ABS Perturbation Methodology Through the Lens of Differential Privacy

James Bailie*, Chien-Hung Chien**

* Methodology Division, Australian Bureau of Statistics, Australia,
james.bailie@abs.gov.au

** Methodology Division, Australian Bureau of Statistics, Australia,
joseph.chien@abs.gov.au

Abstract. The Australian Bureau of Statistics (ABS), like other national statistical offices, is considering the opportunities of differential privacy (DP). This research considers the Australian Bureau of Statistics (ABS) TableBuilder perturbation methodology in a DP framework. DP and the ABS perturbation methodology are applying the same idea – infusing noise to the underlying microdata – to protect aggregate statistical outputs. This research describes some differences between these approaches. Our findings show that noise infusion protects against disclosure risks in the aggregate Census Tables. We highlight areas of future ABS research on this topic.

1 Introduction

The world is witnessing an explosion in the automated collection of personal data; a reduction in the cost of high-powered computational resources; and the increased frequency and sophistication of data attacks. It is a regular occurrence for cyber attacks to make the news. Naturally, this elevates the public concern over privacy and how personal information is used once collected. Public trust in the Australian Bureau of Statistics (ABS) to protect the data it collects from providers, is a cornerstone to the ABS mission. The US Census Bureau (USCB) recently announced – via its Scientific Advisory Committee – that it would protect the publications of the 2018 End-to-End Census Test (E2E) using differential privacy (DP). The E2E test is a dress rehearsal for the 2020 Census' (Abowd, 2018, p.2). In light of this announcement, many National Statistical Offices (NSOs), including the Office for National Statistics and Statistics New Zealand, are investigating DP approaches to protecting their Census outputs.

The ABS has been exploring the possibility of adopting a DP approach to improve existing confidentiality methodologies, particularly those that apply to the ABS TableBuilder. This paper contributes to the current discussion on DP by providing the current ABS perspective.

2 The Five Safes and The ABS TableBuilder

The ABS develops and implements perturbation methodologies in a dynamic tabular environment (via the ABS TabldBuilder) that give a sensible trade-off between disclosure risk and utility. Fraser and Wooton (2005); Marley and Leaver (2011); Thompson et al. (2013) provide details on the ABS tabular confidentiality methodologies adopted in TableBuilder. These methodologies are defensible against newly developed attacks and claims of vulnerabilities. The credibility (and internal consistency) of ABS outputs is critical to maintaining relevance and purpose. The same-cell-same-perturbation principle is implemented to prevent repeated queries on the same group with independent perturbations, which can be averaged to reduce the noise and potentially lead to disclosure risks (Rinott et al., 2018).

The ABS is committed to use the Five Safes governance framework to ensure microdata can be used appropriately by taking into consideration safe people, projects, settings, data and output (Australian Bureau of Statistics, 2016). Desai et al. (2016) propose the Five Safes framework for assessing the privacy risk of published statistics, from a holistic perspective. There are five dimensions in the risk assessment:

1. Safe Data: How sensitive is the data? Are there variables in the dataset which are identifiers or quasi-identifiers? Does the dataset contain confidential information or is the data public knowledge?

2. Safe Outputs: How risky are the published statistics? Are there outputs on private topics? Can respondents be identified from the statistics? Can information about individual respondents be inferred from the statistics, and how hard would it be to do this?

3. Safe People: Who has access to the data? Who has access to the outputs? Are they trusted, or is it possible that they may have malicious intent?

4. Safe Projects: Is the use of the data appropriate? What are the project objectives and methods? Are they safe?

5. Safe Settings: What are the technological and security controls in place to protect the data and outputs? How is access limited to prevent unauthorised use?

In implementing TableBuilder, the ABS assessed disclosure risk through the Five Safes framework. All of the five dimensions were considered when deciding on the appropriate statistical disclosure methods. Thus, the perturbation protection was tailored to the risk context of the release environment and the input dataset. Under this framework, the ABS can use different levels of perturbation protection to different cells in the TableBuilder, with higher risk cells given more protection than low risk cells. The level of protection is dependent on the risk of outputs to maximise utility while maintaining the required level of privacy.

The methodology behind TableBuilder was developed specifically for a dynamic query environment and provides protections for statistical disclosure risks which are unique to this environment. The TableBuilder algorithm has now been used in the ABS dissemination environment for almost a decade, publishing large, complex statistical releases from big datasets (such as the Australian Censuses).

3 What is Differential Privacy?

Differential Privacy is a framework for protecting privacy by adding random noise to released data, such that the level of noise is attuned to a particular choice of privacy budget. As with any perturbation method, the more noise that is added, the greater the privacy protection but the lower the data-utility. The major advantage of DP, is that it limits the total privacy risk from all the aggregate statistics associated with a given dataset. The limit on total privacy risk is known as the privacy budget. The level of noise added is inversely proportional to the privacy budget. In this way, DP considers the privacy risk arising from all publications from a given dataset, considered in totality. In its risk assessment, it makes no assumptions about an adversary's capability and intent, nor the accuracy and availability of external data. Thus, DP provides protections within the 'Safe Outputs' dimension of the Five Safes framework.

Dwork and Roth (2014) define a publishing method M as (ϵ, δ)-differential privacy if for two neighbouring datasets D and D'

$$Pr(M(D) \in S) \leq \exp(\epsilon) \times Pr(M(D') \in S) + \delta, \tag{1}$$

for all sets of outputs $S \subseteq \text{Range}(M)$. Datasets D and D' are said to be neighbours, if they differ in at most one record. Specifying $\delta = 0$ above gives the definition of *strict* differential privacy. The value of δ measures the failure rate of strict $(\epsilon, 0)$-differential privacy; higher δ means less privacy protection. Generally, the publishing method M adds noise sampled from some distribution. Researchers often use the Laplace distribution with parameters $(\mu = 0, b)$ since it is $(b^{-1}, 0)$-differentially private.

DP thus provides a measurable, theoretical limit to the resulting loss of privacy (known in the literature as the 'privacy guarantee') from any data release and uses the parameter, ϵ, to parameterise the perturbation distribution(s) applied to a dataset. The value of ϵ – known as the privacy budget – is a limit on the total 'privacy-leakage' of a dataset (as measured under the DP framework) allowed by the NSOs. It restricts the total number and overall accuracy of allowable queries, with the aim of ensuring that private information is not revealed through multiple statistical outputs.

However, there is little guidance in the literature on the real life meaning of ϵ. There are unresolved questions surrounding the process for choosing ϵ. Dataset and variable sensitivity and the query space are just two considerations in choosing a privacy budget ϵ. The level of trust in the users, is another important consideration. In fact, all aspects of the data release environment – as described in the Five Safes framework – must be considered. Hence, the authors propose further research examining how DP may be integrated into holistic frameworks such as Five Safes.

In its original formulation, there is no differentiation of outputs between trusted and untrusted users under DP. Since knowing multiple different perturbed values of the same query gives more information, trusted users must receive the same perturbed outputs as untrusted users, to maintain the prescribed level of privacy.

For NSOs, the privacy budget must reflect public opinion on both privacy and utility of the data in question. The translation from public opinion to a particular value of ϵ, will likely be a difficult problem, central to any implementation of DP. With changes in public opinion through time, this cannot be a 'set-and-forget' process. In recent years, a number of theoretical approaches for choosing ϵ have been proposed. For example, Hsu et al. (2014) choose ϵ from a private good perspective and Abowd and Schmutte (2019) choose ϵ from a public good perspective. However, we need more practical examples in real-world statistical publications, where privacy budgets have been set using rigorous methodologies. In any case, for every publication, the assignment of the privacy budget will remain a challenging social problem.

There is also a need for more real-world examples examining the impact of DP on utility, particularly in a dynamic query environment. However, these examples may prove to be difficult, since there are currently no established methodologies for maintaining an optimal risk-privacy tradeoff in a typical NSO publication. The application of DP to large datasets and complex publications can be both computationally and theoretically difficult. Using a dynamic query environment adds another layer of difficulty.

Since the details of any differentially-private mechanism can be published without reducing the privacy protections, DP does enable the transparent measurement of the privacy loss for perturbed outputs. Additionally, DP provides a way to measure the accumulating privacy loss across multiple queries. Cumulative queries can then be costed and used to determine a total level of privacy risk across all the performed queries (Dwork et al., 2011).

4 TableBuilder Perturbation in the Lens of a DP Framework

This section describes ABS TableBuilder perturbation methodology through the lens of DP. Since DP is a property of a publishing mechanism (such as

TableBuilder), it does not require using a particular perturbation distribution. In fact, at its core, DP is a way of measuring and bounding the privacy risk associated with any stochastic publishing mechanism (which includes all mechanisms that sample from a perturbation distribution). In this sense, the ABS TableBuilder perturbation methodology can be viewed through the lens of DP, even though the noise infusion mechanisms typically used in DP are significantly different to TableBuilder.

There are two significant differences between typical DP mechanisms and the ABS perturbation method. Firstly, TableBuilder uses the 'same contributors, same perturbation' principle: the noise added to a query is a function of the rows that contribute to that query. Secondly, TableBuilder protects against specific attacks arising under a dynamic query environment (such as drill-down attacks (Chipperfield et al., 2016)), by perturbing all small counts to zero.

A key attribute of DP is that it limits the total privacy-leakage across all table outputs by setting a universal privacy budget. On the other hand, ABS perturbation methodology introduces different privacy budgets for different table outputs. This is because the ABS perturbation methodology considers different output tables as having different disclosure risks; for example, tables with small cells are riskier than tables with larger cells.

Under a DP implementation, the parameters ϵ, δ are chosen and then the mechanism M is determined to satisfy these parameters. In contrast, this paper aims to assess the privacy protection (as measured through the lens of DP) provided by the existing TableBuilder mechanism. The ϵ, δ parameters of the TableBuilder mechanism were not chosen explicitly, but they were set implicitly when the mechanism was implemented. In a sense, the aim of this work is to make explicit what these implicit ϵ, δ values are.

To simplify the analysis, this research restricts the analysis to **a single TableBuilder counting query**, from a census dataset (i.e. every record in the dataset has a weight of 1). ABS research to explore a DP mechanism in a dynamic table environment, is still on-going. Additionally, applying DP to survey datasets, with weighted records, requires further research (Shlomo et al., 2019; Drechsler, 2019).

Under these restrictions, (ϵ,δ)-differential privacy is equivalent to requiring:

$$Pr(M(n) \in S) \leq \exp(\epsilon) \times Pr(M(n-1) \in S) + \delta, \text{ and} \tag{2}$$

$$Pr(M(n) \in S) \leq \exp(\epsilon) \times Pr(M(n+1) \in S) + \delta, \tag{3}$$

for all $n \in \{0, 1, 2, 3, ...\}$ and all subsets of possible outputs S. (Here n represents the true answer to the counting query and $M(n)$ the perturbed answer, published by the mechanism M.) It is useful to note that firstly, $\delta \geq 1$ is not very informative because any mechanism M satisfies $(0, \delta)$-DP for $\delta \geq 1$. Secondly, if $Pr(M(n) \in S) \leq \delta$ the above definition will always be satisfied for that particular n and S, regardless of the value of ϵ. Hence in meaningfully calculating ϵ, we can restrict our consideration to pairs (n, S) which satisfy $n : Pr(M(n) \in S) > \delta$. Thirdly, S is a subset of all possible outputs (usually non-negative integers), rather than a single count because δ must act as slack variable across the set of all outputs.

When assessing (ϵ,δ)-DP of the TableBuilder mechanism M, we can calculate the failure rate δ by:

$$\begin{aligned} \delta &= \max\{\delta_{n-1}, \delta_{n+1} : n = 0, 1, 2, ...\}, \text{ where} \\ \delta_{n-1} &= Pr(M(n)) < Pr(M(n-1)) + Pr(M(n)) > Pr(M(n-1)), \text{ and} \\ \delta_{n+1} &= Pr(M(n)) < Pr(M(n+1)) + Pr(M(n)) > Pr(M(n+1)). \end{aligned} \tag{4}$$

The privacy budget, ϵ, is given by:

$$\begin{aligned} \epsilon &= \max\{\epsilon_{s,n-1}, \epsilon_{s,n+1} : n = 0, 1, 2, ..., s = 0, 1, 2, ...\}, \text{ where} \\ \epsilon_{s,n-1} &= \begin{cases} \log \frac{Pr(M(n)=s)}{Pr(M(n-1)=s)}, & \text{if } Pr(M(n) = s) > \delta \text{ and } Pr(M(n-1) = s) \neq 0 \\ 0, & \text{otherwise,} \end{cases} \\ \epsilon_{s,n+1} &= \begin{cases} \log \frac{Pr(M(n)=s)}{Pr(M(n+1)=s)}, & \text{if } Pr(M(n) = s) > \delta \text{ and } Pr(M(n+1) = s) \neq 0 \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \tag{5}$$

Figure 1 shows that δ represents the cumulative probability across the non-overlapping regions for neighbouring perturbation distributions. Graphically, δ_{n-1} is the area (coloured with dotted blue lines) under the n perturbation distribution which does not overlap with the $n-1$ perturbation distribution. Similarly, δ_{n+1} is the analogous non-overlapping area (coloured with solid light green lines) between the n and $n+1$ perturbation distributions. We call these areas 'failure zones', since strict $(\epsilon, 0)$-differential privacy fails in these regions. We calculate δ as the maximum of these two areas, δ_{n-1} and δ_{n+1}, across all possible true counts n.

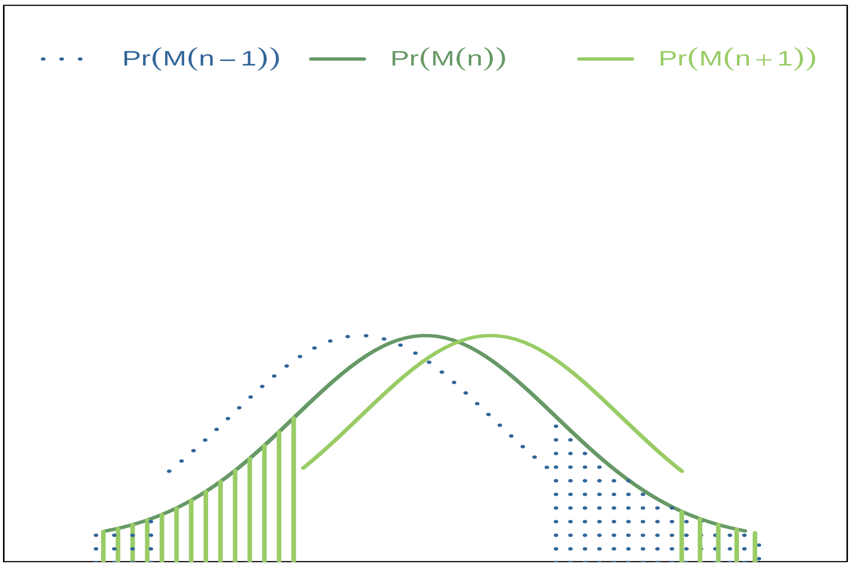

Figure 1: δ for a single query in ABS perturbation methodology

Figure 2 shows that ϵ is calculated as the maximum of $\epsilon_{s,n-1}$ and $\epsilon_{s,n+1}$ across all possible true counts n and all possible outputs s. These values $\epsilon_{s,n-1}, \epsilon_{s,n+1}$ are the log-ratio of probabilities outside the failure zone.

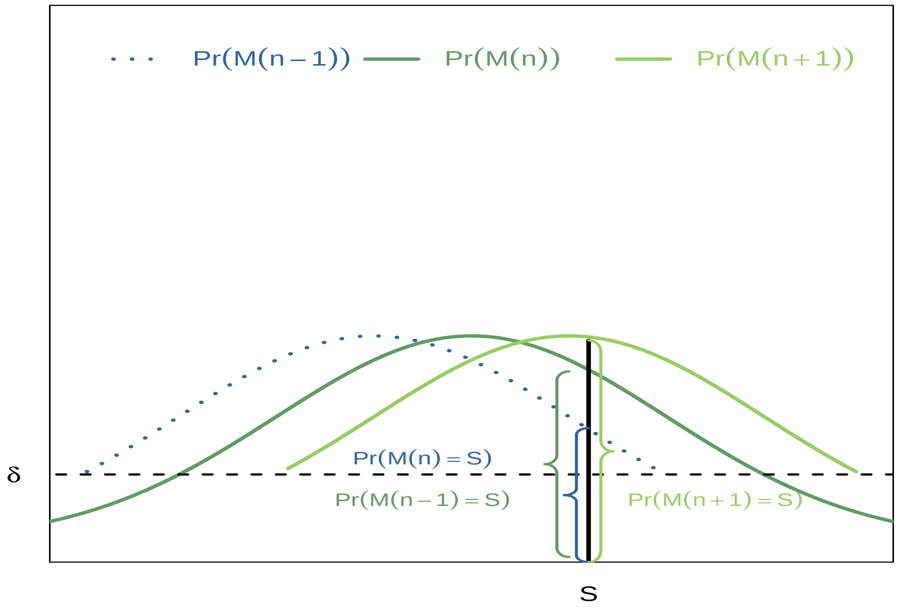

Figure 2: ϵ for single query in ABS perturbation methodology

Table 1 shows the DP parameters for the current TableBuilder perturbation distribution calculated using equations 4 and 5. The value of δ can be interpreted as the probability that the criteria for strict $(\epsilon, 0)$-differential privacy will not be satisfied. However, the δ value of 0.53 for cells less than 7, does not mean that there is a 53% chance that identification will occur. The distinction between attribute disclosure and identity disclosure is a mitigating factor, together with the protections applied at data input and protections within the TableBuilder access system. In addition, TableBuilder obfuscates between real and perturbed zeros. These additional protections are not explicitly taken into account in the DP framework but nevertheless are substantial mitigations against identification risk. In addition, the ABS has determined that these protections do not significantly decrease utility.

	Cell Counts < 7	Cell Counts ≥ 7
ϵ	1.253	0.693
δ	0.53	0.04

Table 1: The DP parameters for the current TableBuilder perturbation distribution.

This work does not show that the current ABS method can be regarded as fully compliant with a DP framework. However, it does help identify the aspects that do or do not accord with a DP framework. To be compliant with DP, the privacy budget ϵ and failure rate δ parameters should be decided upon in advance and from this basis, the perturbation distributions are then determined. There is current research underway at the ABS on how ϵ and δ values can be incorporated in the generation of our perturbation distributions.

Under the 'single query on an unweighted dataset' scenario, we conclude that the TableBuilder perturbation methodology adheres to DP framework. In a dynamic query environment, the current TableBuilder perturbation is (ϵ, δ)-differentially private, only when $\delta \geq 1$ (Rinott et al., 2018). Practically, this means it is not DP. Implementing a dynamic DP mechanism which maintains acceptable levels of both utility and privacy is a challenging methodological problem (Dinur and Nissim, 2003). Additionally, there are challenges to calculate the optimum values of the privacy budget ϵ and the failure rate δ, when users can run a large number of dynamic queries.

5 Conclusions and priorities for future directions

This paper considers how the current TableBuilder perturbation method stacks up against a DP framework. There has been initial exploratory work into a way to augment ABS' current methodologies, to give a way of quantifying the level of differential privacy we currently apply. There are two remaining pieces of work in this research:

- undertaking the reconstruction attack on 2016 Census tabular outputs to identify the risk exposure and the amount of effort required to undertake the attack. Our preliminary results shows the probability of success is low because the ABS only publishes perturbed outputs (See Appendix A); and

- looking further into methodological options for adapting ABS' perturbation distribution into one that is more differentially private.

Beyond these two areas, as outlined in the problem specification, there are other areas of focus to be explored:

- How would we manage the privacy budget in a dynamic query environment?
- How would we address and manage users concerns surrounding differential privacy?
- How would we determine optimum ϵ and δ values in light of the five risk dimensions in the Five Safes framework, in order to optimally balance the trade-off between utility and protection?

As a necessity for effectively progressing this work, the ABS is continuing to build methodological capability in the area of DP. We are working with experts in academia and other organisations (including other NSOs) to build this capability.

A Reconstruction Attacks

The USCB conducted internal research that confirmed that the statistical disclosure limitation systems used for the 2000 and 2010 Censuses had serious vulnerabilities to reconstruction attacks. In these Censuses, a large number of fine level tables were publicly released with no perturbation applied to cell counts. In conducting the internal reconstruction attack they found that 46% of Census records could be reconstructed exactly (on age, sex, race, Hispanic ethnicity and geographic block), and a large number of records could be matched to commercially available databases (Leclerc, 2019). These results - along with the requirement to maintain the same level of publication detail - have motivated the USCB to explore output noise injection as a method of defending against a reconstruction attack. The USBC has given a public undertaking that it is adopting DP for its 2020 Census tabular output releases (Abowd, 2018).

In light of the USCB experience, the ABS has been investigating the risk of a reconstruction attack. The ABS perturbs Census outputs, which means that the same attack algorithm does not work in the Australian context. In light of this, we have investigated modified reconstruction algorithms. Our findings support the US decision to protect its Census outputs by infusing noise. The ABS applies perturbation to its Census outputs and preliminary investigations suggest that a successful attack in the Australian context is low. There are a number of defenses against a attack. In addition to the high level of mathematical and computational expertise required to run such an attack, the attacker would require knowledge of the maximum perturbation bound, which is not publicly available. Since ABS tables are perturbed and true zeroes are obfuscated with perturbed zeros, a large portion of the variable space cannot easily be eliminated (as in the USCB case). This further increases the already significant computational task of the attack.

A.1 Description of ABS Reconstruction Attacks

In a standard reconstruction attack (as conducted by the USCB), the attacker constructs a dataset which is feasible with the published, unperturbed statistics. By feasible, we mean that the attacker's dataset would generate the same published statistics, as were generated by the true dataset. Given that perturbation is applied to Australian Census data before publication, there is an additional layer of complexity to the reconstruction in the ABS context. Instead of taking the published counts as true values, the attacker must assume a bound on the amount of perturbation added and then look for feasible datasets that could be perturbed to produce the published counts (Enderle et al., 2018). For example, given published counts of Male = 3, Female = 5 and Total = 10, along with a maximum perturbation of ±1, the feasible inputs (or plausible true counts) are a) Male = 4, Female = 5; b) Male = 3, Female = 6; or c) Male = 4, Female = 6. Similarly, with published counts of Male = 4, Female = 3, and Total = 5, the feasible inputs are a) Male = 3, Female = 3; b) Male = 4, Female = 2; or c) Male = 3, Female = 2.

As mentioned above, the attacker must assume a bound on the maximum possible perturbation in order to conduct the attack. If the attacker's assumption is incorrect, then the results from the attack will also be incorrect. Hence, for a successful attack, the attacker must know the perturbation bound with certainty. Since this bound is not made publicly available by the ABS, the attacker must obtain it through other means. While it is theoretically possible for an attacker to determine this bound (Asghar and Kaafar, 2019), this is very difficult in a dynamic query environment and even more difficult using only static outputs. This adds another layer of complexity to the attack.

To remove the applied perturbation and determine without doubt what the true count is, the attacker must ascertain that there is only one feasible input that can produce the published count. This can only happen when the perturbation is applied in a particular way (which – as we will show below in Table 2 – happens with a low probability). As a hypothetical example, given true counts of Male = 4, Female = 6, and Total = 10 and a maximum perturbation of ±1, the attacker would be able to remove the perturbation only if the published counts were a) Male = 3, Female = 5 and Total = 11; or b) Male = 5, Female = 7, and Total = 9. Note that perturbation is applied to each cell independently. If other perturbed counts were published, then there would be multiple feasible datasets. So the attacker would be able to reconstruct the true counts (Male = 4, Female = 6, and Total = 10) only if the published counts were a) Male = 3, Female = 5 and Total = 11 or b) Male = 5, Female = 7, and Total = 9.

As the maximum possible perturbation increases, the likelihood of removing the perturbation decreases. In the previous example, assuming a uniform perturbation distribution, the probability of removing the perturbation is $\frac{2}{27}$, since there are 27 possible sets of perturbed counts, but only two sets – a) and b) above – which allow the attack to reconstruct the true counts. Please note that this example does not reflect the perturbation distributions used by the ABS.

We have extended this probability analysis more generally. Table 2 below shows approximate probabilities of an attacker removing the perturbation applied (using the current ABS TableBuilder perturbation distribution) to a single cell with a true count of 1. These success probabilities depend on the complexity of the table used in the attack. Specifically, the probabilities are estimated for tables with m variables and c categories per variable. In calculating these success probabilities, the ABS first ran a reconstruction attack on a number of tables, determining upper and lower bounds on the perturbation applied to each cell. The success probabilities were then derived by extending the results of these preliminary attacks. Apart from the simplifying assumption of constant number of categories per variable, a number of assumptions involving independence between the computed perturbation bounds were made. These assumptions were shown to be approximately valid and unlikely to have a great impact on the success probabilities.

The single cell success probabilities shown in Table 2 increase with the number of variables m used in the tabular request because each new variable adds additional constraints to the perturbation bounds. The probabilities decrease with the number of categories c per variable because more categories increase the degrees of freedom in the constraint equations. For a successful attack on a single cell, the perturbation applied to neighbouring cells must achieve maximal permissible values (contingent on the set of constraint equations) in opposite directions. The chances of this occurring are especially low when category c is large, since the number of neighbouring cells increase with c. The maximum success probability is 3.97797×10^{-3} which occurs in the case of 6 variables with 2 categories per variable.

m	2	3	4	5	6
c					
2	1.327752×10^{-3}	1.990967×10^{-3}	2.653741×10^{-3}	3.316076×10^{-3}	3.977970×10^{-3}
3	3.422734×10^{-5}	5.134057×10^{-5}	6.845351×10^{-5}	8.556616×10^{-5}	1.026785×10^{-4}
4	8.820484×10^{-7}	1.323072×10^{-6}	1.764096×10^{-6}	2.205120×10^{-6}	2.646143×10^{-6}
5	2.273046×10^{-8}	3.409568×10^{-8}	4.546091×10^{-8}	5.682614×10^{-8}	6.819137×10^{-8}
6	5.857655×10^{-10}	8.786482×10^{-10}	1.171531×10^{-9}	1.464414×10^{-9}	1.757296×10^{-9}
7	1.509522×10^{-11}	2.264283×10^{-11}	3.019044×10^{-11}	3.773805×10^{-11}	4.528566×10^{-11}
8	3.890049×10^{-13}	5.835073×10^{-13}	7.780097×10^{-13}	9.725122×10^{-13}	1.167015×10^{-12}
9	1.002468×10^{-14}	1.503703×10^{-14}	2.004937×10^{-14}	2.506171×10^{-14}	3.007405×10^{-14}
10	2.583368×10^{-16}	3.875052×10^{-16}	5.166736×10^{-16}	6.458420×10^{-16}	7.750104×10^{-16}

Table 2: Probability of the attacker determining a single cell with a true count of one, where the number of variables m ranges from two to six, and the number of categories c in each variable ranges from two to ten.

References

Abowd, J. M. (2018). The US Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, page 2867. ACM.

Abowd, J. M. and Schmutte, I. M. (2019). An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review*, 109(1):171–202.

Asghar, H. J. and Kaafar, D. (2019). Averaging attacks on bounded perturbation algorithms. *CoRR*, abs/1902.06414. accessed at http://arxiv.org/abs/1902.06414 on 26/09/2019.

Australian Bureau of Statistics (2016). Information paper transforming statistics for the future. accessed at https://www.abs.gov.au on 01/02/2017.

Chipperfield, J., Gow, D., and Loong, B. (2016). The australian bureau of statistics and releasing frequency tables via a remote server. *Statistical Journal of the IAOS*, 32(1):53–64.

Desai, T., Ritchie, F., and Welpton, R. (2016). Five safes: designing data access for research.

Dinur, I. and Nissim, K. (2003). Revealing information while preserving privacy. pages 202–210.

Drechsler, J. (2019). Differential privacy from a statistical perspective – obtaining valid inferences from differentially private microdata. accessed at https://simons.berkeley.edu/talks/tba-52 on 26/09/2019.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2011). Differential privacy—a primer for the perplexed,". *Joint UNECE/Eurostat work session on statistical data confidentiality 26-28 October*, 11.

Dwork, C. and Roth, A. (2014). *The Algorithmic Foundations of Differential Privacy.* The Algorithmic Foundations of Differential Privacy. now.

Enderle, T., Giessing, S., and Tent, R. (2018). Designing confidentiality on the fly methodology – three aspects. In *Privacy in Statistical Databases*, pages 28–42.

Fraser, B. and Wooton, J. (2005). A proposed method for confidentialising tabular output to protect against differencing. *Monographs of Official Statistics: Work Session on Statistical Data Confidentiality*, pages 299–302.

Hsu, J., Gaboardi, M., Haeberlen, A., Khanna, S., Narayan, A., Pierce, B. C., and Roth, A. (2014). Differential privacy: An economic method for choosing epsilon. In *2014 IEEE 27th Computer Security Foundations Symposium*, pages 398–410. IEEE.

Leclerc, P. (2019). Reconstruction of person level data from data presented in multiple tables. In *Challenges and New Approaches for Protecting Privacy in Federal Statistical Programs*, Washington, United States. National Academies of Sciences, Engineering, and Medicine.

Marley, J. K. and Leaver, V. L. (2011). A method for confidentialising user-defined tables: statistical properties and a risk-utility analysis. In *Proceedings of the 58th Congress of the International Statistical Institute, ISI*, pages 21–26.

Rinott, Y., O'Keefe, C. M., Shlomo, N., Skinner, C., et al. (2018). Confidentiality and differential privacy in the dissemination of frequency tables. *Statistical Science*, 33(3):358–385.

Shlomo, N., Krenzke, T., and Li, J. (2019). Comparison of three post-tabular confidentiality approaches for survey weighted frequency tables. *Transactions on Data Privacy.*

Thompson, G., Broadfoot, S., and Elazar, D. (2013). Methodology for automatic confidentialisation of statistical outputs from remote servers at the australian bureau of statistics. *Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality*, pages 1–37.