Using the Lehmer Mean to Assess Business Data Protection: Statistical Disclosure Control and the Truncated Moment Problem

Stander, Mark; Stander, Julian

Published in:
Transactions on Data Privacy

Publication date:
2024

Document version:
Publisher's PDF, also known as Version of record

TRANSACTIONS ON DATA PRIVACY 18 (2025) 95–112

Using the Lehmer Mean to Assess Business Data Protection: Statistical Disclosure Control and the Truncated Moment Problem

Mark Stander*, **Julian Stander****

*AECOM, United Kingdom.

**Centre for Mathematical Sciences, University of Plymouth, United Kingdom.

E-mail: mark.stander@gmail.com

Received 26 April 2024; received in revised form 30 August 2024; accepted 17 November 2024

Abstract. Confidential business data needs protection against disclosure. Often this data is protected by releasing sample means, variances and higher power moments. Motivated by statistical disclosure control obligations and the need to publish business data safely, we explain how calculating the Lehmer mean from released power moments can lead to the unwanted disclosure of the largest data value. We explain how similar disclosure can apply to smaller data values and provide an approximate solution to the Truncated Moment Problem. We briefly discuss the Gini mean and the relationship between sample central and raw moments.

Keywords. Gini mean; Lehmer mean; Power moments; Statistical disclosure limitation; Truncated Moment Problem

1 Introduction

The paper is motivated by two related problems. The first is the obligation on national statistical offices to publish data in a way that protects the confidentiality of respondents. The second is the need of businesses to release data in a similarly safe manner, for open-source statistical analysis, for example. Often these data sets are protected from the disclosure of sensitive values by releasing sample means, variances and higher power moments.

For positive data, the Lehmer mean is a generalized power mean that can be computed from released consecutive power moments. The Lehmer mean tends upwards to the largest data value as the power increases. The Lehmer mean therefore provides an approximation of the largest data value, the quality of which can be controlled by increasing the power. This means that, if too many power moments are released, there is a risk that the largest data value can be disclosed. The second largest, and smaller data values, can be similarly at risk. We illustrate this using project revenue data from AECOM, an infrastructure consulting firm.

The AECOM data set comprise $8,912$ project revenues from an AECOM business line. Subsets of this data set are representative of data that AECOM shares with third parties. The sample mean revenue

is $49,500$ units (3 significant figures); we do not provide the currency because of commercial confidentiality. The sample median is $35,700$, the standard deviation is $48,400$ and the interquartile range is $72,100$, while the sample skewness and kurtosis are 1.1 and 3.7, indicating considerable asymmetry and peakedness in the project revenue distribution. The sample mean, standard deviation, sample skewness and kurtosis are all commonly released summary statistics from which the first four power moments can be obtained.

To put our work in a legal context, we briefly discuss guidance from the UK and the European Union (EU). The EU's statistical office Eurostat identifies that in the EU there are two data protection frameworks; see [1]. First, the General Data Protection Framework (GDPR) applies to personal data. The second framework concerns 'the protection of data collected for statistical purposes, also called statistical confidentiality', and is 'a fundamental principle of official statistics'. This statistical confidentiality means that 'rules and measures must be taken to prevent disclosure'. One of the examples given by Eurostat ([1], on their web-page 'Statistical confidentiality and personal data protection') of data to which the second framework applies is 'the aggregated turnover of a specific type of company located in a specific region'. The UK Office for National Statistics ([2]) discusses releasing data from both business and social surveys in accordance with the 'Code of Practice for Official Statistics'; see [3].

Our work is relevant to statistical disclosure control (SDC), the science of protecting sensitive information in released data. In their book-length summary, [4] describe SDC as 'the set of methods to reduce the risk of disclosing information on individuals, businesses or other organisations. SDC methods minimise the risk of disclosure to an acceptable level while releasing as much information as possible' (their Section 1.1.2). [4] discuss how SDC methods must preserve statistical properties of the data set being protected. These preserved properties include: means, totals, (higher or lower order) moments, variances and the structure of the data. Our work helps us to understand the disclosure risks of SDC techniques where the preserved properties are moments or their equivalents.

There has been an enormous output related to SDC including, as a small set of theoretical and practical examples, [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. [18] provides a recent overview of disclosure risks and their quantification. She also discusses data utility and common statistical disclosure limitation methods, partially in the context of remote analysis servers. [18] mentions the problems associated with the release of maximum and minimum values, and percentiles. Similar issues are also discussed in [17], for example. [19] discuss SDC in microdata from social surveys released by statistical agencies, mentioning differential privacy, a formal method of privacy protection described in [20] for frequency tables. Our methods would also be useful in other contexts, such as in health settings, where deductive disclosure might lead to heavy fines, although we do not discuss these in detail.

Finally, our work also makes a contribution to the Truncated Moment Problem (TMP). The TMP concerns estimating data values from released power moments and is therefore related to SDC. [21] seems to have been the first to consider – 150 years ago – the TMP, which can be regarded as a finite dimensional version of the Hausdorff Moment Problem. The TMP is important when there is no access to the original data, but sample moments are available. The TMP has a large range of applications including physics, computer science, geography, probability, environmental science, engineering including chemical engineering, and geo-physics; see [22]. One approach to the TMP is to assume a simple distribution shape which is then fitted using the method of moments; see [23] and [24]. For other approaches related to reconstructing a distribution from its moments, see [25], [9] and [26]. The ability of our method to approximate all data values provides an applied solution to the TMP.

1.1 Overview and Structure of the Paper

The essence of our work is the following. Companies and other organisations often perform SDC by releasing power moments. We can approximate the largest values of a data set from released power moments using Lehmer means. Consequently, sensitive values thought to be protected may

not actually be so. We therefore discuss the degree to which released moments can disclose data values.

We believe that the Lehmer mean deserves to be better known and in Section 2 we present some of its key properties that we will use to approximate the largest data values. We describe in detail how to approximate the largest and second largest data values in Section 3, where we also discuss how we can iterate our method to approximate all data values. In Section 4 we apply these approximations to the AECOM project revenue data and assess their performance in the context of data disclosure from released moments. As the protection of multivariate data is often of interest in official statistics, we also discuss the extension of our approach to bivariate data using a simulation study. Brief conclusions and suggestions for further work are given in Section 5. The relationship between central and raw moments, together with proofs of Lehmer mean properties are presented in appendices.

2 Moments, Power Means, and the Gini and Lehmer Means

The Gini and Lehmer means are examples of generalized power means. The Lehmer mean seems to have been introduced by [27]. Before discussing these means, we need some definitions. The *p-th raw moment* of a continuous random variable X with probability density function f is given by:

$$\mu_p = \mathrm{E}[X^p] = \int x^p \, f(x) \, dx,$$

where the integral is over all possible values of X.

The *p-th sample raw moment* for a random sample $x_1, \ldots, x_n$ of n data points is:

$$M_{p;n} = \frac{1}{n} \sum_{k=1}^{n} x_k^p. \tag{1}$$

$M_{0;n} = 1$ and $M_{1;n} = \sum_{k=1}^{n} x_k/n = \bar{x}$ is the sample mean. $M_{p;n}$ is an unbiased estimator of μ_p; see [28]. The quantity $M_{p;n}^{1/p}$ is referred to as a power mean.

The *p-th sample central moment* ([29]) is defined as

$$c_{p;n} = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^p.$$

$c_{0;n} = 1$, $c_{1;n} = 0$ and $c_{2;n} = \sum_{k=1}^{n} (x_k - \bar{x})^2/n$ is one version of the sample variance. $n\, c_{2;n}/(n-1)$ provides an unbiased estimator of $\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$. Raw and central moments are related, as recalled in Appendix A. This relationship means that it does not matter whether sample central or raw moments are released because one can be obtained from the other. For example, one well known relationship is $s_x^2 = c_{2;n} = M_{2;n} - \bar{x}^2$, where s_x is a sample standard deviation. Central moments can be *standardized* leading to the *sample skewness* $c_{3;n}/s_x^3 = c_{3;n}/c_{2;n}^{3/2}$ and the *sample kurtosis* $c_{4;n}/s_x^4 = c_{4;n}/c_{2;n}^2$. Because the sample variance, skewness and kurtosis are defined in terms of central moments, they are invariant to shifts applied to all the data.

Let $x_1, \ldots, x_n > 0$ be a positive data set, and assume that $x_1, \ldots, x_n$ are not all equal. We define the *Gini mean* as:

$$G_{r,s;n} = \left(\frac{M_{r+s;n}}{M_{s;n}} \right)^{\frac{1}{r}} = \left(\frac{\sum_{k=1}^{n} x_k^{r+s}}{\sum_{k=1}^{n} x_k^s} \right)^{\frac{1}{r}}, \; r \neq 0.$$

Let $r = p$ and $s = 0$, then $G_{p,0;n} = M_{p;n}^{1/p}$, a power mean. Hence, the Gini mean is an example of a generalized power mean.

The *Lehmer mean* $L_{p;n}$ is a special case of the Gini mean with $r = 1$ and $s = p - 1$:

$$L_{p;n} = G_{1,p-1;n} = \frac{M_{p;n}}{M_{p-1;n}} = \frac{\sum_{k=1}^{n} x_k^p}{\sum_{k=1}^{n} x_k^{p-1}},$$

and so is also an example of a generalized power mean. It is easy to see that $L_{1;n} = \bar{x}$. Similarly, $L_{0;n} = n/\sum_{k=1}^{n}(1/x_k)$ is the harmonic mean. Calculation of the Lehmer mean requires the consecutive pair of moments $(M_{p;n}, M_{p-1;n})$; the Gini mean may be computed from a non-consecutive pair.

It is convenient to re-define some of our above quantities in terms of ordered data points $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$. To do this, we let $S_{p;n} = x_{(1)}^p + \cdots + x_{(n-1)}^p + x_{(n)}^p$ be the sum of the ordered data values and re-define $M_{p;n} = \sum_{k=1}^{n} x_{(k)}/n$. Then, $L_{p;n} = M_{p;n}/M_{p-1;n} = S_{p;n}/S_{p-1;n}$.

We now state two properties of the Lehmer mean on which we base our approximation of $x_{(n)}$:

Property 1 $L_{p;n}$ is a monotonically increasing function of p.

Property 2 $L_{p;n} \nearrow x_{(n)}$ as $p \nearrow \infty$, meaning that the Lehmer mean tends upwards to the largest data value as p increases.

These properties are proved in Appendix B. From them it follows that $L_{p;n} \leq x_{(n)}$.

A proof similar to that for Property 2 allows us to establish that $G_{r,s;n} \nearrow x_{(n)}$ as $r, s \nearrow \infty$; see also [30]. In this paper, we concentrate on the special case of $L_{p;n}$ because it is easy to control through its single parameter p. For very highly positively skewed data sets the arithmetic mean $\bar{x} = L_{1;n}$ may reflect $x_{(n)}$, but approximations provided by $L_{p;n}$ would offer greater control through the choice of p.

3 Approximating the Largest Value $x_{(n)}$ and the Second Largest Value $x_{(n-1)}$

Because $L_{p;n} \nearrow x_{(n)}$ as $p \nearrow \infty$, $\hat{x}_{(n)}^{(p)} = L_{p;n}$ can be used to approximate $x_{(n)}$, with the accuracy of this approximation increasing as p increases. Released consecutive moments can be used to calculate $\hat{x}_{(n)}^{(p)}$, which can therefore lead to the unwanted disclosure of $x_{(n)}$. We quantify this disclosure risk by the relative error when approximating $x_{(n)}$ by $\hat{x}_{(n)}^{(p)}$, $\left(x_{(n)} - \hat{x}_{(n)}^{(p)}\right)/x_{(n)}$, expressed as a percentage; this quantity is always non-negative because $x_{(n)} \geq \hat{x}_{(n)}^{(p)}$. The smaller this error, the greater is the risk of the unwanted disclosure of $x_{(n)}$. [31] used this quantity to quantify the risks of an 'intruder' learning the largest and second largest values in the context of business data.

Similarly, we can approximate and therefore assess the disclosure risk for the second largest value $x_{(n-1)}$ using $L_{p';n-1} = S_{p';n-1}/S_{p'-1;n-1}$, where the reduced sums $S_{p';n-1}$ and $S_{p'-1;n-1}$ can be approximated as

$$\hat{S}_{p';n-1} = S_{p';n} - (\hat{x}_{(n)})^{p'} = S_{p';n} - (L_{p;n})^{p'} \text{ and } \hat{S}_{p'-1;n-1} = S_{p'-1;n} - (L_{p;n})^{p'-1}.$$

If therefore follows that

$$x_{(n-1)} \approx L_{p';n-1} = \frac{S_{p';n-1}}{S_{p'-1;n-1}} \approx \frac{S_{p';n} - (L_{p;n})^{p'}}{S_{p'-1;n} - (L_{p,n})^{p'-1}},$$

yielding the following approximation $\hat{x}_{(n-1)}^{(p,p')}$ to $x_{(n-1)}$:

$$\hat{x}_{(n-1)}^{(p,p')} = \frac{S_{p';n} - (L_{p;n})^{p'}}{S_{p'-1;n} - (L_{p,n})^{p'-1}}. \tag{2}$$

Note that two pairs of power moments $(M_{p;n}, M_{p-1;n})$ and $(M_{p';n}, M_{p'-1;n})$ are needed to find $\hat{x}_{(n-1)}^{(p,p')}$. It can be shown that, if $p' = p$, then $\hat{x}_{(n-1)}^{(p,p)} = L_{p,n} = \hat{x}_{(n)}$, which is not helpful, meaning that the case $p' = p$ should be avoided.

A heuristic argument outlined in Appendix C and results obtained from the AECOM business revenue data in Section 4 suggest that, if $p' < p$, then $\hat{x}_{(n-1)}^{(p,p')} \to x_{(n-1)}$ as $p, p' \to \infty$. We quantify the error and hence the disclosure risk when approximating $x_{(n-1)}$ by $\hat{x}_{(n-1)}^{(p,p')}$ using $\left(x_{(n-1)} - \hat{x}_{(n-1)}^{(p,p')}\right)/x_{(n-1)}$, expressed as a percentage. If $\hat{x}_{(n-1)}^{(p,p')} > \hat{x}_{(n)}^{(p)}$, as often happens when $p' > p$, then $\hat{x}_{(n-1)}^{(p,p')}$ is set to $\hat{x}_{(n)}^{(p)}$. This is because $x_{(n-1)} \leq x_{(n)}$ by definition, and so having $\hat{x}_{(n-1)}^{(p,p')} > \hat{x}_{(n)}^{(p)}$ is just an artifact of estimation error.

If the data are grouped in the sense that values are repeated, then, for large p, $\hat{x}_{(n)}^{(p)}$ will be closer to $x_{(n)}$ the further $x_{(n)}$ is from $x_{(n-1)}$. If data are based on intervals, then similar considerations apply if we work with interval midpoints.

3.1 An Alternative Way of Approximating $x_{(n-1)}$

An alternative approximation for $\hat{x}_{(n-1)}$, similar to (2) but involving only one pair of consecutive power moments takes the form

$$\hat{x}_{(n-1)}^{(p)} = \frac{S_{p;n} - (L_{p;n}^*)^p}{S_{p-1;n} - (L_{p;n}^*)^{p-1}}, \tag{3}$$

in which $L_{p;n}^*$ is a rounded up version of the Lehmer mean $L_{p;n}$; we round up because $L_{p;n} \nearrow x_{(n)}$.

3.2 Recovering All Data Values

In both the above cases, our approximation of $x_{(n-1)}$ is based on an approximation of $x_{(n)}$. We can therefore iterate to approximate all data values to some degree of accuracy. As an example, using the approximation $\hat{x}_{(n-1)}^{(p)}$ given in (3), our approximation of $x_{(n-2)}$ would be

$$\hat{x}_{(n-2)}^{(p)} = \frac{S_{p;n} - \left(L_{p;n}^*\right)^p - \left(\hat{x}_{(n-1)}^{(p)}\right)^p}{S_{p-1;n} - \left(L_{p,n}^*\right)^{p-1} - \left(\hat{x}_{(n-1)}^{(p)}\right)^{p-1}}. \tag{4}$$

This method provides an approximate solution to the TMP. In the context of SDC, it allows us to understand the risks of unwanted disclosure of smaller data values given release of higher-order moments.

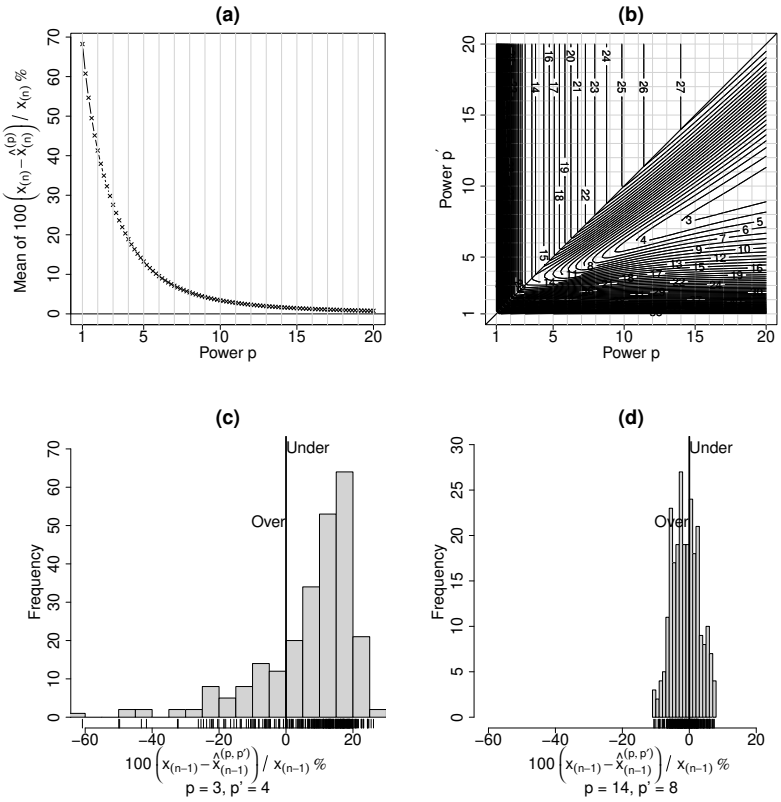
4 Results: Disclosure Risk for Real Business Data

As is often the case, the largest values of the AECOM project revenues data are particularly sensitive because of contractual confidentiality. We therefore use the AECOM data to assess the error associated with the approximations $\hat{x}_{(n)}^{(p)} = L_{p,n}$ and $\hat{x}_{(n-1)}^{(p,p')}$, and hence to understand the risk of disclosure for $x_{(n)}$ and $x_{(n-1)}$. We randomly generated data sets of size $n = 20$ by sampling with replacement. We choose $n = 20$ as an example of the size of data that AECOM shares with third parties.

Figure 1 (a) shows, for values of p from 1 to 20, the mean of the percentage relative error $100(x_{(n)} - \hat{x}_{(n)}^{(p)})/x_{(n)}\%$ associated with using $\hat{x}_{(n)}^{(p)}$ to approximate $x_{(n)}$ over 250 randomly generated data sets. This error tends down towards 0 as p increases, in accordance with Property 2. Hence, the disclosure risk increases as p increases.

Figure 1 (b) presents, for each power pair (p, p') in a grid with sides from 1 to 20, the mean percentage absolute relative error $100\left|x_{(n-1)} - \hat{x}_{(n-1)}^{(p,p')}\right|/x_{(n-1)}\%$ over 250 randomly generated data sets, the absolute value being taken so that negative errors do not cancel out positive ones. We see that $\hat{x}_{(n-1)}^{(p,p')}$

Figure 1: (a) Mean percentage relative error when $\hat{x}_{(n)}^{(p)}$ approximates $x_{(n)}$. (b) Mean absolute percentage relative error when $\hat{x}_{(n-1)}^{(p,p')}$ approximates $x_{(n-1)}$. (c) Percentage relative errors for 250 data sets for $\hat{x}_{(n-1)}^{(3,4)}$. (d) Percentage relative errors for 250 data sets for $\hat{x}_{(n-1)}^{(14,8)}$.



approximates $x_{(n-1)}$ well when $p' < p$ (below the positive diagonal). This is further explored in Figure 1 (c) and (d). Figure 1 (c) quantifies the approximation error when $(p, p') = (3, 4)$ ($p' > p$), by showing a histogram of the 250 relative percentage errors. Figure 1 (d), which is for $(p, p') = (14, 8)$ ($p' < p$), shows smaller errors, as the heuristic argument outlined in Appendix C suggests. The higher is the approximation error, the lower is the disclosure risk for $x_{(n)}$ and $x_{(n-1)}$. Therefore, plots similar to Figure 1 can be used to suggest values of p and p' and therefore the power moments that can be released, whilst maintaining a required level of disclosure protection for $x_{(n)}$ and $x_{(n-1)}$.

4.1 Results for our Alternative Way of Approximating $x_{(n-1)}$

Our alternative way of approximating $x_{(n-1)}$ described in Section 3.1 depends on only one power p and yields $\hat{x}^{(p)}_{(n-1)}$ given in (3). To illustrate the performance of $\hat{x}^{(p)}_{(n-1)}$, we rounded the AECOM project revenue data to the nearest $1,000$ units. Hence, we set $L^*_{p,n}$ to $L_{p,n}$ rounded up to the nearest $1,000$.

Figure 2 (a) shows the mean percentage absolutely relative errors

$$100 \left| x_{(n)} - \hat{x}^{(p=4)}_{(n)} \right| / x_{(n)}\% \text{ and } 100 \left| x_{(n-1)} - \hat{x}^{(p=4)}_{(n-1)} \right| / x_{(n-1)}\%,$$

when $\hat{x}^{(p)}_{(n)}$ and $\hat{x}^{(p)}_{(n-1)}$ approximate $x_{(n)}$ and $x_{(n-1)}$, over the 250 randomly generated data sets. It can be seen that $\hat{x}^{(p)}_{(n-1)}$ performs better than $\hat{x}^{(p)}_{(n)}$ for low values of the power p. This error measure become small as p get large for both approximations. Figure 2 (b) and (c) examine this in greater detail for $p = 4$ by showing histograms of the 250 percentage relative errors $100 \left(x_{(n)} - \hat{x}^{(p=4)}_{(n)} \right) / x_{(n)}\%$ and $100 \left(x_{(n-1)} - \hat{x}^{(p=4)}_{(n-1)} \right) / x_{(n-1)}\%$. As expected, the $100 \left(x_{(n)} - \hat{x}^{(p=4)}_{(n)} \right) / x_{(n)}\%$ values are always positive. The values of $100 \left(x_{(n-1)} - \hat{x}^{(p=4)}_{(n-1)} \right) / x_{(n-1)}\%$ are generally lower, but there is quite a long tail corresponding to over-estimation. Again, plots similar to Figure 2 can be used to suggest the value of p and therefore the power moments that can be released, whilst maintaining a required level of disclosure protection.

4.2 Approximating All Data Values

As mentioned in Section 3.2, we can use an iterative approach to approximate successively smaller values of a positive data set. An example of this for $x_{(n-2)}$ was given in (4). In certain circumstances and using sufficiently large values of p, it may be possible to recover the whole data set. We explored this using a simulation study. We generated 250 random samples $x_1, \ldots, x_n$ of size n from a Poisson distribution with mean μ: $x_i \sim \text{Po}(\mu)$, independently, $i = 1, \ldots, n$. Because the data values are integer, we used rounding up to the nearest integer in our approximations. We then found the smallest value of p needed to reconstruct $x_1, \ldots, x_n$. Our results are shown for sample sizes $n = 1, \ldots, 100$ and for $\mu = 10, 25, 50$ and 75 in Figure 3, which was produced using the `ggplot2` R ([32]) package of [33]. In Figure 3 we plot the points $(n, \text{smallest } p)$ and we add smoothers (blue curves) to show the trends. We have also added the line $2 \times \text{smallest } p = n$ (black lines) to each panel of Figure 3 because a pair of moments $(M_{p;n}, M_{p-1;n})$ is required to find $L_{p;n}$. Values above this line require more higher order moments than the sample size n. When the Poisson mean μ is small, we can perform data set recovery using small values of p for quite small sample sizes. When μ is larger, data set recovery requires considerably higher values of p.

4.3 Bivariate Data

Following the suggestion of a reviewer, we now briefly discuss an extension of our largest value approximation method to bivariate data. We simulated data from a bivariate Student t copula (Ex

Figure 2: (a) The dependence on the power p of the mean percentage absolute relative error when $\hat{x}_{(n)}^{(p)}$ and $\hat{x}_{(n-1)}^{(p)}$ approximate $x_{(n)}$ and $x_{(n-1)}$ over 250 randomly generated data sets. (b) Values of the percentage relative errors $100\left(x_{(n)}-\hat{x}_{(n)}^{(p=4)}\right)/x_{(n)}\%$ from 250 randomly generated data sets. (c) Values of the percentage relative errors $100\left(x_{(n-1)}-\hat{x}_{(n-1)}^{(p=4)}\right)/x_{(n-1)}\%$ from 250 randomly generated data sets.

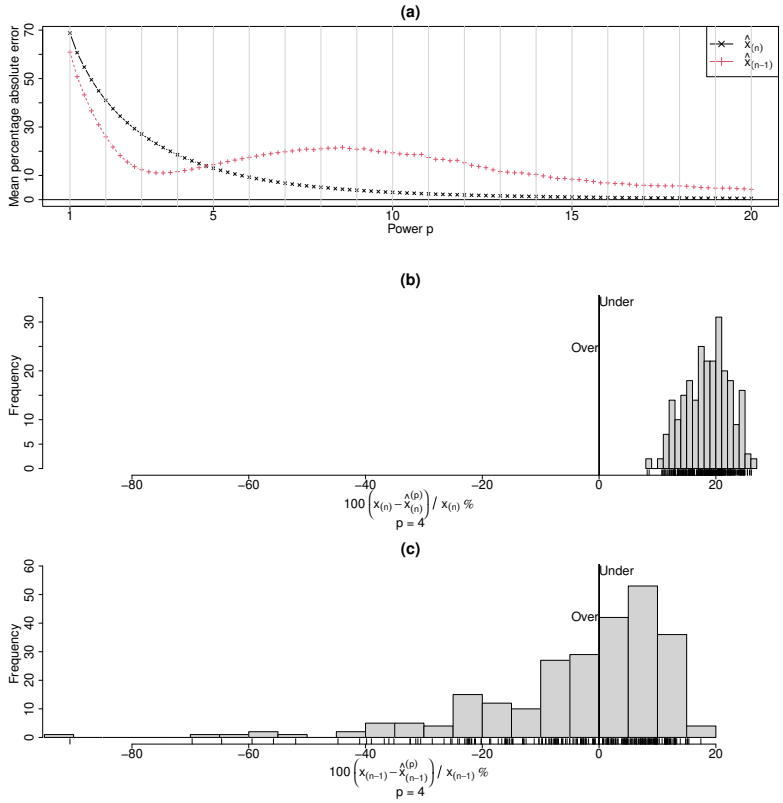
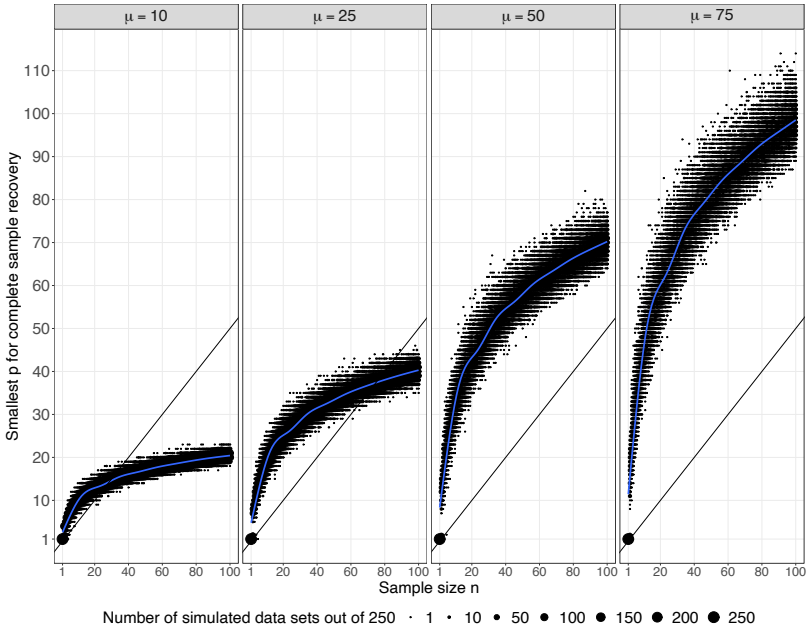


Figure 3: The dependence of the smallest value of p needed for complete sample recovery on sample size n. Blue curves: smoothers through the $(n, \text{smallest } p)$ data. Black lines: these have the equation $2 \times \text{smallest } p = n$.



1.13, [34]) and transformed the margins so that they followed a log-normal distribution. Let the resulting data be (x_i, y_i), $i = 1, \ldots, n$. The log-normal parameters were chosen with reference to the AECOM project revenues data using the method of moments. In this way, there is some match between the simulated $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ data sets, and the project revenues data. For this simulation we used the `BiCopSim` function of the `VineCopula` `R` package [35]. We decided to work with the bivariate Student t copula because, unlike the bivariate Gaussian copula (Ex 1.12, [34]), it has non-zero upper tail dependence (Section 2.3, [34]). Other copulas could have been chosen. See [36] for an example of bivariate copula modelling. We parameterized the bivariate Student t copula in terms of Kendall's τ (Section 2.2, [34]). It can be shown that the value of Kendall's τ only depends on the copula and not on the marginal specification. Let us assume that interest is in protecting the largest values, $x_{(n)}$ in $x_1, \ldots, x_n$ and $y_{(n)}$ in $y_1, \ldots, y_n$. It would be unlikely for $(x_{(n)}, y_{(n)})$ to be a data point. We approximate $x_{(n)}$ as $\hat{x}_{(n)}^{(p)}$ and $y_{(n)}$ as $\hat{y}_{(n)}^{(p)}$ using Lehmer means.

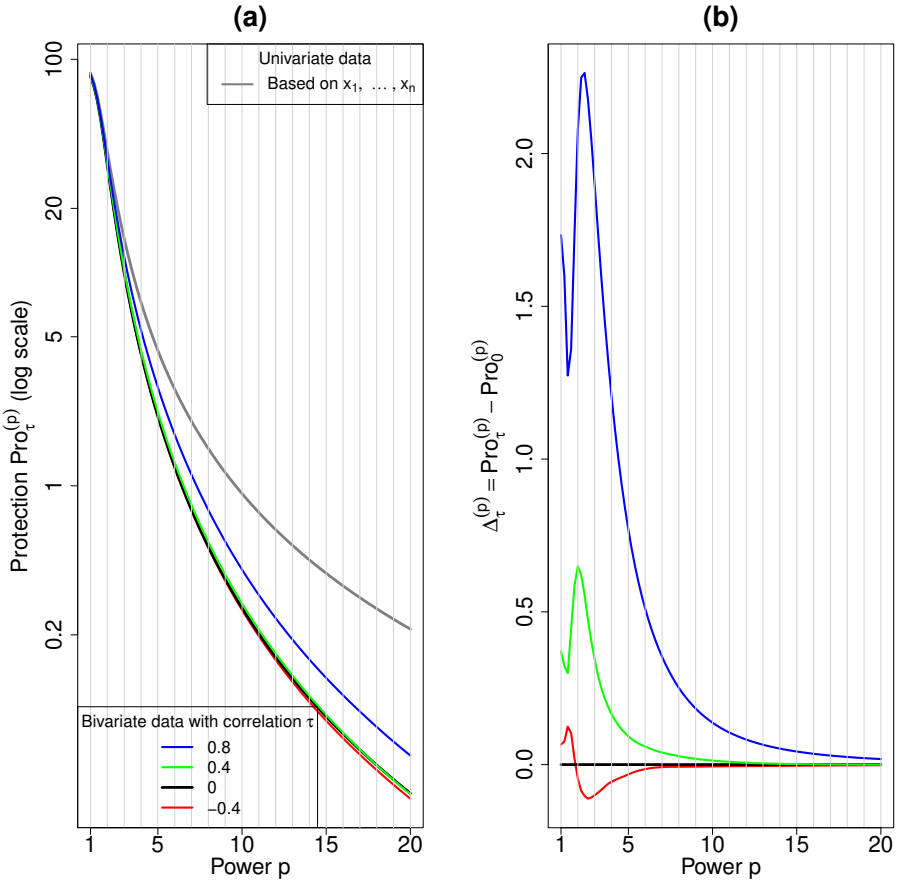
We considered as a measure of joint protection the mean $\text{Pro}_\tau^{(p)}$ of

$$100 \min \left\{ \frac{\left| x_{(n)} - \hat{x}_{(n)}^{(p)} \right|}{x_{(n)}}, \frac{\left| y_{(n)} - \hat{y}_{(n)}^{(p)} \right|}{y_{(n)}} \right\} \%$$

over a sufficient large number of simulated data sets to ensure the stability of our results, each data set being simulated as described using a bivariate Student t copula parameterized by τ. $\text{Pro}_\tau^{(p)}$ reflects the least protected of the two maxima $x_{(n)}$ and $y_{(n)}$. Figure 4 (a) shows how $\text{Pro}_\tau^{(p)}$ decreases as p increases, when $\tau = -0.4$, 0, 0.4 and 0.8. Figure 4 (a) also shows, for comparison, the dependence of the mean of $100 \left(x_{(n)} - \hat{x}_{(n)}^{(p)} \right) / x_{(n)} \%$ over simulated data sets on p (upper curve). We see from Figure 4 (a) that, not surprisingly, more protection is offered to univariate data than to bivariate data. Moreover, higher values of $\text{Pro}_\tau^{(p)}$ occur when $\tau = 0.8$ than when $\tau = 0.4$, 0 and -0.4. To investigate

this further, in Figure 4 (b) we plot $\Delta_\tau^{(p)} = \text{Pro}_\tau^{(p)} - \text{Pro}_0^{(p)}$ for $\tau = -0.4$, 0, 0.4 and 0.8. $\Delta_\tau^{(p)}$ can be thought of as $\text{Pro}_\tau^{(p)}$ centred on $\text{Pro}_0^{(p)}$. We see that $\Delta_{0.8}^{(p)} > \Delta_{0.4}^{(p)}$ for all p, so that $\text{Pro}_{0.8}^{(p)} > \text{Pro}_{0.4}^{(p)}$. Similarly, $\Delta_{0.4}^{(p)} > \Delta_0^{(p)} = 0$ for almost all p, so that $\text{Pro}_{0.4}^{(p)} > \text{Pro}_0^{(p)}$. There is very little difference between $\Delta_0^{(p)}$ and $\Delta_{-0.4}^{(p)}$, possibly due to the fact that the upper tail dependence is almost 0 in both cases, meaning that $\text{Pro}_{0.4}^{(p)}$ and $\text{Pro}_0^{(p)}$ are similar. This simulation study suggests that our method offers higher protection to more strongly correlated bivariate data.

Figure 4: (a) The dependence on p of the mean of $100\left(x_{(n)} - \hat{x}_{(n)}^{(p)}\right) / x_{(n)}\%$ over simulated data sets (upper curve), together with the dependence on p of the protection $\text{Pro}_\tau^{(p)}$ for $\tau = -0.4$, 0, 0.4 and 0.8 (lower curves). (b) The dependence on p of $\Delta_\tau^{(p)} = \text{Pro}_\tau^{(p)} - \text{Pro}_0^{(p)}$ (centred protection) for $\tau = -0.4$, 0, 0.4 and 0.8.



5 Conclusion and Discussion

Working in the context of SDC for business data and using approximations of the largest data values based on Lehmer means, we have proposed a way of understanding the risks of unwanted dis-

closures of sensitive large values when power moments are released. These risks can therefore be controlled in a simple way by the powers that are released. We have illustrated this using AECOM project revenue data. Graphs such as Figure 1 allow businesses to decide which power moments to release. We have also illustrated how our approximations can be iteratively extended to approximate or even recover all data values. In addition, we have briefly discussed properties of the Lehmer mean, the TMP and bivariate data disclosure control.

Data Availability Statement

Computer code, written in R ([32]), to generate the figures and perform the analysis described in this paper has been made available [37]. The data supporting the real business data results presented in this paper is commercially sensitive and so has not been made available. The data supporting the bivariate experiment is synthetic and similar data can be produced, see for example [36]. Although the business data has not been supplied, please note that the computer code can be used with any suitably formatted numeric data set.

Acknowledgements

We are grateful to the Editors Professors Vicenç Torra and Josep Domingo-Ferrer and to two reviewers for their astute suggestions that have substantially improved this article. We thank Professor Silvia Polettini for very insightful and supportive comments, and Professor Mario Cortina Borja and Dr Luisa Franconi for helpful discussions.

Disclosure Statement

The authors report no conflict of interest.

A Appendix A: The Relationship between Sample Central and Raw Moments

It does not matter whether sample central or raw moments are released, because one can be obtained from the other, as we now illustrate. As discussed in [38, 39, 40] for example, to obtain sample central moments from sample raw moments we expand $(x_k - \bar{x})^p$ using the binomial theorem:

$$\begin{aligned} c_{p;n} &= \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^p \\ &= \frac{1}{n}\sum_{k=1}^{n}\left\{\sum_{l=0}^{p}\binom{p}{l} x_k^l\,(-\bar{x})^{p-l}\right\}, \text{ by the binomial theorem} \\ &= \sum_{l=0}^{p}\binom{p}{l}(-1)^{p-l}\left(\frac{1}{n}\sum_{k=1}^{n}x_k^l\right)\bar{x}^{p-l}, \text{ bringing through the sum in } k \\ &= \sum_{l=0}^{p}\binom{p}{l}(-1)^{p-l}\,M_{l;n}\,\bar{x}^{p-l}. \end{aligned} \tag{5}$$

Similarly, sample raw moments can be obtained from sample central moments:

$$\begin{aligned} M_{p;n} = \frac{1}{n}\sum_{k=1}^{n}x_k^p &= \frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x} + \bar{x})^p \\ &= \frac{1}{n}\sum_{k=1}^{n}\{(x_k - \bar{x}) + \bar{x}\}^p \\ &= \frac{1}{n}\sum_{k=1}^{n}\left\{\sum_{l=0}^{p}\binom{p}{l}(x_k - \bar{x})^l\;\bar{x}^{p-l}\right\}, \text{ by the binomial theorem} \\ &= \sum_{l=0}^{p}\binom{p}{l}\left\{\frac{1}{n}\sum_{k=1}^{n}(x_k - \bar{x})^l\right\}\bar{x}^{p-l}, \text{ bringing through the sum in } k \\ &= \sum_{l=0}^{p}\binom{p}{l}\,c_{l;n}\,\bar{x}^{p-l}. \end{aligned}$$

An alternative derivation is based on showing that the binomial transformation is an involution, following [41]. We need a preliminary results: for $0 \le m \le l \le p$:

$$\binom{p}{l}\binom{l}{m} = \binom{p}{m}\binom{p-m}{l-m}.$$

It is easy to establish this result by expanding both sides and cancelling $l!$ from the right side and

$(p-m)!$ from the left. Next, consider $\sum_{l=0}^{p}\binom{p}{l} c_{l;n}\,\bar{x}^{p-l}$ and substitute in for $c_{l;n}$ using (5) to get

$$\begin{aligned}
\sum_{l=0}^{p}\binom{p}{l} c_{l;n}\,\bar{x}^{p-l} &= \sum_{l=0}^{p}\binom{p}{l}\left\{\sum_{m=0}^{l}\binom{l}{m}(-1)^{l-m}\,M_{m;n}\,\bar{x}^{l-m}\right\}\bar{x}^{p-l} \\
&= \sum_{l=0}^{p}\sum_{m=0}^{l}\binom{p}{l}\binom{l}{m}(-1)^{l-m}\,M_{m;n}\,\bar{x}^{p-m} \\
&= \sum_{m=0}^{p}\sum_{l=m}^{p}\binom{p}{m}\binom{p-m}{l-m}(-1)^{l-m}\,M_{m;n}\,\bar{x}^{p-m}, \\
&\quad \text{by the preliminary result and swapping the order of the sum} \\
&= \sum_{m=0}^{p}\binom{p}{m}M_{m;n}\,\bar{x}^{p-m}\sum_{l=m}^{p}\binom{p-m}{l-m}(-1)^{l-m}, \text{ collecting terms} \\
&= \sum_{m=0}^{p}\binom{p}{m}M_{m;n}\,\bar{x}^{p-m}\sum_{l=0}^{p-m}\binom{p-m}{l}(-1)^{l}, \\
&\quad \text{starting the second summation variable } l \text{ at } 0.
\end{aligned}$$

When $m=p$, the second sum is just $(-1)^0=1$. When $m<p$, the second term is the binomial expansion of $(1-1)^{p-m}=0$. So, the only non-zero term on the right side of (6) occurs when $m=p$ and is $\binom{p}{p}M_{p,n}\,\bar{x}^{p-p}=M_{p,n}$, so that

$$M_{p;n}=\sum_{l=0}^{p}\binom{p}{l}c_{l;n}\,\bar{x}^{p-l}, \tag{6}$$

as required. We have effectively inverted the relationship $c_{p;n}=\sum_{l=0}^{p}\binom{p}{l}(-1)^{p-l}\,M_{l;n}\,\bar{x}^{p-l}$ to find (6) without using properties of sample moments.

Similar relationships exist between the raw and central moments of a random variable X.

Let $\mu=\mu_1=\mathrm{E}[X]=\int x\,f(x)\,dx$. Let $\mu'_p=\mathrm{E}[(X-\mu)^p]=\int(x-\mu)^p\,f(x)\,dx$. Then, using an argument analogous to the one that led to (5), we have

$$\begin{aligned}
\mu'_p &= \mathrm{E}\left[(X-\mu)^p\right] \\
&= \mathrm{E}\left[\sum_{l=0}^{p}\binom{p}{l}X^l\,(-\mu)^{p-l}\right] \text{ by the binomial theorem} \\
&= \sum_{l=0}^{p}\binom{p}{l}\mathrm{E}\left[X^l\right]\,(-\mu)^{p-l} \text{ by the linearity of } E \\
&= \sum_{l=0}^{p}\binom{p}{l}(-1)^{p-l}\,\mu_l\,\mu^{p-l}.
\end{aligned}$$

We can easily modify the derivation of (6) to the case of random variables to get $\mu_p=\sum_{l=0}^{p}\binom{p}{l}\,\mu'_l\,\mu^{p-l}$.

B Appendix B: Properties of the Lehmer Mean

We prove Propositions 1 and 2 about the Lehmer mean.

Proof of Proposition 1: First, we find the derivative of $L_{p;n}$ with respect to p. To do this, note that

$\frac{d}{dp}x^p = x^p \log(x)$. Now,

$$\begin{aligned}
\frac{d}{dp} L_{p;n} &= \frac{d}{dp}\left(\frac{\sum_{k=1}^{n} x_k^p}{\sum_{k=1}^{n} x_k^{p-1}}\right) \\
&= \frac{\sum_{k=1}^{n} x_k^{p-1} \frac{d}{dp}\sum_{k=1}^{n} x_k^p - \sum_{k=1}^{n} x_k^p \frac{d}{dp}\sum_{k=1}^{n} x_k^{p-1}}{\left(\sum_{k=1}^{n} x_k^{p-1}\right)^2} \text{ by the quotient rule} \\
&= \frac{\sum_{k=1}^{n} x_k^{p-1} \sum_{k=1}^{n} x_k^p \log(x_k) - \sum_{k=1}^{n} x_k^p \sum_{k=1}^{n} x_k^{p-1} \log(x_k)}{\left(\sum_{k=1}^{n} x_k^{p-1}\right)^2} \text{ by the result just stated.}
\end{aligned}$$

Next, to establish that the derivative of $L_{p;n}$ with respect to p is positive, we need to show that the numerator of this expression is always positive, since the denominator is always positive. We write the sums of the numerator using different indices and argue as follows:

$$\begin{aligned}
&\sum_{i=1}^{n} x_i^{p-1} \sum_{j=1}^{n} x_j^p \log(x_j) - \sum_{i=1}^{n} x_i^p \sum_{j=1}^{n} x_j^{p-1} \log(x_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^{p-1} x_j^p \log(x_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^p x_j^{p-1} \log(x_j) \\
&= \sum_{i\neq j} x_i^{p-1} x_j^p \log(x_j) - \sum_{i\neq j} x_i^p x_j^{p-1} \log(x_j), \\
&\quad \text{because the } \textstyle\sum_{i=1}^{n} x_i^{2p-1} \log(x_i) \text{ terms cancel} \\
&= \sum_{i<j} x_i^{p-1} x_j^p \log(x_j) + \sum_{i>j} x_i^{p-1} x_j^p \log(x_j) - \sum_{i<j} x_i^p x_j^{p-1} \log(x_j) - \sum_{i>j} x_i^p x_j^{p-1} \log(x_j) \\
&= \sum_{i<j} x_i^{p-1} x_j^p \log(x_j) + \sum_{i<j} x_j^{p-1} x_i^p \log(x_i) - \sum_{i<j} x_i^p x_j^{p-1} \log(x_j) - \sum_{i<j} x_j^p x_i^{p-1} \log(x_i) \\
&\quad \text{swapping } i \text{ and } j \text{ in the second and fourth terms} \\
&= \sum_{i<j} (x_i x_j)^{p-1} \left\{x_j \log(x_j) + x_i \log(x_i) - x_i \log(x_j) - x_j \log(x_i)\right\} \\
&= \sum_{i<j} (x_i x_j)^{p-1} \left\{(x_j - x_i)\ (\log(x_j) - \log(x_i))\right\}. \qquad (7)
\end{aligned}$$

Expression (7) is always positive: if $x_j > x_i$, then $\log(x_j) > \log(x_i)$, and the product $(x_j - x_i)\ \{\log(x_j) - \log(x_i)\} > 0$; similarly , the terms $x_j - x_i$ and $\log(x_j) - \log(x_i)$ will both be negative – and hence their product will be positive – when $x_j < x_i$. Since $x_1, \ldots, x_n$ are not all equal, there will be a positive contribution to (7). Hence, $dL_{p;n}/dp > 0$ and $L_{p;n}$ is a monotonically increasing function of p.

Proof of Proposition 2: Let us assume that there is just one largest data value: $x_{(1)} \leq \cdots \leq x_{(n-1)} <$

$x_{(n)}$. Then,

$$\begin{aligned} L_{p;n} &= \frac{x_{(1)}^p + \cdots + x_{(n-1)}^p + x_{(n)}^p}{x_{(1)}^{p-1} + \cdots + x_{(n-1)}^{p-1} + x_{(n)}^{p-1}} \\ &= \frac{x_{(n)}^p}{x_{(n)}^{p-1}} \times \frac{\left(\frac{x_{(1)}}{x_{(n)}}\right)^p + \cdots + \left(\frac{x_{(n-1)}}{x_{(n)}}\right)^p + \left(\frac{x_{(n)}}{x_{(n)}}\right)^p}{\left(\frac{x_{(1)}}{x_{(n)}}\right)^{p-1} + \cdots + \left(\frac{x_{(n-1)}}{x_{(n)}}\right)^{p-1} + \left(\frac{x_{(n)}}{x_{(n)}}\right)^{p-1}} \quad (8) \\ &= x_{(n)} \times \frac{\beta_1^p + \cdots + \beta_{n-1}^p + 1}{\beta_1^{p-1} + \cdots + \beta_{n-1}^{p-1} + 1}, \text{ in which } 0 < \beta_i = \tfrac{x_{(i)}}{x_{(n)}} < 1, i = 1, \ldots, n-1 \quad (9) \\ &\to x_{(n)} \text{ as } p \to \infty, \end{aligned}$$

since $\beta^q \to 0$ as $q \to \infty$ when $0 < \beta < 1$. It is easy to extend this argument to the case when there is more than one maximum data value, because $L_{p;n}$ is a monotonically increasing function of p, $L_{p;n} \nearrow x_{(n)}$ as $p \nearrow \infty$. Similarly, $\lim_{r,s\to\infty} G_{r,s;n} = x_{(n)}$; see [42].

C Appendix C: Approximating and Protecting the Second Largest Value $x_{(n-1)}$:

We present an argument that suggests that $\hat{x}_{(n-1)}^{(p,p')} \to x_{(n-1)}$ as $p, p' \to \infty$, provided $p > p'$.
By considering the leading terms in (9), we can approximate $L_{p;n}$ as

$$\begin{aligned} L_{p;n} \approx x_{(n)} \frac{1+\beta_{n-1}^p}{1+\beta_{n-1}^{p-1}} &= x_{(n)} \left(1+\beta_{n-1}^p\right)\left(1+\beta_{n-1}^{p-1}\right)^{-1} \\ &\approx x_{(n)} \left(1+\beta_{n-1}^p\right)\left(1-\beta_{n-1}^{p-1}\right) \\ &\approx x_{(n)} \left(1-\beta_{n-1}^{p-1}\right), \quad (10) \end{aligned}$$

since $(1+x)^{-1} = 1 - x + x^2 - \cdots$ for $0 < x < 1$ and by retaining the lowest power of β_{n-1}.
Let $0 < \theta_i = x_{(i)}/x_{(n-1)} < 1$, $i = 1, \ldots, n-1$. Then, using an approach similar to the one that led to (8), we have that

$$S_{p';n} = x_{(n-1)}^{p'} \left\{ \theta_1^{p'} + \cdots + \theta_{n-2}^{p'} + 1 + \left(\frac{1}{\beta_{n-1}}\right)^{p'} \right\},$$

since $\beta_{n-1} = x_{(n-1)}/x_{(n)}$, from which we obtain, using (10):

$$\begin{aligned} S_{p';n} - (L_{p,n})^{p'} &\approx x_{(n-1)}^{p'} \left\{ \theta_1^{p'} + \cdots + \theta_{n-2}^{p'} + 1 + \left(\frac{1}{\beta_{n-1}}\right)^{p'} - \left(\frac{1}{\beta_{n-1}}\right)^{p'} \left(1-\beta_{n-1}^{p-1}\right)^{p'} \right\} \\ &\approx x_{(n-1)}^{p'} \left\{ \theta_1^{p'} + \cdots + \theta_{n-2}^{p'} + 1 + \left(\frac{1}{\beta_{n-1}}\right)^{p'} - \left(\frac{1}{\beta_{n-1}}\right)^{p'} \left(1-p'\beta_{n-1}^{p-1}\right) \right\} \\ &\quad \text{by a binomial approximation} \\ &= x_{(n-1)}^{p'} \left(\theta_1^{p'} + \cdots + \theta_{n-2}^{p'} + 1 + p'\beta_{n-1}^{p-p'-1} \right). \end{aligned}$$

Hence,

$$\hat{x}_{(n-1)}^{(p,p')} = \frac{S_{p';n} - (L_{p;n})^{p'}}{S_{p'-1;n} - (L_{p;n})^{p'-1}} \approx \frac{x_{(n-1)}^{p'} \left(\theta_1^{p'} + \cdots + \theta_{n-2}^{p'} + 1 + p'\beta_{n-1}^{p-p'-1}\right)}{x_{(n-1)}^{p'-1} \left(\theta_1^{p'-1} + \cdots + \theta_{n-2}^{p'-1} + 1 + (p'-1)\beta_{n-1}^{p-p'}\right)},$$

which tends to $x_{(n-1)}$ as $p, p' \to \infty$, provided $p > p'$. For fast convergence, we require p, p' and $p - p'$ to be large. This can be achieved to a certain extent by setting $p' = p/2$ so that $p - p' = p/2$.

References

[1] Eurostat. Statistical confidentiality and personal data protection. European Commission Eurostat. https://ec.europa.eu/eurostat/web/microdata/statistical-confidentiality-and-personal-da Accessed 12 July 2023.

[2] Office for National Statistics. Policy for social survey microdata. Office for National Statistics. https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontro Census 2021. Accessed 12 July 2023.

[3] UK Statistics Authority. Code of practice for statistics. https://code.statisticsauthority.gov.uk/, 2023. Accessed 12 July 2023.

[4] A. Hundepool, J. Domingo-Ferrer, L. Franconi, S. Giessing, E. Schulte Nordholt, K. Spicer, and P.-P. de Wolf. *Statistical Disclosure Control*. Wiley, 2012.

[5] C. Skinner. Statistical disclosure risk: separating potential and harm. *International Statistical Review*, 80:349–368, 2012.

[6] F. Ritchie and M. J. Elliot. Principles- versus rules-based output statistical disclosure control in remote access environments. *International Association for Social Science Information Service and Technology*, 39(2):5, 2014.

[7] L. H. Cox, A. F. Karr, and S. K. Kinney. Risk-utility paradigms for statistical disclosure limitation: How to think but not how to act (with discussion). *International Statistical Review*, 79:160–183, 2011.

[8] J. Lei, A.-S. Charest, A. Slavkovic, A. Smith, and S. Fienberg. Differentially private model selection with penalized and constrained likelihood. *Journal of the Royal Statistical Society Series A*, 181:609–633, 2018.

[9] S. Polettini. Maximum entropy simulation for microdata protection. *Statistics and Computing*, 13:307–320, 2003.

[10] Statistics Netherlands. Mu-argus. Statistics Netherlands. https://research.cbs.nl/casc/mu.htmand https://github.com/sdcTools/muargus Accessed 12 July 2023.

[11] Statistics Netherlands. Tau-argus. Statistics Netherlands. https://research.cbs.nl/casc/tau.htmand https://github.com/sdcTools/tauargus Accessed 12 July 2023.

[12] M. Templ, A. Kowarik, and B. Meindl. Statistical disclosure control for micro-data using the R package sdcMicro. *Journal of Statistical Software*, 67(4):1–36, 2015.

[13] L. Franconi and J. Stander. A model-based method for disclosure limitation of business microdata. *The Statistician*, 51:51–61, 2002.

[14] L. Franconi and J. Stander. Spatial and non-spatial model-based protection procedures for the release of business microdata. *Statistics and Computing*, 13:295–305, 2003.

[15] S. Polettini, L. Franconi, and J. Stander. Model based disclosure protection. In Josep Domingo-Ferrer, editor, *Inference Control in Statistical Databases*, volume 2316 of *LNCS*, pages 83–96. Springer-Verlag, 2002.

[16] Office for National Statistics. Disclosure control: Best practice for applying disclosure control to data. Office for National Statistics. https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontro Accessed 12 July 2023.

[17] E. Griffiths, C. Greci, Y. Kotrotsios, S. Parker, J. Scott, R. Welpton, A. Wolters, and C. Woods. *Handbook on Statistical Disclosure Control for Outputs*. UK Data Service, 2019.

[18] N. Shlomo. Statistical disclosure limitation: new directions and challenges. *Journal of Privacy and Confidentiality*, 8(1), 2018.

[19] C. Carota, M. Filippone, and S. Polettini. Assessing Bayesian semi-parameteric log-linear models: an application to disclosure risk estimation. *International Statistical Review*, 90:165–183, 2022. https://onlinelibrary.wiley.com/doi/full/10.1111/insr.12471 Accessed 12 July 2023.

[20] Y. Rinott, C. M. O'Keefe, N. Sholmo, and S. Skinner. Confidentiality and differential privacy in the dissemination of frequency tables. *Statistical Science*, 33:358–385, 2018.

[21] P. Chebyshev. Sur les valeurs limites des intégrales. *Journal de Mathématiques Pures et Appliquées*, 19:157–160, 1874.

[22] R Curto. Truncated moment problems: An introductory survey. https://www.birs.ca/workshops/2019/19w5137/files/Curto.pdf, April 2019. Slides from Multivariable Spectral Theory and Representation Theory, BIRS Workshop. Accessed 12 July 2023.

[23] V. John, I. Angelov, A. A. Öncül, and D. Thévenin. Techniques for the reconstruction of a distribution from a finite number of its moments. *Chemical Engineering Science*, 62:2890–2904, 2007.

[24] N. Lebaz, A. Cockx, M. Spérandio, and J. Morchain. Reconstruction of a distribution from a finite number of its moments: A comparative study in the case of depolymerization process. *Computers and Chemical Engineering*, 84:326–337, 2016.

[25] J. R. Lewis, S. B. MacEachern, and Y. Lee. Bayesian restricted likelihood methods: conditioning on insufficient statistics in Bayesian regression (with discussion). *Bayesian Analysis*, 16(4):1393–1462, 2022.

[26] J. Burridge. Information preserving statistical obfuscation. *Statistics and Computing*, 13:321–327, 2003.

[27] D. H. Lehmer. On the compounding of certain means. *Journal of Mathematical Analysis and Applications*, 36:183–200, 1971.

[28] E. W. Weisstein. Sample raw moment. From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/SampleRawMoment.html Accessed 12 July 2023.

[29] E. W. Weisstein. Sample central moment. From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/SampleCentralMoment.html Accessed 12 July 2023.

[30] S. Simić. A simple proof of monotonicity for stolarsky and gini means. *Kragujevac J. Maths.*, 32:75–79, 2009.

[31] H. J. Kim, J. P. Reiter, and A. F. Karr. Simultaneous edit-imputation and disclosure limitation for business establishment data. *Journal of Applied Statistics*, 45(1):63–82, 2018.

[32] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2024.

[33] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[34] C. Czado. *Analyzing Depedent Data with Vine Copulas: A Practical Guide with R*. Springer, 2019.

[35] T. Nagler, U. Schepsmeier, J. Stoeber, E. Brechmann, B. Graeler, and T. Erhardt. *VineCopula: Statistical Inference of Vine Copulas*, 2023. R package version 2.5.0.

[36] J. Stander, L. Dalla Valle, C. Taglioni, B. Liseo, A. Wade, and M. Cortina-Borja. Analysis of paedi-

atric visual acuity using Bayesian copula models with sinh-arcsinh marginal densities. *Statistics in Medicine*, 38:3421–3443, 2019.

[37] J. Stander and M. Stander. Code for Using the Lehmer Mean to Assess Business Data Protection: Statistical Disclosure Control and the Truncated Moment Problem, 2023.

[38] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 2nd edition, 1984.

[39] E. W. Weisstein. Central moment. From MathWorld–A Wolfram Web Resource. `https://mathworld.wolfram.com/CentralMoment.html` Accessed 12 July 2023.

[40] E. W. Weisstein. Raw moment. From MathWorld–A Wolfram Web Resource. `https://mathworld.wolfram.com/RawMoment.html` Accessed 12 July 2023.

[41] Robert Israel. Proof that the binomial transform is involution. Mathematics Stack Exchange, 2017. `https://math.stackexchange.com/questions/2082924/proof-that-the-binomial-transform-is-i` Accessed 12 July 2023.

[42] P. S. Bullen. *Handbook of Means and their Inequalities*. Springer-Verlag, New York, 2003.