Counter Examples to the New EIA/USBC Confidentiality Auditing Method

Gordon Sande, Sande & Associates, Inc.
600 Sanderling Court, Secaucus, N.J. 07094

Abstract: In the conference paper "*Statistical Methods to Limit Disclosure for the Manufacturing Energy Consumption Survey: Collaborative Work of the Energy Information Administration and the U. S. Census Bureau*" by Ramesh Dandekar of the Energy Information Administration there is a new method proposed for confidentiality auditing. The proposed new method was intended to reduce over suppression of publications. It has several problems. It may both report residual disclosures which are not present and fail to find residual disclosures which are present. Simple examples of both failure modes can be constructed. Results of a confidentiality audit of the publication to which the proposed new method was applied are reported.

Introduction.

In the paper "*Statistical Methods to Limit Disclosure for the Manufacturing Energy Consumption Survey: Collaborative Work of the Energy Information Administration and the U. S. Census Bureau*" by Ramesh Dandekar of the Energy Information Administration there is a new method for confidentiality auditing. The new method has several problems.

Confidentiality auditing is used to provide an independent evaluation of a cell suppression activity. The input data to the process is the published table and the representation of the external user knowledge. A lax estimate of external user knowledge is the assertion that the withheld cell values are positive, but otherwise unknown. This would be the situation for an evaluation which had no subject matter knowledge as would be done outside of a statistical agency and away from relevant subject matter experts. A tight estimate of external user knowledge would be the assertion that the user could estimate the value to be between 0% and 200%, 50% and 150% or some other specified percentages of the true value. The most important feature of the tight estimate is that it provides a small upper bound on small cells and can be established with relatively limited subject matter knowledge. Such an extensive and precise estimate of user knowledge would be possible within a statistical agency where the true value is available and would not require access to subject matter experts. With this data, self consistent completions of the table are constructed by filling in possible values for the withheld cells. The smallest and largest possible values are noted and reported as lower and upper bounds on the external user estimates. This is readily implemented as a linear programming problem where we seek the minimum and maximum of simple objective functions. For a large collection of related tables, there can be many withheld cells and the total time to calculate all the objective functions can mount up. Both the lax and tight confidentiality audits can detect exact residual disclosures. Approximate residual disclosures of sensitive cells can be detected if the sensitivities of the cells are also available as they would be for a tight confidentiality audit within a statistical agency. The results of a tight confidentiality audit are not usually published as the center of many of the ranges is the true value. The results of a lax confidentiality audit are implicit in the publication.

Confidentiality auditing was carried for the 1994 Manufacturing Energy Consumption Survey (MECS). The main interest in this survey is in the Tables A1, A3, A4 and A5 which provide four definitions of manufacturing energy consumption. These tables are four differing combinations of six components of energy consumption. The symbolic relationships are

$$\begin{aligned} A1 &= c1 + c2 + c3 + c4 \\ A3 &= c2 + c4 \\ A4 &= c1 + c3 + c5 \\ A5 &= c1 \end{aligned}$$

where c6 is out of scope for MECS and defined to be zero. These symbolic equations can be solved for the components to yield

$$\begin{aligned} c1 &= A5 \\ c3 &= A1 - A3 - A5 \\ c5 &= - A1 + A3 + A4 \\ c2 + c4 &= A3. \end{aligned}$$

The main MECS tables are three dimensional with classification by fuel type, geography and SIC code. The components can be arranged in a two by three table of production by consumption type. The published tables are selected three dimensional sub tables from a five dimensional collection of tables. Most of the five way table is not presented in any publication, and we may say it is *hidden* from publication. Some of the hidden values will be well determined, while others will not be as a result of the withholding in the published tables.

The USBC cell suppression system is based on network theory. This theory only applies to two dimensional tables, with some restrictions. The extension to three dimensions is by treating the third dimension as a collection of independent one dimensional problems. The result is a system that

> *I knew that this program would not do a theoretically correct job of disclosures analysis, but I hoped it would at least give the respondents some protection and would run within a reasonable amount of time.* (Jewett 1993)

which requires a confidentiality audit procedure to detect defective patterns. The structure of the main MECS tables is more complex than three dimensions with the relations between the tables represented in the USBC cell suppression system by inequalities. In tests run with the 1991 micro data (Kirkendall 1996), there were exact residual disclosures when the main MECS tables were jointly analyzed using a linear programming confidentiality audit.

The new method is a replacement for the linear programming confidentiality audit that is intended to lower the execution cost.

New confidentiality audit method.

The Dandekar paper provides a full description of the method applied to a small illustrative example. It also provides an incomplete description of the method applied to MECS, as it is too bulky for the paper's format. We will follow Dandekar and examine the small example. The data has two components which lead to a publication of two tables. The first table, called A, includes both components and the second table, called B, includes only the first component. This is the relationship of Tables A1 and A3 above. In the example, there is no table for the second component but we can define a table, called C, which is hidden and includes the data for the second component. We would have the symbolic relationships

$$A = c1 + c2$$
$$B = c1$$

which can be solved for

$$c1 = B$$
$$c2 = A - B$$

which we might also express by

$$A = c1 + c2$$
$$B = c1$$
$$C = c2.$$

We seek to confidentiality audit the released tables A and B without setting up the explicit structure of A = B + C and components c1 and c2.

The new method uses four sets of equations. Each equation in each set is evaluated independently. Two sets of equations are existing equations and two sets are constructed equations. The construction can only be done, like the tight confidentiality audit, with the availability of the original tables as would be true within the originating agency. The first set of equations are the tables A and B. We would verify that

$$A0 = A1 + ... + An \quad (1.1)$$
$$B0 = B1 + ... + Bn \quad (1.2)$$

do not have any disclosures. This also serves to establish the notation that the table A has total column A0 and internal columns A1 through An. Similarly for table B. The second set of equations are the relations

$$B0 \leq A0 \quad (2.0)$$
$$B1 \leq A1 \quad (2.1)$$
$$...$$
$$Bn \leq An \quad (2.n)$$

The first set of constructed equations are

$$B0 = A1 + ... + An - [c0] \quad (3.0)$$
$$A0 = B1 + ... + An + [c1] \quad (3.1)$$
$$...$$
$$A0 = A1 + ... + Bn + [cn] \quad (3.n)$$

in which we substitute a column of B for the corresponding column of A. The resulting equations will not balance and a correction term must be supplied. (Following Dandekar, there is a problem in notation with components, table C and correction terms which is reduced somewhat by the use of the redundant brackets on the correction terms.) The correction term is the difference between the corresponding columns of the true tables. This information is available within an agency. The second set of constructed equations is

$$A0 = B1 + ... + Bn + [c0] \quad (4.0)$$
$$B0 = A1 + ... + Bn - [c1] \quad (4.1)$$
$$...$$
$$B0 = B1 + ... + An - [cn] \quad (4.n)$$

in which we substitute a column of A for the corresponding column of B. The correction term for these equations is the negative of the correction term for the first set of constructed equations. We would verify that each set of equations has no disclosures and also that the effect of the lower and upper bounds across the equations do not provide for a disclosure.

The simplest example is the 2 by 2 by 2 table. The true values would be

Table A (true)

8	4	4
4	2	2
4	2	2

Table B (true)

4	2	2
2	1	1
2	1	1

Table C (true)

4	2	2
2	1	1
2	1	1

and the released tables would be

Table A

8	4	4
4	2	2
4	2	2

Table B

4	2	2
2	x	x
2	x	x

with Table C defined but not released.

The equations (1.1), (1.2), (2.0), (2.1) and (2.2) are the existing equations and readily verified. For equation (3.0) we may arrange the equations into a table as

Equation (3.0)

4	4	4	-4
2	2	2	-2
2	2	2	-2

where the last column has the correction terms. Equations (3.1) and (3.2) are

Equation (3.1)

8	2	4	2
4	x	2	1
4	x	2	1

Equation (3.2)

8	4	2	2
4	2	x	1
4	2	x	1

Similarly equations (4.0), (4.1) and (4.2) are

Equation (4.0)

8	2	2	4
4	x	x	2
4	x	x	2

Equation (4.1)

4	4	2	-2
2	2	x	-1
2	2	x	-1

Equation (4.2)

4	2	4	-2
2	x	2	-1
2	x	2	-1

We see that equations (3.1), (3.2), (4.1) and (4.2) all indicate disclosures. These are all false indications of the presence of disclosures.

The complements which protect a sensitive cell form a path. (In two dimensions these are often called circuits but there is no common terminology for higher dimensions so we choose a neutral word to allow for the more general case.) If any of the cells on the path is given a value, we will be able to determine values for all the other cells on the path. The replacement of a column in a table by a new column is very prone to cause this cascade of known values. The above example of multiple false indications of disclosures is an example of this.

We may also use the example

Table A (true)

13	5	5	3
5	2	2	1
5	2	2	1
3	1	1	1

Table B (true)

9	3	3	3
3	1	1	1
3	1	1	1
3	1	1	1

Table C (true)

4	2	2	0
2	1	1	0
2	1	1	0
0	0	0	0

with a defective release pattern

Table A

13	5	5	3
5	2	x	x
5	x	2	x
3	x	x	x

Table B

9	3	3	3
3	1	x	x
3	x	y	x
3	x	x	x

We may show the problem by calculating the margins and one internal entry of C = A - B as

Table C (step 1)

4	2	2	0
2	1	-	-
2	-	-	-
0	-	-	-

which is completed from the zero margins and filling in the three remaining cells as

Table C (step 2)

4	2	2	0
2	1	1	0
2	1	1	0
0	0	0	0

which in turn yields that y is disclosed to be 1.

Equations (1.1), (1.2), (2.0), (2.1), (2.2) and (2.3) are all readily verified. We have

Equation (3.0)

9	5	5	3	-4
3	2	x	x	-2
3	x	2	x	-2
3	x	x	x	0

Equation (3.1)

13	3	5	3	2
5	1	x	x	1
5	x	2	x	1
3	x	x	x	0

Equation (3.2)

13	5	3	3	2
5	2	x	x	1
5	x	y	x	1
3	x	x	x	0

Equation (3.3)

13	5	5	3	0
5	2	x	x	0
5	x	2	x	0
3	x	x	x	0

Equation (4.0)

13	3	3	3	4
5	1	x	x	2
5	x	y	x	2
3	x	x	x	0

Equation (4.1)

9	5	3	3	-2
3	2	x	x	-1
3	x	y	x	-1
3	x	x	x	0

Equation (4.2)

9	3	5	3	-2
3	1	x	x	-1
3	x	2	x	-1
3	x	x	x	0

Equation (4.3)

9	3	3	3	0
3	1	x	x	0
3	x	y	x	0
3	x	x	x	0

which are the same configurations as (1.1) and (1.2) and show no disclosures. This is a false indication of the absence of disclosures.

We have examples of using the new method where disclosures are falsely reported and where disclosures are not reported when present. The claim in the Dandekar paper that the new method is equivalent to a full linear programming confidentiality audit is not true.

The well known example of a defective suppression pattern in a two way table is

Defective pattern

16	4	4	4	4
4	1	1	x	x
4	1	1	x	x
4	x	x	1	1
4	x	x	1	y

displayed in more symmetrical form than is usual. The defect in this pattern can be shown by noticing that the 2 by 2 block of 4 values of 1 has the effect of determining the total of the other internal 2 by 2 blocks. In particular, this is true for the diagonally opposite block of 3 values of 1 and a y which has a total of 4. Thus the value of y is 1. This construction can be done in three dimensions in which a 2 by 2 by 2 cube of 8 values of 1 will determine the total of the other internal 2 by 2 by 2 cubes. A diagonally opposite cube of 7 values of 1 and a y will have a known total and thus a value of 1 for the y, even when the six other cubes contain all xs. This can be repeated for any number of dimensions. The interest in the displayed example is that all the rows and columns have at least 2 withheld cells. We can also say that there are no disclosures in the one dimensional equations. In the three dimensional example there are no disclosures in any of the two dimensional equations. The example given by Zayatz (1992) is equivalent to this example. The symmetrical form makes the construction easier and also makes the analysis easier as any result for rows will also apply to columns or slices. In general for any number of dimensions, we have an example in which there are no disclosures in any of the equations with one less dimension. This provides a constructive example of the need to use the equations for the full number of dimensions to detect disclosures.

The Dandekar paper gives no motivation for or derivation of the new method beyond the specification of the computational procedure. A standard problem in operations research is the reduction in the number of variables involved in s system of constraints. A possible objective of the new method is to reduce the number of variables in the constraint systems used for the confidentiality audit. The Fourier-Motzkin elimination method is often proposed for this purpose. It, or any other method, is rarely used as the number of constraints grows quickly and the Fourier-Motzkin method often introduces redundant constraints beyond those that are required. The growth in the number of constraints usually increases the costs faster than the reduction in costs from decreasing the variable count. Exceptions are special cases such as systems with only two variables per equation, some theoretical demonstrations or vertex generation methods, none of which apply here. The column substitution scheme would seem prone to disrupt the structure of the paths of complements and the simple nature of the counter examples is not surprising. The need for the full dimension structure of the equations is well known in lower dimensions and readily demonstrated for all dimensions as above. These examples to not immediately apply here but counter examples that do apply can be constructed as minor modifications as seen above. The two modes of operation of the new method to either show the presence or absence of disclosures have counter examples which follow examples of standard facts about cell suppression.

Sensitive cell only speed up.

The Dandekar paper also reports that it is only necessary to confidentiality audit the sensitive cells. A cell suppression program may use several paths of complements to protect a sensitive cell. It is possible that only one of these paths is defective. The result will be some exact residual disclosures of of complements and an approximate residual disclosure of the sensitive cell which has less protection than was intended. We may use the example

Table A (true)

17	5	6	6
5	1	2	2
6	2	2	2
6	2	2	2

Table B (true)

9	3	3	3
3	1	1	1
3	1	1	1
3	1	1	1

Table C (true)

8	2	3	3
2	0	1	1
3	1	1	1
3	1	1	1

where the zero in Table C is assumed known. We release

Table A (released)

17	5	6	6
5	1	x	x
6	x	x	x
6	x	x	x

Table B (released)

9	3	3	3
3	z	y	1
3	y	X	x
3	1	x	x

in which the *X* is in two paths in this table, one with the *x*s and the other with the two *y*s and *z*. We get

Table C (calculated)

8	2	3	3
2	0	x	x
3	x	x	x
3	x	x	x

which allows the value of *z* to be determined which yields the values of the two *y*s. If the *X* is the only cell being evaluated because it is known to be sensitive, then we will miss the residual disclosures of the *y*s and *z*. If there is a tight confidentiality audit the limits on *X* will be tighter than intended and we will show an approximate residual disclosure. A lax confidentiality audit may not show the approximate residual disclosure.

Any cell suppression procedure will tend to use multiple paths if either it is intended to reduce the total value of suppressed cells or it has to deal with highly concentrated large cells. Failures will be associated with paths so that multiple exact residual disclosures of non sensitive complements may be associated with an approximate residual disclosure of a sensitive cell. The presence of exact residual disclosures of non sensitive cells in a publication is over suppression. Such over suppression lowers the value of the publication to the end users. It also serves to indicate that there are technical problems in the production of the publication in the area of the confidentiality protection provided to the respondents. One of the purposes of a confidentiality audit is to both detect such failures and to display the information that an end user could also display. The restriction to evaluation of sensitive cells as a computational economy loses this important capability of the confidentiality audit.

Reduction of over suppression.

The Dandekar paper also reports that confidentiality auditing was used to reduce over suppression in the MECSs publication. A cell suppression method may have over suppression because either it uses a configuration of cells which has more suppression than an alternate (possibly quite different) configuration of cells or it uses a configuration of cells that included more cells than were necessary. The alternate configuration mode of over suppression is associated with use of multiple single cell protection steps to achieve multiple cell protection. The excess cells mode of over suppression is associated with the use of continuous optimization methods rather than discrete optimization methods. In one dimension this latter problem is the well known knapsack problem of discrete optimization in which there is an attempt to find the subset of a set of objects whose sum is closest to a specified value. Confidentiality auditing of final suppression patterns will neither identify nor ameliorate either version of the over suppression problem. These are technical definitions of over suppression but the term many also have a common meaning of excessive suppression.

The USBC cell suppression system partitions the full problem into multiple smaller network problems. The small problems influence each other through a process called backtracking. The intermediate results can be confidentiality audited and the backtracking can be terminated if all sensitive cells are seen to be protected. When backtracking operates it is prone to have more alternate configuration mode over suppression than would be present in other cell suppression methods and confidentiality auditing will not reduce such over suppression. The Dandekar paper identifies the reduction in over suppression with the reduction in the amount of suppression resulting from backtracking. This suggests that over suppression may be being used to mean excessive suppression which is a term also used in the Dandekar paper.

Each single cell protection step acts to determine the additional complementary cell suppression required to protect the current sensitive cell. The determination may be that the sensitive cell is currently protected and no further complementary suppression is required. A confidentiality audit of the intermediate results, under the same confidentiality model as used by a cell suppression system with backtracking, would produce the same result and the only effect would be to stop the backtracking before a stage which would produce no new complementary suppression. The partitioning of the cell suppression problem causes the USBC cell suppression system to discover more protection than is actually present. Either the upper bounds will be higher, the lower bounds will be lower or possibly both because it does not use all of the equations possible but only those which represent the network in the current partition. We would expect to see situations in which the confidentiality audit of intermediate, or final, results indicates the need for more protection but the system determines that it has finished. This possibility is the reason why a confidentiality audit is required to detect defective patterns when the USBC cell suppression system is used for three, or more, dimensions.

For the confidentiality auditing to stop the backtracking before the USBC cell suppression system there must be a difference in the confidentiality model. The easiest difference to introduce is the lax confidentiality audit versus the tight confidentiality audit which is a difference in the model of the external user knowledge. The Dandekar paper does not address this issue. This form of difference can readily result from minor implementation differences. Another easy to introduce difference is the use of a cell by cell confidentiality model rather than a model which accounts for multiple cells. The simplest example of this difference is a single respondent cell with a complement of another single respondent cell of approximately the same size. Under a cell by cell model this would be permitted. Under a multiple cell model it would be recognized that the sum of these two cells, with an exact value as they would be the only cells withheld in a row or column, would be sensitive as it has only two respondents and would not be permitted. The USBC cell suppression system uses a multiple cell confidentiality model which identifies pairs of cells and would not permit this example to be published. The Dandekar paper does not address this issue. The multiple cell confidentiality model of the USBC cell suppression system would require that numerical coefficients, called protection capacities, that implement the model of user knowledge be adjusted for each sensitive cell. Omitting this repeated adjustment of the model of user knowledge would simplify the implementation of the confidentiality audit. Other forms of difference in the confidentiality model are also possible, but they are not as easy to introduce and they would require a much more explicit change in the confidentiality model. A possible example would be for the USBC cell suppression system to use an *n-k%* type sensitivity rule and for the confidentiality audit to use a *p%* type sensitivity rule, or the reversed situation. Another form of difference would be for the confidentiality auditing software to be defective and to incorrectly report that no further suppression is required. The results above indicate that the latter is quite possible.

The Dandekar paper indicates that confidentiality auditing of intermediate results was used to reduce the amount of suppression that was required from the backtracking process. A flexibility feature of the USBC cell suppression system is the use of *fudge factors*. This permits the system to lower its need for additional complements by lowering their required size by applying numerical coefficients. This flexibility was used in the MECS processing to match a weaker confidentiality model in the confidentiality audit software (Jewett 1998). At the 1996 Annual Joint Statistical Meetings a USBC speaker indicated that the use of such *fudge factors* was being discouraged. They are now no longer in use for surveys processed in 1998 or later (Clark 1998). A natural method of reducing suppression during backtracking would be to apply suitable *fudge factors* so that the confidentiality audit of intermediate results and the USBC cell suppression system agreed on the need for no further backtracking.

There are two broad solutions to the problem of a publication having excessive suppression. One is to lower the amount of confidentiality protection provided to the respondents so that the amount of suppression is no longer excessive. This is not acceptable if the original determination of the level of confidentiality protection is correct. This solution then becomes one of either revising the specification of the level of confidentiality protection provided or revising the assessment of excessive suppression. The other solution is to increase the quality of the protection procedures and reduce over suppression. The removal of superfluous suppressions is one quality improvement method. The use of alternate suppression algorithms that have a wider choice of complements or that make the choice of complements in an improved way is another form of quality improvement. The mode of protection may be revised so that range publication is used as a quality improvement over cell suppression. The USBC cell suppression system has various forms of over suppression. Sullivan (1993) provides examples of over suppression in the backtracking. The example given applies to the previous version of the system which did not have the ability to allow one classification variable to have a hierarchical code but the principles still apply. Kirkendall (1996) reports that superfluous suppressions were removed in some of the trials with the 1991 MECS micro data. Neither of these were detected or ameliorated by the use of a confidentiality audit.

The solution to over suppression used in the Dandekar paper of reducing the amount of backtracking addresses excessive suppression by changing the specification of the confidentiality protection provided to the respondents. This has been done by use of an unspecified, but apparently weaker, confidentiality model which was then approximately matched by appropriate *fudge factors* in the USBC cell suppression system.

Conclusions.

The Dandekar paper reports that the new method for confidentiality auditing was applied to the 1994 MECS. Reduction in over suppression is reported as a successful result of use of the new method. The new method is claimed to be equivalent to a linear programming confidentiality audit of the full structure. From the examples we see that this claim is false. The restriction of the confidentiality auditing to only the sensitive cells means that a partial failure may leave non sensitive complementary suppressed cells with exact residual disclosures unreported. The sensitive cells will usually be subject to approximate residual disclosure in this case. The claimed reduction of over suppression may be stopping because of the use of a different confidentiality model than the one used by the USBC cell suppression system or stopping because of the use of a defective confidentiality audit which did not recognize disclosures. Both cases would have the effect of lowering the protection provided to the respondents.

Applying a confidentiality audit to the published tables can address some of these ambiguities. The confidentiality audit can only be a lax confidentiality audit under a cell by cell confidentiality model as the true tables are not published. The published tables have heavy

1994 MECS - Tables A1, A3, A4 and A5
- SIC Major Groups by Fuel Type by Census Region (945 cells only)

Cell statuses under separate and joint confidentiality audits.

	A1	A3	A4	A5	Joint	
Published as value	438	111	524	523	1596	
Published as zero or *	160	578	159	162	1059	
Withheld as q or h	109	6	0	5	120	
Withheld as w	238	250	262	255	1005	
W not disclosed	195	81	215	203	48	= 694 - 646
W disclosed	43	169	47	52	957	= 311 + 646
W disclosed = value	4	13	4	6	287	= 27 + 260
W disclosed = 1	4	18	4	5	202	= 31 + 171
W disclosed = 0	35	138	39	41	468	= 253 + 215
						= sum + extra

independent rounding of their entries. The publication unit of measure is trillion (10^{15}) BTU for micro data collected in million (10^{6}) BTU. The MECS tables have a large number of *w* symbols to indicate withholding. This indicates that there has been an attempt made to provide the end users with as much information as can be released from this data collection. If there are not a large number of withholding symbols, then the data could be more finely classified to provide more detail to the end users. A confidentiality audit must deal with the presence of the *rounded to zero* symbols of * and the independent rounding of the tables. The quickest solution is to replace the *rounded to zero* symbols by *zero* and to force the tables to be exactly additive by a adjustment with a minimum sum of absolute adjustments procedure. This has the disadvantage of collapsing some small protection intervals. These are a result of not coordinating the cell suppression activity with the rounding for publication activity. There are multiple *w*s that are trivially *rounded to zero* symbols and thus zero for the analysis. Some of the withheld values may be negative as physical activities such as *Net Electricity* may be negative as a result of cogeneration. When the MECS Tables A1, A3, A4 and A5 are subjected to stand alone confidentiality audits they show large numbers of exact residual disclosures. Many of the disclosed values are small, which is consistent with the collapsing of small protection intervals and the tendency of small cells to be sensitive. Table A1 with 5400 cells has 646 exact residual disclosures in 1518 withheld cells (i.e. 43%), Table A3 with 1992 cells has 329 exact residual disclosures in 495 withheld cells (66%), Table A4 with 4590 cells has 768 exact residual disclosures in 1450 withheld cells (53%) and Table A5 with 4590 cells has 898 exact residual disclosures in 1407 withheld cells (64%).

It is also of interest to do a joint analysis of the tables as this was the motivation for the development of the new auditing method. The highest level SIC aggregations are the 20 major groups. We may collect these for the Census Regions and define the components c3 and c5 as well as the total of all the components. The resulting higher way table can be subjected to a full linear programming audit. The results are displayed in a table of separate and joint analyses. To allow the joint analysis we must include a column of zeros in A3 for Net Electricity and treat the 105 Totals in A1 as if they were hidden due to their slightly different treatment which breaks out sales of energy in that table. The resulting table does not use the full SIC and geographical structure so we would expect to see fewer residual disclosures than in tables which did use the full structure. The use of the full structure has a tendency to identify more residual disclosures in high level tables such as this one. For this portion of the publication under the separate analysis there 311 exact residual disclosures in 1005 withheld cells (31%). Under the joint analysis there is an increase by 646 to 957 exact residual disclosures in 1005 withheld cells (95%). One would expect that there should be fewer disclosures under the separate analysis as is observed here. The number of disclosures under the joint analysis seems to be considerably elevated from what one would expect. However given the various problems which are present in Dandekars's new auditing method, this may not be as surprising as it might be.

References.

Clark, C. (1998). private letter, 2 March 1998.

Dandekar, Ramesh (1998) "Statistical Methods to Limit Disclosure for the Manufacturing Energy Consumption Survey: Collaborative Work of the Energy Information Administration and the U. S. Census Bureau", In J. Domingo-Ferrer, ed., *Statistical Data Protection '98,* Conference Proceedings 25-27 March 1998, Lisbon, Portugal.

Jewett, R. (1993). "Disclosure Analysis for the 1992 Economic Census", Economic Programming Division, U. S. Bureau of the Census, Washington, DC.

Jewett, R. (1998). private e-mails, February 1998.

Kirkendall, N. et al. (1996). "Report on EIA-Census Evaluation of Disclosure Limitation Methods", Office of Statistical Standards, Energy Information Administration,Washington, DC.

Sullivan, C. (1993). "A Comparison of Cell Suppression Methods" ESMD-9301, Economic Statistics Methods Division, U. S. Bureau of the Census, Washington, DC.

Zayatz, L. (1992). "Linear Programming Methodology for Disclosure Avoidance Purposes at the Census Bureau," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 679-684.