (i.e., it is always assumed to be equal to one). The extent to which this reduces the disclosure risk has not been measured here. A way of taking this into account would be to allow the magnitude values m_i in P^* to depend upon the weight of the K largest contributors, w_i, for $i = 1, \ldots, K$. It is likely that a unit with a high weight would require a much smaller (possibly equal to zero) magnitude value than a unit with a small weight.
2. Sampling (e.g., sampling of employees in Section 4) reduces the risk of disclosure because the attacker does not know if the target unit is selected in the sample. This is important since a benefit of the remote server is that, unlike the release of micro-data, attacks may be required to even establish whether the target is selected in the sample. Chipperfield and O'Keefe (2014) showed even establishing whether or not a target is in the sample can require a significant number of attacks. The reduction in disclosure risk due to sampling could be off-set by a reduction in the degree of perturbation, leading to an increase in utility.
3. Sampling error can reduce the disclosure risk (e.g., in Section 4 we ignored the protection provided to a TAU due to selecting only a sample, rather than all, of its employees). It would be interesting to allow the magnitude of the perturbation to depend upon the degree of protection already provided by sample error (e.g., if sampling employees within a TAU provides sufficient protection, is there a need to perturb the TAU's contribution to estimates)?
4. Preventing a differencing attack from occurring in the first place. This would mean supressing a cell if it, together with a previously released cell, met the condition of a differencing attack.

5. The optimal magnitude parameters (Subsection 3.4) assumed y took only positive values. This could be extended to allow for negative values of y.
6. As mentioned, the current functionality of TableBuilder fixes K, m_i for i $= 1, \ldots,$ K and the distributions of d_i^* and h_i^*. Further work could consider allowing these to depend upon the perceived sensitivity of y and the distribution of y in the cell (e.g., if the top three contributors' relative values of y were approximately -10, 20 and 90).
7. Should the agency release the perturbation parameters underlying P^*? Releasing the parameters would, under any attack scenario, allow an attacker to put a bound on y_1 for each cell total that contained unit 1. The risk and the utility of releasing the parameters would need to be measured. Instead, an indication (perhaps in ranges) of the size of the perturbation or the MSE of the perturbation would be released but, again, any impact on risk and utility should be measured.

8. References

Abrahams, C. and K. Mahony. 2008. "2New Policy and Procedures Governing the Release of Microdata Derived from ONS Social Surveys." *13th GSS Methodology Conference*, London, June 23, 2008. Available at: https://www.ons.gov.uk/ons/media-centre/events/past-events/thirteenth-gss-methodology-conference–23-june-2008 (accessed January 2018).

Chipperfield, J.O. 2014. "Disclosure-Protected Inference with Linked Micro-data using a Remote Analysis Server." *Journal of Official Statistics* 30: 123–146. Doi: http://dx.doi.org/10.2478/jos-2014-0007.

Chipperfield, J.O. and C. O'Keefe. 2014. "Disclosure-Protected Inference using Generalised Linear Models." *International Statistical Review* 82: 371–391. Doi: https://doi.org/10.1111/insr.12054.

Chipperfield, J.O., D. Gow, and B. Loong. 2016. "The Australian Bureau of Statistics and releasing frequency tables via a remote server." *Statistical Journal of the IAOS* 1: 53–64. Doi: https://doi.org/10.3233/SJI-160969.

Domingo-Ferrer, J. and V. Torra. 2001. "Disclosure Protection Methods and Information Loss for Microdata." In *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies*, edited by P. Doyle, J.I. Lane, J.J.M. Theeuwes, and L. Zayatz, 91–110. Amsterdam: North-Holland.

Dwork, C., F. McSherry, K. Nissim, and A. Smith. 2006. "Calibrating Noise to Sensitivity in Private Data Analysis." In *Theory of Cryptography TCC*, edited by S. Halevi and R. Rabin, 265–284. Heidelberg: Springer.

Evans, T., L. Zayatz, and J. Slanta. 1998. "Using Noise for Disclosure Limitation of Establishment Tabular Data." *Journal of Official Statistics* 4: 537–551. Available at: https://www.scb.se/contentassets/f6bcee6f397c4fd68db6452fc9643e68/using-noise-for-disclosure-limitation-of-establishment-tabular-data.pdf (accessed January 2019).

González, J.J.S. 2005. "A Unified Mathematical Programming Framework for different Statistical Disclosure Limitation Methods." *Operations Research* 53: 819–829. Doi: https://doi.org/10.1287/opre.1040.0202.

Krsinich, F. and A. Piesse. 2002. "Multiplicative Microdata Noise for Confidentialising Tables of Business Data." *Statistics New Zealand*. Available at: http://archive.stats.govt.nz/browse_for_stats/businesses/business_characteristics/multiplicative-microdata-noise-for-business-data.aspx (accessed January 2019).

Lucero, J., L. Zayatz, L. Singh, J. You, M. DePersio, and M. Freiman. 2011. "The Current Stage of the Microdata Analysis System at the U.S. Census Bureau." *Proceedings of the World Congress of the International Statistical Institute*, 3115–3133. Dublin. Available at: http://2011.isiproceedings.org/papers/650103.pdf (accessed January 2019).

Miranda, J. and L. Vilhuber. 2013. "Looking back on three years of Synthetic LBD Beta." Cornell University. Available at: http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1013&context=ldi (accessed January 2019).

O'Keefe, C. and J. Chipperfield. 2013. "A Summary of Attack Methods and Confidentiality Protection Measures for Fully Automated Remote Analysis Systems." *International Statistical Review* 81: 426–455. Doi: https://doi.org/10.1111/insr.12021.

Reuter, W.H. and J.M. Museux. 2010. "Establishing an Infrastructure for Remote Access to Microdata at Eurostat." In *Privacy in Statistical Databases*, edited by J. Domingo-Ferrer and E. Magkos, 249–257. Berlin, Heidelberg: Springer.

Tambay, J. 2017. "A layered perturbation method for the protection of tabular outputs." *Survey Methodology* 43: 31–40. Available at: https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2017001/article/14818-eng.pdf?st=qzA3QL0u (accessed January 2019).

Tambay, J.-L. and J.M. Fillion. 2013. "Strategies for processing tabular data using the G-Confid cell suppression software." *Proceedings of the Survey Research Methods Section*. American Statistical Association Joint Statistical Meetings, Montreal, August 3–8, 2013. Available at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2017/7_gconfid.pdf (accessed January 2019).

Thompson, G., S. Broadfoot, and D. Elazar. 2013. "Methodology for the Automatic Confdentialisation of Statistical Outputs from Remote Servers at the Australian Bureau of Statistics." *UNECE Work Session on Statistical Data Confidentiality*, Ottawa, October. Available at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_1_ABS.pdf (accessed January 2019).

Yancey, W.E., W.E. Winkler, and R.H. Creecy. 2002. "Disclosure Risk Assessment in Perturbative Micro-data Protection." In *Inference Control in Statistical Databases*, edited by J. Domingo-Ferrer, 135–151. New York: Springer.

Received September 2016
Revised September 2018
Accepted January 2019