An Improved Formulation of the Disclosure Auditing Problem for Secondary Cell Suppression

Jacco Daalmans*, Ton de Waal**
* Statistics Netherlands, P. O. Box 24500, 2490 HA The Hague
E-mail: j.daalmans@cbs.nl
** Statistics Netherlands, P. O. Box 24500, 2490 HA The Hague
E-mail: t.dewaal@cbs.nl

Abstract. Statistical agencies have to ensure that respondents' private information cannot be revealed from the tables they release. A well-known protection method is cell suppression, where values that provide too much information are left out from the table to be published. In a first step, sensitive cell values are suppressed. This is called primary suppression. In a second step, other values are suppressed as well to exclude that primarily suppressed values can be re-calculated from the values published in the table. This second step is called secondary cell suppression.

In this article we explain that the problem of checking whether a pattern of secondary cell suppressions is safe for release or not is generally described in a slightly inconsistent way in the literature. We illustrate with examples that the criteria that are often applied to judge whether a table can be safely published or not do not always give satisfactory results. Furthermore, we present a new criterion and explore some of its consequences. The new criterion is an extension of the well-known (p,q)-prior-posterior rule. This extension is for aggregations of suppressed cells for which a value can be derived from the table. Finally, we provide a method to apply the new criterion in practice.

1 Introduction

Before publishing a magnitude table a statistical agency has to ensure that the privacy of the individual contributors to the table is not endangered. The privacy of individual entities, such as persons or businesses, may be endangered if data that are considered sensitive can be disclosed from the table to be released. Suppose, for instance, that one wants to release a table containing the turnover of enterprises cross-classified by branch of industry and region. If there is only one enterprise in the population with a certain combination of branch of industry and region, the turnover of this enterprise can be disclosed, simply by looking at the corresponding cell in the table. The turnover of enterprises is generally considered sensitive information, so this table cannot be released in its full form.

The aim of statistical disclosure control (SDC) is to prevent sensitive information on entities from being disclosed. Statistical disclosure control can be applied to both frequency count tables and magnitude tables. In this article we focus on tabular magnitude data only.

The basis for most SDC techniques for tabular magnitude data is a sensitivity measure for individual cells. Such a sensitivity measure determines whether a cell is sensitive or not, and hence whether its value must be censored or may be published. A table containing only non-sensitive cells is called safe; a table containing at least one sensitive cell is called unsafe. The statement that a table is safe cannot be seen in isolation from the sensitivity measure used; a table may be considered safe according to one rule, but unsafe according to another rule.

An unsafe table has to be protected against statistical disclosure. A well-known method to protect an unsafe table is cell suppression, where the value of one or more cells is deleted. Instead of publishing the value of a cell, a special character, such as a cross (×), is published.

This article deals with the problem of finding out whether a table with suppressed cell values is safe to be released or not. In the literature, e.g. [5] and [10], this problem is usually referred to as the disclosure auditing problem. This problem is distinct from the problem of finding the cell values of a table that have to be suppressed, although both problems are closely connected. Namely, the problem of finding the cell values that have to be suppressed may be formulated as: find suppressed cells that minimize a certain objective function, such as the total suppressed value, the number of suppressed cell values, or the number of contributions to the suppressed cells, subject to the condition that the resulting table is safe.

There are several classes of sensitivity measures. The best-known sensitivity measures are the prior/posterior rule, where a cell is considered sensitive if one of the contributions can be calculated to within a certain percentage of its value, and the dominance rule, where a cell is considered sensitive if a substantial part of its value is due to only a few contributors. Section 2 discusses sensitivity measures for individual cells.

Once the sensitive cells have been determined, their values are suppressed. This is referred to as primary cell suppression. In addition, usually a number of non-sensitive cells have to be suppressed in order to prevent the possibility of re-calculation of the suppressed sensitive cell values from the values published in the table. This phenomenon is called secondary cell suppression. The secondary cell suppression problem amounts to finding a good set of secondarily suppressed cells. Section 3 describes the basic form of the usual formulations of the secondary cell suppression problem. In this section it will be explained that these standard descriptions of the secondary cell suppression problem, as given in the literature, e.g. [4], [7], [9], [11], [17] and section 4 in [5], are generally slightly inconsistent. Even the second author of the present article is himself guilty in this respect (see [30] and [31]). Implicitly, Sande [23], [24] and [25] already gave a nearly accurate description of the cell secondary suppression problem in the 1970's. In our point of view, his definition contains a minor flaw, however.

Cox [5] and Giessing [12] briefly describe the secondary suppression problem quite accurately. They refer to the resulting problem as the multicell disclosure problem. They, however, do not attempt to describe how this multicell disclosure problem should be handled and do not provide any details. In fact, Cox [5] notes that the multicell disclosure problem "is an unmanageable problem if approached directly". Salazar-González [22] refers to the problem as the multi-attacker cell suppression problem, and makes an attempt to solve it. We return to the approach by Salazar-González later in this article.

The material contained in Sections 2 and 3 are for a substantial part based on the papers referred to above as well as some others, such as [19], [20] and [21].

In Section 4 we give a formulation for the disclosure auditing problem that is correct in our point of view. This formulation includes a criterion to judge whether a table is sufficiently protected. This criterion could – and should in our opinion – also form the basis of methods for secondary cell suppression. As far as we are aware this is the first time that the new formulation for the disclosure auditing problem is given in such detail.

In the literature the discussion of the disclosure auditing problem and cell suppression for magnitude tables is usually restricted to tables with nonnegative contributions only. Furthermore, the contributors to different cells of the table are assumed not to co-operate, which could happen if they are all enterprises of the same holding. The new formulation of the disclosure auditing problem that is given in Section 5 does allow for negative contributions and holdings.

In Section 6 a formulation of a Mixed Integer Programming (MIP) problem is given that can be used to apply our idea for the disclosure auditing problem in practice. That is, by solving a MIP it can be judged whether some table is sufficiently protected. This MIP formulation requires the sensitivity measure for individual cells to be extended to aggregations of cells. Therefore, Section 5 first discusses this extension. Section 7 concludes the article with a brief discussion.

2 Sensitivity measures for individual cells

As mentioned in the Introduction, one generally uses a sensitivity measure to determine whether individual cells in a table are sensitive or not. A widely used sensitivity measure is the (*p*,*q*)-prior/posterior rule.

Rule 1 (The (*p*,*q*)-prior/posterior rule for separate cells). This sensitivity measure is based on two nonnegative parameters, p and q with $p < q$. It is assumed that, prior to the publication of the table, everyone can estimate the contribution of each contributor to the table to within q percent. A cell is considered sensitive if the contribution of an individual contributor to that cell can be estimated, for instance by one of the other contributors to that cell, to within p percent after (posterior to) publication of the table.

This criterion is called the prior/posterior rule because it involves both prior knowledge (q-percent estimates of the individual contributions) and posterior knowledge provided by the published table. In practical applications, one reformulates the general criterion (Rule 1) into an operational one.

For convenience we assume in this section that all contributions to the table are nonnegative, and later relax this condition in Section 5. We denote the number of contributors to the table by R. The number of cell values will be denoted by N_C. We denote a cell value by x_i, $i = 1,\ldots,N_C$ and the contribution of contributor r to x_i by x_i^r ($i = 1,\ldots,N_C$; $r = 1,\ldots,R$). No distinction is made between a respondent that does not contribute to a cell and a contribution of zero, in both cases $x_i^r = 0$. In practice, most contributors only contribute to one cell, so most x_i^r will be equal to zero. We can write $x_i = \sum_{r=1}^{R} x_i^r$. We let $x_i^{[r]}$, ($i = 1,\ldots,N_C$, $r = 1,\ldots,R$) represent decreasingly ordered contributions to cell x_i, i.e. $x_i^{[1]} \geq x_i^{[2]} \geq \ldots \geq x_i^{[R]} \geq 0$.

To reformulate the (*p*,*q*)-prior/posterior rule into an operational criterion we consider the accuracy with which contribution t can be estimated by a contributor s. This so-called intruder (or attacker) s can calculate an upper bound for x_i^t as follows. Knowing the cell value $x_i = \sum_{r=1}^{R} x_i^r$, he subtracts his own contribution

x_i^s and estimates for the other contributions from this cell value. By assumption, the lower bound on the estimate for the contribution of contributor r amounts to $x_i^r - (q/100)x_i^r$, for $r \neq s$, t. This results in the following upper bound on x_i^t from the perspective of contributor s

$$U_s(x_i^t) = x_i^t + \frac{q}{100} \sum_{r \neq s,t} x_i^r .$$

The cell would be sensitive if $U_s(x_i^t) \leq x_i^t + (p/100)x_i^t$ for some s and t. It can be derived that all contributions are sufficiently protected, i.e.

$$\frac{q}{100} \sum_{r \neq s,t} x_i^r > \frac{p}{100} x_i^t \qquad \text{for all } s \text{ and } t$$

if the largest contribution is sufficiently protected for the respondent with the second largest contribution, that is if:

$$\frac{q}{100} \sum_{r \neq 1,2} x_i^{[r]} > \frac{p}{100} x_i^{[1]} . \tag{2.1}$$

This immediately follows from:

$$\frac{q}{100} \sum_{r \neq s,t} x_i^r \geq \frac{q}{100} \sum_{r \neq 1,2} x_i^{[r]}$$

and

$$\frac{p}{100} x_i^{[1]} \geq \frac{p}{100} x_i^t .$$

Note that the second largest contributor to cell i – assuming his contribution to this cell is non-zero – can estimate the largest contribution to cell i more accurately than someone who does not contribute to cell i at all, as he can use his own contribution as additional information for making the estimate. To apply the (p,q)-prior/posterior rule in practice one therefore checks whether $U_2(x_i^{[1]}) > x_i^{[1]} + (p/100)x_i^{[1]}$, or equivalently whether $q\sum_{r=3}^{R} x_i^{[r]} > px_i^{[1]}$ for all i=1,…,N_C.

One can also consider the lower bound on contribution x_i^t of contributor t that can be derived by contributor s, and base a sensitivity criterion on that, instead of on the upper bound. This exactly leads to the same operational criterion, however.

Note that when all contributions to a cell equal zero, the cell is considered to be unsafe according to the operational criterion. This is in accordance with our intuitive feelings with respect to a sensitivity measure as everyone can derive the exact value of the contributors to the cell. For instance, if the cell gives the turnover of enterprises in a certain period, everyone can derive that none of the contributors to that cell had any turnover in that period, which can be highly sensitive information.

This sensitivity rule can be translated into a so-called sensitivity function $S_{p,q}$. The sensitivity function for the (*p*,*q*)-prior/posterior rule is

$$S_{p,q}(x_i) = px_i^{[1]} - q\sum_{r=3}^{R} x_i^{[r]}.$$

A cell *i* is sensitive, if and only if $S_{p,q}(x_i) \geq 0$.

A special case of the prior/posterior rule is the so-called *p*%-rule, where one assumes that all that an intruder knows is that the value of each contribution to the table is nonnegative. In terms of the (*p*,*q*)-prior/posterior rule this is more or less equivalent to assuming that $q = 100$, in the sense that the *p*%-rule and the (*p*,*q*)-prior/posterior rule with $q = 100$ lead to the same operational definition of a sensitive cell.

Another common way to determine whether a cell is considered sensitive is by means of a dominance rule. An (*n*,*k*)-dominance rule states that if the values of the data of a certain number of contributors, *n*, constitute more than a certain percentage, *k*, of the total value of the cell, then this cell is sensitive. The choice of *n* and *k* depends on the desired level of protection. Whereas in the past the dominance rule used to be the generally preferred rule, nowadays the prior/posterior rule, in particular its special case the *p*% rule, seems to be the generally preferred rule, because of its more appealing properties, e.g. [5], [14], [15] and [32]. In this article we do not consider the dominance rule anymore, but focus on the prior/posterior rule instead.

3 Secondary cell suppression

After the sensitive cells have been determined and suppressed, an intruder may still be able to calculate the original value of a sensitive cell through a close examination of the published cells and marginal totals. Consider for example Table 1 below, where all contributions are known to be nonnegative, and x_{11} and x_{21} are primary suppressions. It is easy to see that both x_{11} and x_{21} must have the value 100.

Table 1. A table with primary suppressions.

	C_1	C_2	C_3	Total
R_1	x_{11}	1	3	104
R_2	x_{21}	2	1	103
R_3	70	3	2	75
Total	270	6	6	282

In general we have to suppress some additional cells to adequately protect the sensitive cells. These additionally suppressed cells are the secondary suppressions.

Now we turn to the problem of an intruder who makes estimates of cell values. Consider Table 1 again. After entry $R_1 \times C_3$ and $R_2 \times C_3$ are chosen as secondary suppressions, Table 2 results. Both x_{11} and x_{21} cannot be calculated exactly. The following equations hold true

$$x_{11} + x_{13} = 103,$$
$$x_{13} + x_{23} = 4,$$
$$x_{11} + x_{21} = 200,$$
$$x_{21} + x_{23} = 101.$$

Combining the above equations with the nonnegativity of the contributions yields a range of possible values for x_{11}. We can deduce that $99 \le x_{11} \le 103$. The interval [99,103] is the so-called suppression interval of x_{11}.

Table 2. Primary and secondary suppressions.

	C_1	C_2	C_3	Total
R_1	x_{11}	1	x_{13}	104
R_2	x_{21}	2	x_{23}	103
R_3	70	3	2	75
Total	270	6	6	282

In order to avoid the possibility of calculating a suppressed sensitive cell in a (nonnegative) table too closely, one usually requires that the suppression interval for such a cell should be sufficiently wide. Example 1 below illustrates the procedure.

Example 1. Suppose we apply a p% rule with $p = 20$. We also demand that the suppression interval of each sensitive cell should have a width of at least 50% of the cell value. We apply these rules to Table 3.

Table 3. An unsafe table.

	I	*II*	*III*	Total
A	100	200	150	450
B	250	150	300	700
C	600	450	500	1150
Total	950	800	950	2700

Suppose the sensitive cells are *A*×*I* and *A*×*III*. We also suppose that to each of these two cells there is only one contributor. We protect the table by suppressing these cells, and a number of additional cell values. Suppose we obtain Table 4.

Table 4. "Protected version" of Table 3.

	I	*II*	*III*	Total
A	×	200	×	450
B	×	150	×	700
C	600	450	500	1150
Total	950	800	950	2700

To determine the suppression intervals of the suppressed cells for the suppression pattern in Table 4 we consider the following set of equations that can be derived from Table 4.

$$x_{11} + x_{13} = 250, \quad (3.1)$$

$$x_{21} + x_{23} = 550, \quad (3.2)$$

$$x_{11} + x_{21} = 350, \quad (3.3)$$

$$x_{13} + x_{23} = 450, \quad (3.4)$$

$$x_{11}, x_{13}, x_{21}, x_{23} \geq 0, \quad (3.5)$$

where x_{ij} denotes the value of the suppressed cell in row *i* and column *j*. For instance, the upper, respectively lower, bound on the suppression interval of x_{11} can be found by maximizing, respectively minimizing, x_{11} subject to (3.1) to (3.5). This is a simple linear programming problem, and can, for example, be solved by means of the simplex algorithm, e.g. [2]. In a similar way, the lower and upper bounds on the suppression intervals of the other suppressed cell values can be found. The suppression intervals are given in Table 5.

Table 5. Suppression intervals corresponding to Table 4.

	I	*II*	*III*	Total
A	*[0, 250]*	200	*[0, 250]*	450
B	*[100, 350]*	150	*[200, 450]*	700
C	600	450	500	1150
Total	950	600	950	2700

Neither of the two sensitive cell values can be determined to within 50% of its actual cell value. Table 4, i.e. the "protected version" of Table 3, is hence considered safe according to the applied criterion on the widths of the suppression intervals. □

The basic form of the standard formulation of the secondary cell suppression problem is: find the "best" set of secondary suppressions such that the suppression interval for each sensitive cell is sufficiently wide. Here "best" is defined by means of some target function such as: minimize the total suppressed value, minimize the number of suppressed cell values, or minimize the number of contributions to the suppressed cells. Depending on how the "best" set of secondary suppressions and on how "sufficiently wide suppression intervals for each sensitive cell" are precisely defined and operationalized one obtains various formulations for the secondary cell suppression problem. The standard criterion on the width of the suppression interval is described below.

Operational criterion 1 (The suppression width rule). The upper bound on the suppression interval has to be at least equal to that value for which the cell would be safe according to the sensitivity measure for individual cells (for instance the (p,q)-prior/posterior rule).

Implicitly, this criterion assumes that a sensitive cell is protected by suppressing some additional cells with relatively small contributions (see Example 4 in this article and, e.g., [3] and [5]).

Note that Operational criterion 1 is aimed at the protection of sensitive cell values, whereas the more stringent Rule 1 is directed at protecting the underlying contributions of the respondents. In practice, the operational criterion can be used as an approximation of Rule 1. In many cases this criterion is sufficient for protecting the sensitive cell at hand.

Example 2. Consider Table 6 below in which only cell $A \times I$ is assumed to be sensitive. Suppose that the largest contribution to this cell equals 155, and the second largest to 4. We suppress the value of cell $A \times I$ and demand that the upper bound on its suppression interval is at least equal to that value for which the cell would be safe according to the p%-rule with $p = 20$.

Table 6. An unsafe table.

	I	*II*	*III*	Total
A	160	380	340	880
B	40	80	60	180
C	610	800	270	1680
Total	810	1260	670	2740

According to the sensitivity measure, the upper bound on the suppression interval should be at least 190. Namely, when the upper bound on the suppression interval equals 190, the second largest contributor can derive that the upper bound on the largest contribution is 186 (=190–4). This upper bound exceeds the actual value (i.e. 155) by exactly 20%. An upper bound of 190 or more on cell $A \times I$ is achieved by the following suppression pattern.

Table 7. "Protected version" of Table 6.

	I	*II*	*III*	Total
A	×	×	340	880
B	×	×	60	180
C	610	800	270	1680
Total	810	1260	670	2740

It can easily be derived that the suppression interval of cell $A \times I$ is [80, 200]. The cell is sufficiently protected, since the upper bound implied by the interval (=200) exceeds the critical value (=190). □

However, if all individual cells of a table are sufficiently protected, according to Rule 1, and the table as a whole is safe on the basis of Operational criterion 1, the respondents' contributions are not necessarily sufficiently protected according to Rule 1, as will be shown in Example 3.

Example 3. As in Examples 1 and 2, we will use a p% rule with $p = 20$. We return to Table 4, i.e. the "protected version" of Table 3. In this table the cell values $A \times I$ and $A \times III$ have been suppressed. To each of these two cells there is only one contributor. One can derive the total value of the cells $A \times I$ and $A \times III$. The value of this "ad-hoc" cell, or aggregation, is 250. Each of the contributors to $A \times I$ and $A \times III$ can easily derive the contribution of the other. Thus, the "protected version" of Table 3 is not protected at all, according to Rule 1! Note that this conclusion does not depend on the required size of the suppression interval. □

Sande [28] calls the above phenomenon an "ad-hoc roll up", and gives a number of examples in publications of various statistical offices. From Example 3 it is clear that – if one wants to use the concept of suppression intervals – tables in which contributions of respondents may be recalculated exactly by other respondents can sometimes be considered safe! Accordingly, there is a need for a better operational criterion. In Section 4 a new criterion will be given, that is indeed better from our point of view.

In Example 3 we considered an aggregation of cells for which the exact value can be derived. The set of explicit aggregations, i.e. the aggregations that can be directly read off from the "protected" table, will be denoted by

$$\sum_{i=1}^{N_S} a_{ik} x_i = b_k \quad \text{for } k = 1,\ldots,K, \tag{3.6}$$

where N_S is the number of suppressed cells, and K denotes the number of explicit aggregations. An explicit aggregation (3.6) is a linear combination of suppressed cells of which the value can be derived from one row or column of a table. The coefficients of this linear combination are denoted by a_{ik} and b_k in (3.6). Usually $a_{ik} = 0$ or $a_{ik} = 1$, but this is no strict limitation.

For instance, for Table 4 the set of equations (3.6) is given by (3.1) to (3.4). The aggregations and their total values that can be derived from (3.6) are given by

$$\sum_{k=1}^{K} \mu_k \left(\sum_i a_{ik} x_i \right) = \sum_{k=1}^{K} \mu_k b_k ,$$

where the μ_k $(k = 1,\ldots,K)$ are coefficients. These coefficients may be positive or negative. For instance: by subtracting (3.3) from (3.1), one obtains the aggregation:

$$x_{13} - x_{21} = -100. \tag{3.7}$$

An aggregation j is defined by its coefficients $(\mu_1^j, \mu_2^j, \ldots, \mu_K^j)$. For notational convenience we will write the coefficients of aggregation j as $\lambda_i^j = \sum_{k=1}^{K} \mu_k^j a_{ik}$.

Aggregations are defined up to a constant factor, i.e. if we have an aggregation

$$\sum_{k=1}^{K} \mu_k \left(\sum_i a_{ik} x_i \right) = \sum_{k=1}^{K} \mu_k b_k ,$$

then

$$\sum_{k=1}^{K} (\alpha\mu_k) \left(\sum_i a_{ik} x_i \right) = \sum_{k=1}^{K} (\alpha\mu_k) b_k ,$$

is basically the same aggregation for $\alpha \neq 0$. This allows us to scale the $(\mu_1^j, \mu_2^j, \ldots, \mu_K^j)$. By dividing the μ_k^j $(k = 1,\ldots,K)$ by the maximum of their absolute values, we can ensure that they lie between –1 and 1, i.e. that

$$-1 \le \mu_k^j \le 1 \qquad \text{for all } k = 1,\ldots,K.$$

4 A new criterion

The approach of Section 3 is to apply a (p,q)-sensitivity rule to individual cells and a criterion on the width of the suppression interval to judge the safety of aggregations. However, from a conceptual point of view separate cells and aggregations are the same: they are just a collection of contributions of which the total value is known. From this perspective the approach of using a sensitivity rule for separate cells and a suppression interval criterion for aggregations is inconsistent. It is more logical that cells and aggregations of cells should be subjected to the same sensitivity rule. This leads to Operational criterion 2 below.

Operational criterion 2. A table is safe if and only if all aggregations of suppressed cells are safe, on the basis of the same sensitivity rule that is applied to separate cells.

In the literature sensitivity measures are defined for separate cells only. However, in order to apply Operational criterion 2, sensitivity measures have to be extended to aggregations. The extension for the (p,q)-sensitivity rule will be formally defined in Section 5.

Below the sensitivity rule for the most simple form of aggregations will be illustrated: sums of cells of which the value can be derived from one row or column of the table.

Example 4. We continue with Example 2 and again use the p% rule with $p = 20$. The largest two contributions to cell $A \times I$ in Table 6 equal 155 and 4, respectively. Suppose that the largest two contributions to cell $B \times I$ equal 28 and 10. Cell $A \times I$ is hence indeed sensitive and cell $B \times I$ non-sensitive (according to the p% rule with $p = 20$). If we consider the suppression pattern of Table 7, we can merge cells $A \times I$ and $B \times I$ into one imaginary cell. This cell has a total value of 200. The respondent that makes the second largest contribution to this aggregation, with value 28, can derive that the largest contribution to the aggregation is at most $200 - 28 = 172$. This is within 20% of the actual value (=155) of the largest contribution. According to the extended p%-rule, that will be formally described in Section 5, the "protected version" of Table 6, would hence be unsafe. □

The idea applied to Example 4 is that suppressed cells that appear in one row or column of a table are seen as one imaginary cell and the sensitivity rule is applied to this cell. On the basis of this rule the table from Example 2 would be unsafe, since one of the contributions can be recalculated too accurately (i.e. to within 20% of its actual value). However, on the basis of Operational criterion 1 the table would be classified as safe (see Section 3).

In Operation criterion 1 the implicit assumption is made that the attacker contributes to the same cell as the contribution under attack. The minimum width of the suppression interval of one particular cell only depends on the sensitivity rule that is applied to that cell. However, an attacker that contributes to a different cell may derive a lower upper bound than an attacker that contributes to the same cell, by using an aggregation that involves both cells. This may especially occur if the second largest contribution to the sensitive cell is less than the largest contribution to another suppressed cell within the same aggregation.

As a result Operational criterion 2 is more stringent than Operational criterion 1 in some cases (e.g. in Example 4). The opposite may also occur, as will be shown in Example 5.

Example 5. Consider Table 8 below. Suppose that three respondents contribute to cell $A \times I$ and three respondents contribute to cell $B \times I$. The contributions to $A \times I$ amount to 1000, 500 and 100 and the contributions to $B \times I$ to 100, 30 and 20. According to the p% rule, with $p = 20$, each upper bound on $A \times I$ should be at least 1700 (since 1700 – 500 = 1200 = 1000 + 20%×1000). The largest contributor to $B \times I$ can derive an upper bound on $A \times I$ of 1650 (=3750 – 2000 – 100). Hence, Table 8 is unsafe according to Operational criterion 1.

Table 8. A table with suppressions

	I	*II*	Total
A	×	×	2500
B	×	×	2500
C	2000	1000	3000
Total	3750	4250	8000

However, on the basis of Operational criterion 2 the table would be classified as safe. For this criterion the cells $A \times I$ and $B \times I$ are combined into one imaginary cell, with a value of 1750 and contributions of 1000, 500, 100, 100, 30 and 20. The second largest contributor to this combined cell can derive an upper bound of 1250 (1750 – 500) for the largest contribution, which exceeds its true value of 1000 by more than 20%. □

Operational criterion 1 says that the value of a primary suppressed cell may not be estimated too closely. As mentioned before, this criterion is at the cell level. Operational criterion 2, however, is at the level of the underlying contributions. In Example 5 it is shown that an attacker who can derive a close bound on a cell value, cannot always approximate its underlying contributions with a similar high precision. Thus, in some cases Operational criterion 1 is more restrictive than Operational criterion 2.

Although Operational criterion 1 is not always sufficient, it does give a good approximation for Operational criterion 2. In several cell suppression software packages, such as CONFID [19], [20] and [21], ACS [26] and [27] and τ-ARGUS [15], this approximation is used. The problem with the standard formulation, i.e. Operational criterion 1, has been acknowledged by others besides us, such as by Giessing, [12] and [13], who notes that implicit aggregates of cells may be sensitive. Giessing [12] notes that the sensitivity measure should also be applied to any linear combination of suppressed cells within a row or column of the table, but does not consider combinations of suppressed cells that are not within the same row or column. Moreover, she focuses on the situation where one has singletons, i.e. cells with one contribution, in the table, whereas we also consider cells with more contributions.

Most sensitivity measures, such as the (p,q)-prior/posterior rule and the (n,k)-dominance rule, as described in literature only make sense for sums of cells. Therefore, to apply our definition the sensitivity measure used has to be extended in order to make sense for more general combinations of cells, for example for differences of cells, such as (3.7). This extension of the (p,q)-prior/posterior rule will be made in the next section. The minor flaw in Sande's formulation referred to in the Introduction is that in his formulation he only considers aggregations with positive coefficients, whereas also aggregations with negative coefficients occur.

Cox [5] and Giessing [12] seem to use Operational criterion 2 when they refer to what they call the multicell disclosure problem. They do not examine, or even mention, the extension of the sensitivity measure to general linear aggregations of cells. Cox notes that "the multicell is important because it is at this level that actual respondents, and not artificial cell totals, are being protected for unauthorized disclosure". He also notes that "all current methods do not protect against multicell disclosure".

Salazar-González [22] uses an approach similar to Operational criterion 1. For each attacker, lower and upper bounds on all suppressed cells are assumed. The attacker knows that the true value of each suppressed cell value lies between these bounds. Different attackers are assumed to know different lower and upper bounds, reflecting different levels of knowledge of the true value of the suppressed cells. A table is considered safe, if the attackers cannot approximate a primary suppressed cell value "too closely". Salazar-González [22] assumes that the difference between the upper (or lower) bound and the true value should be at least as large as some protection level that has to be specified beforehand. For each combination of an attacker and a suppressed cell different protection levels may be used. He does not specify the meaning of "too closely", however.

The criterion that is used by Salazar-González [22] is on the cell level rather than on the level of the underlying contributions. By choosing the protection levels for the cell values in some appropriate way, it might be possible to protect their underlying contributions as well, according to the sensitivity measure used. As we already mentioned, Salazar-González does not discuss the topic of choosing appropriate values for the protection levels.

5 Extending the sensitivity measure

In this section the extension of the (p,q)-prior/posterior rule to aggregations, i.e. Operational criterion 2, will be formally described. From this point on negative contributions and holdings will be allowed.

As in Section 3 aggregations, i.e. linear combinations of suppressed cells with a known value, will be denoted by $X_j = \sum_{i=1}^{N_S} \lambda_i^j x_i$, where λ_i^j (i =1,…,N_S) are coefficients and the superscript j is used to identify the aggregation. Here, X_j will be used both for the definition of an aggregation, in terms of its constituent parts, as well as the value at the right-hand side of this aggregation. We hope that it will be clear from the context whether the definition of the aggregation or its value is meant. The contribution of respondent r, r = 1,…,R to aggregation X_j will be denoted by the linear combination $X_j^r = \sum_{i=1}^{N_S} \lambda_i^j x_i^r$. So, this contribution is defined as a linear combination of the contributions of contributor r to the underlying cell values.

Further, the *absolute contribution* of a respondent r, (r = 1,…,R) to an aggregation X_j will be defined as the linear combination $X_{jA}^r = \sum_{i=1}^{N_S} | \lambda_i^j x_i^r |$. So, absolute contributions can be obtained from contributions by replacing each term, consisting of a contribution multiplied by its coefficient, by the corresponding absolute value. Note that the value of an absolute contribution is nonnegative. For later reference we define the *absolute aggregation* X_{jA} as

$$X_{jA} = \sum_{r=1}^{R} X_{jA}^{r} = \sum_{r=1}^{R} \sum_{i=1}^{N_S} | \lambda_i^j x_i^r |.$$

Just as for the sensitivity measure for single cell values, a sensitivity measure for an aggregation X_j can be derived by computing an upper (or lower) bound on the contribution to X_j of a specific contributor from the perspective of another contributor[1]. Suppose a contributor s wants to approximate the value of the contribution X_j^t of contributor t. In order to check whether X_j^t is adequately protected for contributor s, we derive an upper bound on the value of X_j^t from the perspective of contributor s. This upper bound is obtained by subtracting the value of X_j^s and lower bounds for the values of X_j^r, $r \neq s,t$ from X_j, i.e. the value at the right-hand side of the aggregation. The lower bound $L_s(X_j^r)$ on the contribution of respondent r to aggregation j from the perspective of respondent s is

$$L_s(X_j^r) = (1 - \frac{q}{100}) \sum_{i:\lambda_i^j X_i^r > 0} \lambda_i^j X_i^r + (1 + \frac{q}{100}) \sum_{i:\lambda_i^j X_i^r < 0} \lambda_i^j X_i^r = X_j^r - \frac{q}{100} X_{jA}^r.$$

This results in the following upper bound $U_s(X_j^t)$ for the value of X_j^t from the perspective of contributor s

$$U_s(X_j^t) = X_j^t + \frac{q}{100} \sum_{r \neq s,t} X_{jA}^r . \quad (5.1)$$

Operational criterion 2 says that X_j^t is not sufficiently protected for contributor s if and only if

$$U_s(X_j^t) \leq X_j^t + \frac{p}{100} X_{jA}^t . \quad (5.2)$$

Combining the equation (5.1) with the inequality (5.2) yields that the contribution of contributor t to aggregation X_j is adequately protected for contributor s if and only if

$$q \sum_{r \neq s,t} X_{jA}^r > p X_{jA}^t . \quad (5.3)$$

[1] A sensitivity measure for aggregations may also be based on alternative ideas. Such alternative sensitivity measures for aggregations are not explored in this paper. For more details on these alternative measure we refer to Daalmans [6].

Analogous to (2.1), one can derive that all contributions to X_j are sufficiently protected, i.e.

$$q\sum_{r\neq s,t} X_{jA}^{r} > pX_{jA}^{t} \qquad \text{for every } s,\, t = 1,\ldots R,\; s \neq t\,.$$

if

$$q\sum_{r\neq 1,2} X_{jA}^{[r]} > pX_{jA}^{[1]},$$

which means that the largest contribution to X_{jA} is sufficiently protected for the contributor with the second largest contribution to X_{jA}. This leads to the following operational definition of sensitivity of aggregations. An aggregation X_j is sensitive if and only if

$$q\sum_{r=3}^{R} X_{jA}^{[r]} \leq pX_{jA}^{[1]}. \tag{5.4}$$

Example 6. We return to Tables 6 and 7 of Example 2. We use a (p,q)-prior/posterior rul with $p = 20$ and $q = 100$. In contrast to Example 2, this time we assume that cell $B\times II$ is also sensitive and has only one contribution with value 80. The suppression pattern in Table 7 consists of four suppressed cells: $A\times I$, $A\times II$, $B\times I$ and $B\times II$. These will henceforth be denoted by x_1, x_2, x_3, x_4, respectively. This notation differs from that of the earlier examples, but will be convenient when discussing Example 7. By subtracting the explicit aggregations

$$x_1 + x_3 = 200,$$
$$x_3 + x_4 = 120,$$

we obtain the implicit aggregation

$$x_1 - x_4 = 80. \tag{5.5}$$

Recall from Example 2 that $x_1 = 160$ and $x_4 = 80$. Thus, the translation of the implicit aggregation to its absolute form is

$$x_1 + x_4 = 240.$$

Since the largest contribution to x_1 amounts to 155 and there is only one contributor to x_4 with a contribution of 80, we obtain that $X_{jA}^{[1]}=155$, $X_{jA}^{[2]}= 80$ and $\sum_{r=3}^{R} X_{jA}^{[r]}= 5$ (i.e. 240 – 155 – 80). We obtain that $pX_{jA}^{[1]} - q\sum_{r=3}^{R} X_{jA}^{[r]} = 3100 - 500 = 2600$. Since this number is positive we conclude that the aggregation is sensive. Note that aggregation (5.5) allows the (only) contributor to x_4 to determine a lower bound of $80 + 80 - 5\times 1.2 = 154$ for the largest contribution of x_1, which is within 20% of its actual value. This demonstrates that this aggregation is indeed senstive. □

Note that the sensitivity rule in (5.4) resembles the sensitivity rule for individual cells. Namely, the sensitivity rule for aggregations can be obtained from the sensitivity rule for individual cell values by replacing each contribution to an individual cell by the absolute contribution to the aggregation.

We can also examine the lower bound on the value of X_j^t from the perspective of contributor s. As for individual cells, it can easily be seen that this leads to the same operational definition of the sensitivity measure as the one derived from the upper bound.

An appropriate sensitivity function for the (p,q)-prior/posterior rule extended to aggregations is

$$S_{p,q}^{a}(X_j) = pX_{jA}^{[1]} - q\sum_{r=3}^{R} X_{jA}^{[r]}.$$

An aggregation X_j is sensitive if and only if $S_{p,q}^{a}(X_j) \geq 0$. Note that

$$S_{p,q}^{a}(X_j) = S_{p,q}(X_{jA}).$$

The concept of absolute contributions has been introduced to extend the (p,q)-prior/posterior rule to aggregations that involve both positive and negative coefficients. Analogous to aggregations, the extended sensitivity rule can also be applied to separate cells. This leads to the following sensitivity function for individual cells

$$S_{p,q}^{a}(x_j) = px_{jA}^{[1]} - q\sum_{r=3}^{R} x_{jA}^{[r]}.$$

where $x_{jA}^{[r]}$ denotes the r largest contribution in absolute value to cell x_j. This is almost the same sensitivity function as the one given in Section 2; the only difference is that this definition allows for negative contributions.

For the (p,q)-prior/posterior rule it is implicitly assumed that an intruder knows the sign of all contributions. This assumption may not be realistic in practice, however. Therefore there may be a need for more appropriate sensitivity measures that can be applied to tables that have positive as well as negative contributions. We leave these measures open for further research.

Without any further assumptions the p%-rule becomes meaningless for tables with negative contributions as any contribution can then a priori assume any value. For any table only cells with at most two contributors would be sensitive, whereas all other cells would be non-sensitive, see [12]. We propose to operationalize the p%-rule for tables with negative contributions as the above (p,q)-prior/ posterior rule with q = 100[2].

Operational criterion 2 says that a table is sufficiently protected if all aggregations are sufficiently protected. One might fear, however, that there are situations, in which individual contributions to cell values involved in aggregations can be estimated too accurately, although all aggregations are non-sensitive. Fortunately, these situations cannot occur. This is the content of Theorem 1 below.

Theorem 1. If all contributions to an aggregation X_j are sufficiently protected according to the (p,q)-prior/posterior rule, all contributions to individual cell values involved in X_j are also sufficiently protected.

Proof. See the Appendix. □

As a consequence of Theorem 1, it is not necessary to apply a sensitivity measure to individual contributions to cells involved in aggregations.

It is not necessary to check whether all possible aggregations are sensitive or not: one only has to consider aggregates that involve at least one sensitive cell. Namely, Theorem 2 below says that aggregations that only involve non-sensitive cells are non-sensitive.

Theorem 2. Consider an aggregation $X_j = \sum_{i=1}^{N_S} \lambda_i^j x_i$. Then $S_{p,q}^a(x_i) < 0$ for every i with $\lambda_i^j \neq 0$ implies $S_{p,q}^a(X_j) < 0$.

Proof. See the Appendix. □

The property stated in Theorem 2 that an aggregation of non-sensitive cells is non-sensitive is an aspect of what is called subadditivity in the literature, e.g. [3] and [5]. Theorem 2 shows that this aspect of subadditivity holds true for our extended (p,q)-prior/posterior rule.

[2] An alternative way of operationalising the p%-rule for tables with negative contributions would be by assuming that an attacker a priori (only) knows whether each contributions is positive, zero or negative. This alternative way is not explored further in this article.

6 How to determine the most sensitive aggregation

In order to apply Operational criterion 2, and check whether a table is sufficiently protected by means of cell suppression, we apply a simple idea: we determine the most sensitive aggregation. If that aggregation is safe, the table is safe. To make this simple idea work, some technical "machinery" is required, however. In this section a mathematical model will be presented to implement our simple idea.

As before an aggregation will be denoted by

$$\sum_{k=1}^{K} \hat{\mu}_k \left(\sum_i a_{ik} x_i \right) = \sum_{k=1}^{K} \hat{\mu}_k b_k \ . \qquad (6.1)$$

where as usual K stands for the number of explicit aggregations, a_{ik} and b_k are coefficients of the explicit aggregations, and $\hat{\mu}_k$ are the variables in the mathematical model, i.e. these will be obtained as output. For ease of notation the superscript ^ will be used to denote all variables. The sum of all absolute contributions to some aggregation X will be denoted by $\hat{T}$, i.e.

$$\hat{T} = \sum_{r=1}^{R} X_A^r,$$

which can be rewritten as

$$\hat{T} = \sum_{r=1}^{R} \sum_{i=1}^{N_S} \left| \sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r \right|.$$

Further, the absolute contribution under attack will be denoted by $\hat{A}_1$ and the absolute contribution of the attacker will be expressed by $\hat{A}_2$.

Criterion (5.4) can be rewritten by: a cell is sensitive if

$$q(\hat{T} - \hat{A}_1 - \hat{A}_2) \le p\hat{A}_1,$$

or equivalently if

$$(p+q)\hat{A}_1 + q\hat{A}_2 - q\hat{T} \ge 0, \qquad (6.2)$$

We can, in principle, find 'the most' unsafe aggregation by maximizing

$$(p+q)\hat{A}_1 + q\hat{A}_2 - q\hat{T}. \qquad (6.3)$$

However, the technical problem has to be solved that it is not known beforehand which contributors yield the values of $\hat{A}_1$ and $\hat{A}_2$. Put differently, beforehand it is not clear which respondent will be the attacker and which contribution will be attacked.

Therefore the attacker and the contribution under attack will have to be identified by the model. For this purpose 0-1 variables $\hat{u}_r$ and $\hat{v}_r$ are introduced, that satisfy

$$\hat{u}_r = \begin{cases} 1 & \text{if respondent } r \text{ is under attack} \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

and

$$\hat{v}_r = \begin{cases} 1 & \text{if respondent } r \text{ is the attacker} \\ 0 & \text{otherwise} \end{cases} \tag{6.5}$$

A Mixed-Integer Programming (MIP) problem can be solved in order to find the most sensitive aggregation. The formulation of this optimization problem is

$$\text{Maximize } (p+q)\hat{A}_1 + q\hat{A}_2 - q\hat{T}, \tag{6.6}$$

$$\text{subject to } \hat{T} = \sum_{r=1}^{R}\sum_{i=1}^{Ns}(\hat{\xi}_{ir}^{+} + \hat{\xi}_{ir}^{-}), \tag{6.7}$$

$$\hat{\xi}_{ir}^{+} \geq \sum_{k=1}^{K}\hat{\mu}_k a_{ik} x_i^r \qquad \text{for } i=1,\ldots Ns,\ r=1,\ldots,R, \tag{6.8}$$

$$\hat{\xi}_{ir}^{-} \geq -\sum_{k=1}^{K}\hat{\mu}_k a_{ik} x_i^r \qquad \text{for } i=1,\ldots Ns,\ r=1,\ldots,R, \tag{6.9}$$

$$\hat{A}_1 \leq \sum_{i=1}^{Ns}\hat{\alpha}_{ir} + M(1-\hat{u}_r) \qquad \text{for } r=1,\ldots,R, \tag{6.10}$$

$$\hat{A}_2 \leq \sum_{i=1}^{Ns}\hat{\alpha}_{ir} + M(1-\hat{v}_r) \qquad \text{for } r=1,\ldots,R, \tag{6.11}$$

$$\hat{\alpha}_{ir} \leq \sum_{k=1}^{K}\hat{\mu}_k a_{ik} x_i^r + M(1-\hat{\beta}_{ir}) \qquad \text{for } i=1,\ldots Ns,\ r=1,\ldots,R, \tag{6.12}$$

$$\hat{\alpha}_{ir} \leq -\sum_{k=1}^{K}\hat{\mu}_k a_{ik} x_i^r + M\hat{\beta}_{ir} \qquad \text{for } i=1,\ldots Ns,\ r=1,\ldots,R, \tag{6.13}$$

$$\sum_{r=1}^{R}\hat{u}_r = 1, \tag{6.14}$$

$$\sum_{r=1}^{R}\hat{v}_r = 1, \tag{6.15}$$

$$\hat{v}_r + \hat{u}_r \leq 1 \qquad \text{for } r=1,\ldots,R, \tag{6.16}$$

$$\xi_{ir}^{+}, \hat{\xi}_{ir}^{-} \geq 0 \qquad \text{for } i=1,\ldots Ns,\ r=1,\ldots,R, \tag{6.17}$$

$$-1 \leq \hat{\mu}_k \leq 1 \qquad \text{for } k=1,\ldots,K, \tag{6.18}$$

$$\hat{u}_r, \hat{v}_r \in \{0,1\} \qquad \text{for } r=1,\ldots,R, \tag{6.19}$$

$$\hat{\beta}_{ir} \in \{0,1\} \qquad \text{for all } i,r, \tag{6.20}$$

where as before Ns is the number of suppressed cell values, R is the number of respondents, and M is some sufficiently large, positive number.

Below we first explain some of the symbols and the meaning of the constraints. Thereafter, we discuss the working of the model in more detail.

As an outcome of the model, optimal values of $\hat{\mu}_k$ are obtained. These values identify the aggregation $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$. According to constraint (6.18) $\hat{\mu}_k$ are restricted to values between -1 and +1. This constraint can be used in the model, since each aggregation can be scaled, such that all absolute values of $\hat{\mu}_k$ are smaller than or equal to 1 (see also the end of Section 3).

In (6.7) $\hat{T}$ is a variable whose optimal value is the sum over all absolute contributions to the aggregation. This property is achieved by constraints in (6.7), (6.8), (6.9) and (6.17), as will be explained below.

The optimal values for the 0-1 variables $\hat{u}_r$ and $\hat{v}_r$ define the respondent with the largest and second largest absolute contribution to the aggregation. As mentioned before, the first respondent is the one whose contribution is most vulnerable to disclosure and the second one is the most dangerous attacker. From now on these respondents will be denoted by "the attacked unit" and the "attacker".

The constraints (6.14) – (6.15) ensure that there is exactly one attacker and one unit under attack. Each respondent may be the attacker or the attacked unit, except that a respondent cannot be both at the same time. For this reason in (6.16) the values of $\hat{u}_r$ and $\hat{v}_r$ cannot be one for the same *r*.

The optimal values of $\hat{A}_1$ and $\hat{A}_2$ define the absolute contribution of the attacked unit and the attacker. As shown hereafter, this property is implied by the constraints (6.10)-(6.13).

Finally, the variables $\hat{\xi}_{ir}^-$, $\hat{\xi}_{ir}^+$, $\hat{\alpha}_{ir}$ and $\hat{\beta}_{ir}$ are incorporated in the model to deal with absolute values in the problem. As shown below, the optimal values of $\hat{\xi}_{ir}^- + \hat{\xi}_{ir}^+$ and $\hat{\alpha}_{ir}$ denote the contribution of respondent *r* to the most sensitive aggregation that can be attributed to cell x_i.

We now elaborate on the working of the model in more detail.

First, we demonstrate that the optimal value of $\hat{T}$ is the sum of all absolute contributions to the aggregation. As mentioned in (6.7) $\hat{T}$ is set equal to a sum of $\hat{\xi}_{ir}^+ + \hat{\xi}_{ir}^-$ over all i=1,…N_S and r=1,…,R. The inequalities in (6.8), (6.9) and (6.17) imply a lower bound on $\hat{\xi}_{ir}^+ + \hat{\xi}_{ir}^-$. We consider two cases, depending on the outcome of $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$. If, for some *i* and *r*, $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r \geq 0$, the tightest lower bound on $\hat{\xi}_{ir}^+$ is $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$, see (6.8), and the tightest lower bound on $\hat{\xi}_{ir}^-$ is 0, cf. (6.17).

In the other case, if $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r < 0$, a lower bound on $\hat{\xi}_{ir}^{+}$ is 0, cf. (6.17), and a lower bound on $\hat{\xi}_{ir}^{-}$ is $-\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$, cf. (6.9). In both cases a lower bound on $\hat{\xi}_{ir}^{+} + \hat{\xi}_{ir}^{-}$. is $\left|\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r\right|$. Substitution of this upper bound into (6.7) gives

$$\hat{T} \geq \sum_{r=1}^{R} \sum_{i=1}^{Ns} \left|\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r\right|,$$

indicating that $\hat{T}$ is at least as large as the sum of all absolute contributions to the aggregation. Due to the maximization of $-q\hat{T}$ in (6.6), the optimal value of $\hat{T}$ will be at its lower bound. This completes our explanation that the optimal value of $\hat{T}$ equals the sum of all absolute contributions to an aggregation.

We proceed to show that in the optimal solution $\hat{u}_r$ and $\hat{v}_r$ identify the attacked unit and the attacker and that the optimal values of $\hat{A}_1$ and $\hat{A}_2$ denote the absolute contributions of these two respondents.

Before showing this main result, we first need to derive an upper bound on $\sum_{i=1}^{Ns} \hat{\alpha}_{ir}$ in (6.10) and (6.11). In (6.12) and (6.13) two upper bounds on $\hat{\alpha}_{ir}$ are given. The value of the 0-1 variable $\hat{\beta}_{ir}$ determines which of the two bounds on $\hat{\alpha}_{ir}$ is the tightest.

Again, we consider two cases, depending on the outcome of $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$. If, for some i and r, $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r \geq 0$, the largest possible upper bound on $\hat{\alpha}_{ir}$ is achieved for $\hat{\beta}_{ir} = 1$ and this upper bound equals $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$.

In the other case, if $\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r < 0$, the largest possible upper bound is achieved for $\hat{\beta}_{ir} = 0$ and this bound equals $-\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r$.

In both cases the largest possible upper bound on $\hat{\alpha}_{ir}$ can be written as $\left|\sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r\right|$. As a consequence, we obtain the result that

$$\sum_{i=1}^{Ns} \hat{\alpha}_{ir} \le \sum_{i=1}^{Ns} \left| \sum_{k=1}^{K} \hat{\mu}_k a_{ik} x_i^r \right|,$$

i.e. that $\sum_{i=1}^{Ns} \hat{\alpha}_{ir}$ in (6.10) and (6.11) is bounded by the absolute contribution of respondent *r* to the aggregation.

We now return to the main explanation: the meaning of $\hat{u}_r$, $\hat{v}_r$, $\hat{A}_1$ and $\hat{A}_2$. The constraints in (6.10) and (6.11) impose *R* upper bounds on $\hat{A}_1$ and $\hat{A}_2$, one for each respondent. Note that the most restrictive upper bound in (6.10) amounts to $\sum_{i=1}^{N_S} \hat{\alpha}_{ir}$ where *r* is the respondent with $\hat{u}_r = 1$. Similarly, the most restrictive upper bound in (6.11) is $\sum_{i=1}^{N_S} \hat{\alpha}_{ir}$, for the respondent *r* with $\hat{v}_r = 1$. As mentioned before, $\sum_{i=1}^{N_S} \hat{\alpha}_{ir}$ is bounded by the absolute contribution to the aggregation of respondent *r*.

Due to the maximization of the objective function in (6.6) the optimal values of $\hat{A}_1$ and $\hat{A}_2$ are equal to their highest possible upper bounds. These are the absolute contribution to the aggregation of the respondent *r* for which $\hat{u}_r$ and $\hat{v}_r$ is one, respectively. The highest possible upper bounds are attained when both $\hat{u}_r$ and $\hat{v}_r$ are one for the respondent with the largest absolute contribution to the aggregation. However, this is not feasible, since $\hat{u}_r$ and $\hat{v}_r$ cannot be one for the same respondent *r*, due to the constraint in (6.16). Since (6.6) is maximized and $(p+q) > q$ it follows that $\hat{A}_1 \ge \hat{A}_2$ in the optimal solution. Thus, $\hat{u}_r$ will be one for the respondent with the largest absolute contribution to the aggregation (the attacked unit), and $\hat{v}_r$ will be one for the respondent with the second largest contribution (the attacker). As mentioned before, the optimal values of $\hat{A}_1$ and $\hat{A}_2$ are the absolute contributions to the aggregation of both respondents. Finally, note that, since the $\hat{\mu}_k$ can be scaled so they obtain finite values – in our case they are scaled so they lie between -1 and 1 (see (6.18) and the end of Section 3) – a finite optimum to (6.6) exists. This finishes our explanation of optimization model (6.6) to (6.20).

We are now ready for Theorem 3 below.

Theorem 3. A table is sufficiently protected according to the (*p*,*q*)-prior/posterior rule if and only if maximizing (6.6) subject to (6.7) to (6.20) yields a negative value.

Proof. By our criterion of a safe table, a nonnegative value of (6.3) can be obtained, if and only if a table involves a sensitive aggregation. Through the maximization of the objective function (6.6), the existence of a sensitive aggregation will lead to a nonnegative objective function value. Also the reverse holds true: a nonnegative objective function value of (6.6) means that the table contains a sensitive aggregation. This directly follows from the definition (6.3). □

The number of variables of the model can be very large. However, a reduced model can be applied to check whether some pre-defined set of contributions is sufficiently protected for some pre-defined selection of attackers. In this reduced model the variables $\hat{u}_r$ and $\hat{v}_r$ and the restrictions (6.10) – (6.13) are only defined for a subset of all "relevant" contributors, i.e. potential attackers and the contributors that are potentially being attacked.

Moreover, it is not necessary to define $\hat{\xi}_{ir}^{-}, \hat{\xi}_{ir}^{+}$ and $\hat{\alpha}_{ir}$ for some combination of cell x_i and respondent *r* if $x_i^r = 0$, since their optimal values will be zero.
Although this is not strictly necessary, it may be worthwhile to add the constraint

$$\hat{A}_1 \geq \hat{A}_2$$

to the formulation of the optimization model as this may speed up solving the problem.

Example 7. Consider Table 7 again. As in Example 6 the four suppressed cells: *A*×*I*, *A*×*II*, *B*×*I* and *B*×*II* will be denoted by x_1, x_2, x_3, x_4, respectively. The explicit aggregations of this table are:

1: $x_1 + x_2 = 540$,
2: $x_1 + x_3 = 200$,
3: $x_3 + x_4 = 120$,
4: $x_2 + x_4 = 460$.

All aggregations can be written as a linear combination of these four explicit aggregations, as in (6.1), with *K*=4. For instance, $\hat{\mu}_1 = 1,\ \hat{\mu}_2 = 0,\ \hat{\mu}_3 = 0,\ \hat{\mu}_4 = -1$, is the implicit aggregation:

$$1\,(x_1 + x_2) - 1\,(x_2 + x_4) = x_1 - x_4 = 540 - 460 = 80.$$

We suppose that there are 6 respondents. All nonzero contributions are given in Table 9.

Table 9. Individual contributions to Table 7

Cell	**Respondent**	**Contribution**
x_1	R_1	155
x_1	R_2	5
x_2	R_3	200
x_2	R_4	180
x_3	R_3	28
x_3	R_5	12
x_4	R_6	80

Note that in this example all individual cells are sensitive. Note also that respondent R_3 is a holding; it contributes to two cells, x_2 and x_3. We apply the p%-rule, with p = 20. As mentioned before, we operationalize the p%-rule, as a (p,q)-prior/posterior rule with q = 100. The objective function in (6.6) is

$$\text{Maximize } 120\hat{A}_1 + 100\hat{A}_2 - 100\hat{T}$$

The constraint (6.7) becomes:

$$\hat{T} = \hat{\xi}_{11}^{+} + \hat{\xi}_{11}^{-} + \hat{\xi}_{12}^{+} + \hat{\xi}_{12}^{-} + \hat{\xi}_{23}^{+} + \hat{\xi}_{23}^{-} + \hat{\xi}_{24}^{+} + \hat{\xi}_{24}^{-} + \hat{\xi}_{33}^{+} + \hat{\xi}_{33}^{-} + \hat{\xi}_{35}^{+} + \hat{\xi}_{35}^{-} + \hat{\xi}_{46}^{+} + \hat{\xi}_{46}^{-}.$$

Note that $\hat{\xi}_{ir}^{-}$ and $\hat{\xi}_{ir}^{+}$ are defined for the combinations of cells x_i and respondents r with nonzero contributions only, i.e. for the combinations that appear in Table 9. For the combination of cell x_1 and respondent R_1, constraint (6.8) becomes

$$\hat{\xi}_{11}^{+} \geq 155(\hat{\mu}_1 + \hat{\mu}_2)$$

and constraint (6.9) is given by

$$\hat{\xi}_{11}^{-} \geq -155(\hat{\mu}_1 + \hat{\mu}_2),$$

where $(\hat{\mu}_1 + \hat{\mu}_2)$ is the coefficient of x_1 in the aggregation and 155 is the contribution of respondent R_1 to x_1. Other constraints of types (6.8) and (6.9) are defined in a similar way.

The constraints in (6.10) on the contribution of the attacked respondent $\hat{A}_1$ are

$$\hat{A}_1 \leq \hat{\alpha}_{11} + M(1 - \hat{u}_1),$$
$$\hat{A}_1 \leq \hat{\alpha}_{12} + M(1 - \hat{u}_2),$$
$$\hat{A}_1 \leq \hat{\alpha}_{23} + \hat{\alpha}_{33} + M(1 - \hat{u}_3),$$
$$\hat{A}_1 \leq \hat{\alpha}_{24} + M(1 - \hat{u}_4),$$
$$\hat{A}_1 \leq \hat{\alpha}_{35} + M(1 - \hat{u}_5),$$
$$\hat{A}_1 \leq \hat{\alpha}_{46} + M(1 - \hat{u}_6).$$

The constraints in (6.11) resemble the constraints in (6.10). The only difference is that $\hat{A}_1$ is replaced by $\hat{A}_2$ and that $\hat{u}_r$ is replaced by $\hat{v}_r$. For instance, the first constraint in (6.11) is

$$\hat{A}_2 \le \hat{\alpha}_{11} + M(1-\hat{v}_1).$$

The upper bounds on $\hat{\alpha}_{ir}$ in (6.12) and (6.13) for cell x_1 and respondent R_1 become

$$\hat{\alpha}_{11} \le 155(\hat{\mu}_1 + \hat{\mu}_2) + M(1-\hat{\beta}_{11}),$$
$$\hat{\alpha}_{11} \le -155(\hat{\mu}_1 + \hat{\mu}_2) + M\hat{\beta}_{11}.$$

The other upper bounds on $\hat{\alpha}_{ir}$ are similar.

The constraints in (6.14), (6.15) and (6.16) are respectively

$$\hat{u}_1 + \hat{u}_2 + \hat{u}_3 + \hat{u}_4 + \hat{u}_5 + \hat{u}_6 = 1,$$
$$\hat{v}_1 + \hat{v}_2 + \hat{v}_3 + \hat{v}_4 + \hat{v}_5 + \hat{v}_6 = 1,$$
$$\hat{u}_r + \hat{v}_r \le 1 \qquad r = 1,\ldots,R.$$

The first and second constraints above mean that there is only one attacker and one unit under attack. The third one tells us that both units cannot be the same. Finally, the bounds in (6.17) – (6.20) become

$$\hat{\xi}^+_{11,},\ \hat{\xi}^+_{12},\ \hat{\xi}^+_{23},\ \hat{\xi}^+_{24},\ \hat{\xi}^+_{33},\ \hat{\xi}^+_{35},\ \hat{\xi}^+_{46} \ge 0,$$
$$\hat{\xi}^-_{11,},\ \hat{\xi}^-_{12},\ \hat{\xi}^-_{23},\ \hat{\xi}^-_{24},\ \hat{\xi}^-_{33},\ \hat{\xi}^-_{35},\ \hat{\xi}^-_{46} \ge 0,$$
$$\hat{u}_1,\ \hat{u}_2,\ \hat{u}_3,\ \hat{u}_4,\ \hat{u}_5,\ \hat{u}_6 \in \{0,1\},$$
$$\hat{v}_1,\ \hat{v}_2,\ \hat{v}_3,\ \hat{v}_4,\ \hat{v}_5,\ \hat{v}_6 \in \{0,1\},$$
$$\hat{\beta}_{11,},\ \hat{\beta}_{12},\ \hat{\beta}_{23},\ \hat{\beta}_{24},\ \hat{\beta}_{33},\ \hat{\beta}_{35},\ \hat{\beta}_{46} \in \{0,1\},$$
$$-1 \le \hat{\mu}_k \le 1 \qquad k = 1,\ldots,4.$$

The optimal solution to the above problem is given by: $\hat{\mu}_1 = -1$, $\hat{\mu}_2 = 1$, $\hat{\mu}_3 = 1$, $\hat{\mu}_4 = -1$, meaning that the most sensitive aggregation is

$$-2x_2 + 2x_3 = -680. \qquad (6.21)$$

Furthermore, we obtain the following absolute contributions to the aggregation: $\hat{\xi}^+_{33} = 56$, $\hat{\xi}^+_{35} = 24$, $\hat{\xi}^-_{23} = 400$, $\hat{\xi}^-_{24} = 360$. These figures are twice the figures of the contributions in Table 9, due to the coefficients of 2 and –2 in the aggregation. In the optimal solution $\hat{T}$ = 840, which is the sum over all absolute contributions (56 + 24 + 400 + 360).

The attacked unit is respondent 3 ($\hat{u}_3$=1) and respondent 4 is the attacker ($\hat{v}_4$=1). Their absolute contributions to the aggregation are given by $\hat{A}_1 = 456$ and $\hat{A}_2 = 360$. Note that respondent 3 contributes to two cells of the aggregation, x_2 and x_3, and therefore we have that $\hat{A}_1 = \hat{\xi}^+_{33} + \hat{\xi}^-_{23}$. Respondent 4 contributes to x_3 only, hence we obtain $\hat{A}_2 = \hat{\xi}^-_4$.

Since $120\hat{A}_1 + 100\hat{A}_2 - 100\hat{T} = 6720 \geq 0,$ we conclude that the table is insufficiently protected according to Operational criterion 2.

By looking at the aggregation (6.21), we can indeed see that the table is unsafe. Namely, respondent 4 can derive that the total contribution to the aggregation (6.21) of respondent 3 is at most –680 + 360 = –320. This upper bound approximates the true contribution to the aggregation $56 - 400 = -344$, too closely, i.e. within 20%. As a further illustration that aggregation (6.21) is sensitive, note that after multiplying (6.21) by –1/2 respondent 3 can derive that the contribution of respondent 4 to cell x_2 is at most $340 - 180 + 28 = 168$, which is within 20% of the true value. □

7 Discussion

In this article we have demonstrated that the commonly used definition to check whether a table with suppressed cell values is safe to be released or not is slightly inconsistent. We have presented a definition of a safe table that is correct in our opinion. In this article we have explored some of the consequences of our new definition, such as extending the sensitivity measure to general linear aggregations of cells.

Although we claim that our definition, i.e. Operational criterion 2, should replace the commonly used definition, i.e. Operational criterion 1, we can understand that some statistical agencies may not want to abandon using Operational criterion 1 completely, in any case not immediately. Such statistical agencies may consider using both criteria simultaneously. However, this would possibly lead to overprotecting a table. In this article we have not explored the combined use of different auditing criteria.

We have also given a method to check whether a table with suppressed cell values is safe to be released or not. As we already mentioned, such a checking method is usually referred to as a disclosure auditing method (see, e.g., [5] and [10]). A disclosure auditing method is, however, only a first step towards a method for constructing a safe table given an unsafe one.

Developing a cell suppression method that constructs a safe table, i.e. a table that satisfies Operational criterion 2, with (close to) minimal loss of information appears to be much more complicated than developing a disclosure auditing method. In theory it may be possible to base such a cell suppression method on the mixed-integer programming approach described in Section 6. For instance, an objective function could be minimized that depends on the number or the total of the suppressed cells (see also [30] and [31]) and one of the constraints would be that the resulting table is safe, i.e. that the value of the objective function (6.6) is negative subject to (6.7) to (6.20). Note that this constraint involves the result of a second optimization problem. In the literature this kind of model is known as a bi-level model. For more on bi-level models we refer to [1], [8], and [29].

Another theoretical possibility to base a cell suppression method on the disclosure auditing model described in Section 6 is an iterative procedure. During each iteration an objective function is minimized that, for instance, depends on the number or the total of suppressed values. Next, it is checked whether the resulting table is safe by means of the model described in Section 6. If not, one or more constraints, so-called cuts, are added to the problem, which basically say that the current solution is no longer allowed. For more on cuts and optimization algorithms based on cuts in general we refer to [18].

Unfortunately, both approaches sketched above are likely to be exceedingly slow in practice

Some alternative approaches not based on the disclosure auditing model of Section 6 have also been developed already. Sande [23], [24] and [25] has done work on an approach based on so-called elementary aggregations. The idea of this approach is that a table is safe when all elementary aggregations are safe, and hence that one only has to focus on protecting the elementary aggregations. Unfortunately, the number of elementary aggregations can be exceedingly high. At the U.S. Bureau of the Census an alternative approach has been followed to obtain safe suppression patterns [16]. The idea of this approach is to determine to what extent a cell can protect a sensitive cell before a suppression pattern is generated. The extent to which a cell can protect a certain sensitive cell is called the protection capacity of the former cell (for this sensitive cell). If the protection capacities are calculated correctly, one can ensure in this way that the final table with suppressed cell values is safe. However, determining the correct protection capacities can be very complicated for practical situations. One therefore often opts for a conservative approach where one may be too strict but is sure to produce a safe table.

Both the method based on elementary aggregations and the method based on protection capacities are not entirely satisfactory. Perhaps the approach by Salazar-González [22] may be used to construct safe tables, if appropriate protection levels can somehow be defined, but this would lead to very large and very complex optimization problems.

Currently, a cell suppression method that finds secondarily suppressed cells for tables arising in practice and that yields safe tables according to our definition in this article appears to be lacking. We encourage experts on operations research to develop better methods for constructing safe table by means of cell suppression. Developing such methods appears to be a major problem.

References

[1] Bard, J.F. (1998), *Practical Bilevel Optimization: Algorithms and Application*. Kluwer Academic Publishers, Dordrecht.

[2] Chvátal, V. (1983), *Linear Programming*. W.H. Freeman and Company, New York.

[3] Cox, L.H. (1981), Linear Sensitivity Measures in Statistical Disclosure Control. *Journal of Statistical Planning and Inference 5*, 153-164.

[4] Cox, L.H. (1995), Protecting Confidentiality in Business Surveys. In: *Business Survey Methods* (eds. B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge and P.S. Kott), John Wiley & Sons, Inc., New York, 443-473.

[5] Cox, L.H. (2001), Disclosure Risk for Tabular Economic Data. In: *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies* (eds. P. Doyle, J.I. Lane, J.J.M. Theeuwes and L.V. Zayatz), North-Holland Elsevier, Amsterdam, 167-183.

[6] Daalmans, J. (2002), *Deriving and Protecting Elementary Aggregations in the Cell Suppression Problem*. Report (research paper 0203), Statistics Netherlands, Voorburg.

[7] Dellaert, N.P. and W.A. Luijten (1999), Statistical Disclosure in General Three-Dimensional Tables. *Statistics Neerlandica 53*, 197-221.

[8] Dempe, S. (2002), *Foundations of Bilevel Programming*. Kluwer Academic Publishers, Dordrecht.

[9] Duarte De Carvalho, F., N.P. Dellaert and M. De Sanches Osório (1994), Statistical Disclosure in Two-Dimensional Tables: General Tables. *Journal of the American Statistical Association 89*, 1547-1557.

[10] Duncan, G.,T., S.E. Fienberg, R. Krishnan, R. Padman and S.R. Roehrig (2001), Disclosure Limitation Methods and Information Loss for Tabular Data. In: *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies* (eds. P. Doyle, J.I. Lane, J.J.M. Theeuwes and L.V. Zayatz), North-Holland Elsevier, Amsterdam, 135-166.

[11] Fischetti, M. and J.J. Salazar-González (2000), Models and Algorithms for Optimizing Cell Suppression in Tabular Data with Linear Constraints. *Journal of the American Statistical Association 95*, 916-928.

[12] Giessing, S. (2001), Nonperturbative Disclosure Control Methods for Tabular Data. In: *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies* (eds. P. Doyle, J.I. Lane, J.J.M. Theeuwes and L.V. Zayatz), North-Holland Elsevier, Amsterdam, 185-213.

[13] Giessing, S. (2004), Survey on Methods for Tabular Data Protection in ARGUS. In: *Privacy in Statistical Databases* (eds. J. Domingo-Ferrer and V. Torra), Springer-Verlag, Berlin, 1-13.

[14] Giessing, S. and S. Dittrich (2006), Harmonizing Table Protection: Results of a Study. In: *Privacy in Statistical Databases* (eds. J. Domingo-Ferrer and L. Franconi), Springer-Verlag, Berlin, 35-47.

[15] Hundepool, A (2006), The ARGUS Software in CENEX. In: *Privacy in Statistical Databases* (eds. J. Domingo-Ferrer and L. Franconi), Springer-Verlag, Berlin, 334-346.

[16] Jewett, R. (1993), *Disclosure Analysis for the 1992 Economic Census*. Unpublished Manuscript. Economic Programming Division, U.S. Bureau of the Census, Washington, DC.

[17] Kelly, J.P., B.L. Golden and A.A. Assad (1992), Cell Suppression: Disclosure Protection for Sensitive Tabular Data. *Networks 22*, 397-417.

[18] Nemhauser, G.L. and L.A. Wolsey (1988), *Integer and Combinatorial Optimisation.* Wiley, New York.

[19] Robertson, D. (1992), Cell Suppression at Statistics Canada. *Proceedings of the Annual Research Conference*, US Bureau of the Census, Washington DC, 107-135.

[20] Robertson, D. (1995), Automated Disclosure Control at Statistics Canada. *Proceedings of the Second International Seminar on Statistical Confidentiality*, Luxembourg.

[21] Robertson, D. (2000), Improving Statistics Canada's Cell Suppression Software (CONFID). In: *Proceedings in Computational Statistics 2000* (eds. J.G. Bethlehem and P.G.M. Van der Heijden), Physica-Verlag, New York, 403-408.

[22] Salazar-González, J.J. (2002), Extending Cell Suppression to Protect Tabular Data against Several Attackers. In: *Inference Control in Statistical Databases, From Theory to Practice* (editor J. Domingo-Ferrer), Springer, 34 – 58.

[23] Sande, G. (1977), *Towards Automated Disclosure Analysis for Establishment Based Statistics*. Report, Statistics Canada.

[24] Sande, G. (1978a), *A Theorem Concerning Elementary Aggregations*. Report, Statistics Canada.

[25] Sande, G. (1978b), *Confidentiality and Polyhedra – An Analysis of Suppressed Entries and Cross-Tabulations*. Report, Statistics Canada.

[26] Sande, G. (1984), Automated Cell Suppression Software to Preserve Confidentiality of Business Statistics. *Statistical Journal of the United Nations ECE 2*, 33-41.

[27] Sande, G. (1999), Structure of the ACS Automated Cell Suppression System. *Joint ECE/Eurostat Work Session on Statistical Data Confidentiality*.

[28] Sande, G. (2000), *Blunders by Official Statistical Agencies While Protecting the Confidentiality of Business Statistics* (unpublished paper).

[29] Shimizu, K., Y. Ishizuka, and J.K. Bard (1997), *Nondifferentiable and Two-Level Mathematical Programming*. Kluwer Academic Publishers, Dordrecht.

[30] Willenborg, L. en T. de Waal (1996), *Statistical Disclosure Control in Practice*. Springer-Verlag, New York.

[31] Willenborg, L. en T. de Waal (2001), *Elements of Statistical Disclosure Control*. Springer-Verlag, New York.

[32] Zayatz, L. (2007), Disclosure Avoidance Practices and Research at the U.S. Census Bureau: An Update. *Journal of Official Statistics 23*, 253-265.

Appendix

Proof of Theorem 1.

Consider an aggregation X_j that involves a cell x_{i_0}, i.e. $\lambda_{i_0}^j \neq 0$. Below it will be shown that all contributions to x_{i_0} are sufficiently protected according to the (*p*,*q*)-prior/posterior rule, if all contributions to X_j are sufficiently protected.

A contribution to x_{i_0} of a respondent *t* is sufficiently protected for an attacker *s* if this attacker cannot estimate an upper or lower bound for $x_{i_0}^t$ to within *p*% of its actual value. A mathematical expression for the upper bound will be given below. It will be derived from the bounds on $\lambda_{i_0}^j x_{i_0}^t$. From the perspective of attacker *s* these bounds are given by

$$U_s(\lambda_{i_0}^j x_{i_0}^t) = \lambda_{i_0}^j x_{i_0}^t + \frac{q}{100}\sum_{r \neq s,t} X_{jA}^r + \frac{q}{100}\sum_{i \neq i_0} \left|\lambda_i^j x_i^t\right|, \tag{A.1}$$

$$L_s(\lambda_{i_0}^j x_{i_0}^t) = \lambda_{i_0}^j x_{i_0}^t - \frac{q}{100}\sum_{r \neq s,t} X_{jA}^r - \frac{q}{100}\sum_{i \neq i_0} |\lambda_i^j x_i^t|. \tag{A.2}$$

Analogous to the derivation of (5.1), the attacker can derive these bounds from the value at the right-hand side of the aggregation and bounds on the contributions of all other respondents, except from the contribution under attack. By definition these estimates differ *q* percent from the actual values, which explains the expressions (A.1) and (A.2).

The second term at the right-hand side of (A.1) and (A.2) stands for all absolute contributions from respondents other than *s* and *t*, and the third term denotes all contributions of respondent *t*, except for the contribution to x_{i_0} that is under attack.

From (A.1) and (A.2) an upper bound for the contribution $x_{i_0}^t$ can be derived from the perspective of respondent *s*. This upper bound is

$$U_s(x_{i_0}^t) = x_{i_0}^t + \frac{q}{100\left|\lambda_{i_0}^j\right|}\sum_{r \neq s,t} X_{jA}^r + \frac{q}{100\left|\lambda_{i_0}^j\right|}\sum_{i \neq i_0} \left|\lambda_i^j x_i^t\right|. \tag{A.3}$$

The derivation depends on the sign of $\lambda_{i_0}^j$, i.e. $U_s(x_{i_0}^t) = U_s(\lambda_{i_0}^j x_{i_0}^t)/\lambda_{i_0}^j$ if $\lambda_{i_0}^j \geq 0$, else $U_s(x_{i_0}^t) = L_s(\lambda_{i_0}^j x_{i_0}^t)/\lambda_{i_0}^j$. In both cases we obtain (A.3).

The contribution $x_{i_0}^t$ is sufficiently protected for an attacker *s* if

$$U_s(x_{i_0}^t) > x_{i_0}^t + \frac{p}{100}|x_{i_0}^t|,$$

which implies that $x_{i_0}^t$ is sufficiently protected for an attacker s if

$$\frac{q}{\left|\lambda_{i_0}^j\right|}\sum_{r\neq s,t} X_{jA}^r + \frac{q}{\left|\lambda_{i_0}^j\right|}\sum_{i\neq i_0}\left|\lambda_i^j x_i^t\right| > p\,|x_{i_0}^t|. \tag{A.4}$$

Under the assumption that the contribution of contributor t to the aggregation X_j is sufficiently protected for contributor s, i.e.

$$q\sum_{r\neq s,t} X_{jA}^r > pX_{jA}^t,$$

see (5.3), it follows that

$$\frac{q}{\left|\lambda_{i_0}^j\right|}\sum_{r\neq s,t} X_{jA}^r + \frac{q}{\left|\lambda_{i_0}^j\right|}\sum_{i\neq i_0}\left|\lambda_i^j x_i^t\right| \geq \frac{q}{\left|\lambda_{i_0}^j\right|}\sum_{r\neq s,t} X_{jA}^r > \frac{p}{\left|\lambda_{i_0}^j\right|} X_{jA}^t = \frac{p}{\left|\lambda_{i_0}^j\right|}\sum_{i=1}^{N_S}\left|\lambda_i^j\right|\left|x_i^t\right| \geq$$

$$\frac{p}{\left|\lambda_{i_0}^j\right|}\left|\lambda_{i_0}^j\right|\left|x_{i_0}^t\right| = p\left|x_{i_0}^t\right|,$$

which shows that condition (A.4) is indeed satisfied, and hence that $x_{i_0}^t$ is sufficiently protected for attacker s.

Since t denotes an arbitrary contributor, s an arbitrary attacker, x_{i_0} denotes an arbitrary cell and X_j an arbitrary aggregation, we obtain the result that all contributions to cells involved in a safe aggregation are sufficiently protected. □

Proof of Theorem 2.

Consider an aggregation X_j and suppose $S_{p,q}^a(x_i) < 0$, for every i with $\lambda_i^j \neq 0$, which means that this aggregation only involves non-sensitive cells. We will prove that $S_{p,q}^a(X_j) < 0$, where

$$S_{p,q}^a(X_j) = pX_{jA}^{[1]} - q\sum_{r=3}^{R} X_{jA}^{[r]}, \tag{A.5}$$

That is, we will prove that the aggregation X_j is non-sensitive.

Theoretically, the largest possible absolute contribution to an aggregation is obtained from a holding that makes the largest absolute contribution to each of the underlying cells of the aggregation. This implies

$$X_{jA}^{[1]} \leq \sum_{i=1}^{N_S}\left|\lambda_i^j x_i^{[1]}\right|. \tag{A.6}$$

Analogously, the largest possible sum of the two largest contributions to X_j could come from two holdings that would make the largest and the second largest contribution to each of the underlying cells, i.e.

$$X_{jA}^{[1]} + X_{jA}^{[2]} \leq \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[1]}\right| + \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[2]}\right|.$$

Since

$$X_{jA} = \sum_{r=1}^{R} X_{jA}^{[r]} = \sum_{r=1}^{R} \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[r]}\right|,$$

it follows that

$$\sum_{r=3}^{R} X_{jA}^{[r]} \geq \sum_{r=3}^{R} \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[r]}\right|, \tag{A.7}$$

Combining (A.5) – (A.7) gives

$$S_{p,q}^{a}(X_j) \leq p \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[1]}\right| - q \sum_{r=3}^{R} \sum_{i=1}^{N_S} \left|\lambda_i^j x_i^{[r]}\right| = \sum_{i=1}^{N_S} \left|\lambda_i^j\right| \left(p\left|x_i^{[1]}\right| - \sum_{r=3}^{R} q\left|x_i^{[r]}\right| \right) =$$

$$\sum_{i=1}^{N_S} \left|\lambda_i^j\right| S_{p,q}^{a}(x_i) < 0,$$

which shows that $S_{p,q}^{a}(X_i) < 0$, under the assumption that $S_{p,q}^{a}(x_i) < 0$, for every i with $\lambda_i^j \neq 0$. □