**WP.6**
ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE**
**CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION**
**STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)**

**Joint UNECE/Eurostat work session on statistical data confidentiality**
(Manchester, United Kingdom, 17-19 December 2007)

Topic (i): Microdata

# EVALUATING THE DISCLOSURE RISKS OF REPORTING QUALITY MEASURES TO THE PUBLIC

**Invited Paper**

Prepared by Jerome P. Reiter, Duke University,
Anna Oganian and Alan F. Karr, National Institute of Statistical Sciences, United States of America

# Evaluating the disclosure risks of reporting quality measures to the public

Jerome P. Reiter*, Anna Oganian**, Alan F. Karr***

* Department of Statistical Science, Duke University, Durham, NC, USA, (jerry@stat.duke.edu)

** National Institute of Statistical Sciences, Research Triangle Park, NC, USA, (aoganian@niss.org)

*** National Institute of Statistical Sciences, Research Triangle Park, NC, USA, (karr@niss.org)

**Abstract**. To protect confidentiality, statistical agencies typically alter data before releasing them to the public. Ideally, although rarely done, the agency releasing data also provides a way for secondary data analysts to assess the quality of inferences obtained with the released data. Quality measures can help secondary data analysts to disregard inaccurate conclusions resulting from the disclosure limitation procedures, as well as have confidence in accurate conclusions. We propose an interactive computer system that analysts can query for measures of data quality. We focus on potential disclosure risks of providing these quality measures.

## 1 Introduction

Many national statistical agencies, survey organizations, and researchers disseminate microdata, i.e. data on individual units, to the public. Wide dissemination of data greatly benefits society, enabling broad subsets of the research community to access and analyze the collected data. Often, however, data disseminators cannot release microdata as collected, because doing so could reveal survey respondents' identities or values of sensitive attributes.

Data disseminators therefore limit what they release, typically by altering the collected data. Common strategies include recoding variables, such as releasing ages or geographical variables in aggregated categories; reporting exact values only above or below certain thresholds, for example reporting all incomes above $100,000 as "$100,000 or more"; swapping data values for selected records, e.g., switch the quasi-identifiers for at-risk records with those for other records to discourage users from matching, since matches may be based on incorrect data; and, adding noise to numerical data values to reduce the possibilities of exact matching on key variables

or to distort the values of sensitive variables (Willenborg and de Waal, 2001). Another approach is to replace sensitive values with multiple imputations, often called synthetic data (Little, 1993; Rubin, 1993; Reiter, 2003, 2004; Abowd and Woodcock, 2004).

Generally, increasing the amount of alteration decreases the risks of disclosures; but, it also decreases the accuracy of inferences obtained from the released data. Unfortunately, with most disclosure limitation strategies it is difficult for data users to determine how much their particular estimation has been compromised by the data alteration. This is especially true when data disseminators do not release detailed information about the disclosure limitation strategy.

Secondary data analysts would be greatly helped if statistical agencies provided some way for them to learn about the quality of inferences based on the released data. Benefits of providing data quality measures include (i) analysts can avoid publishing (in the broad sense) results with poor quality, (ii) analysts can feel more confident about results with good quality, and (iii) agencies can claim that they provided information about quality, so that they are not responsible if analysts arrive at inaccurate conclusions. Ideally, the measures are specific to particular inferential quantities rather than broad measures. For example, reporting comparisons of means, variances, and correlations in the observed and masked data does little to help analysts estimating complex models.

In this article, we discuss an approach that enables users of secondary data to assess the quality of inferences made on disclosure-proofed public use data. The idea is to create a *verification server*. This server, located at the statistical agency, would store the original and masked datasets. Analysts, who have only the masked data, would submit queries to the server for measures of data quality for certain estimands. The server would run the analysis on both the original and masked data, and report back to the analyst a measure of data quality that compares the inferences obtained from both sources. The server also serves as a data collection mechanism for the agency, capturing what quantities analysts care most about. Agencies might be able to utilize this information to improve the quality of future data releases.

Verification servers are not the proverbial free lunch. As we shall illustrate, providing measures of data quality also provides ill-intentioned users (henceforth called intruders) with more information for disclosure attacks. The usefulness of this additional information for intruders can be reduced by applying disclosure limitation strategies on the quality measures released by the server.

The remainder of this article is organized as follows. In Section 2, we give a formal representation of the verification server. In Section 3, we describe a data quality measure for the server. In Section 4, we illustrate potential disclosure attacks using the masked data and quality measures. In Section 5, we suggest some approaches for reducing the additional risks from releasing quality measures.

## 2 Formal description of verification server

The agency wants to release some version $M$ of original microdata, $D$, to the public in a way the protects confidentiality of survey respondents' identities and sensitive attributes. As is typical of disclosure settings, we assume that the agency does not reveal precise details about the disclosure limitation strategy, except in broad terms like, "noise was added to some variables, while other variables were swapped."

The analyst seeks inferences about some quantity $Q$, such as a confidence interval for a regression coefficient. Let $Q(M)$ be the estimate of $Q$ obtained from using $M$, and let $Q(D)$ be the estimate of $Q$ obtained from using $D$. The analyst can compute only $Q(M)$. In general, whenever $Q$ is based on units whose values were altered for disclosure protection, we expect that $Q(M) \neq Q(D)$. Of course, this is not always the case: $Q(M)$ could equal $Q(D)$ for some $Q$.

The agency wants to enable secondary data analysts to learn about the differences between $Q(M)$ and $Q(D)$. Given instructions from the analyst, the agency could manually compute $Q(M)$ and $Q(D)$ and describe the differences to the analyst, but this is a labor-intensive process. A preferable approach is to let analysts query a verification server for measures of the quality of their particular $Q(M)$. Let $FM(a, b)$ represent a numerical summary comparing $a$ to $b$. We call $FM$ a *fidelity measure*, as it represents the degree to which the quantity $b$ is faithful to the quantity $a$. For queries for acceptable $Q$, the server reports back a value of the fidelity measure to the user. It never reports $D$ or $Q(D)$.

The verification server has some advantages over model servers, also known as remote access systems, that give analysts $Q(D)$. As we shall illustrate, providing infinitely precise information about analyses of $D$, whether in the form of $Q(D)$ or $FM(Q(D), Q(M))$, can lead to high disclosure risks. Arguably, it is easier to coarsen $FM(Q(D), Q(M))$ than $Q(D)$. The fidelity measure is qualitative and so can be coarsened while still providing meaningful information about data quality. The $Q(D)$ is precise; altering $Q(D)$ essentially defeats the purpose of providing it. Additionally, model servers must limit the scope of analyses and details of output, since clever queries can reveal individual data values (Gomatam *et al.*, 2005). The verification server with coarsened $FM(Q(D), Q(M))$ arguably is not as susceptible to such tricks.

## 3 A proposed fidelity measure

Existing utility measures are of two types: (i) comparisons of broad differences between the original and released data, and (ii) comparisons of differences in specific models between the original and released data. Broad difference measures essentially quantify some statistical distance between the distributions of the data on the original and released files, for example a Kullback-Leibler or Hellinger distance. As the distance between the distributions grows, the overall quality of the released

data generally drops. Another approach is based on how well one can discriminate between the original and altered data. For example, Woo *et al.* (2007) stack the original and altered data sets in one file, and estimate probabilities of being "assigned" to the original data conditional on all variables in the data set. When the distributions of probabilities are similar in the original and altered data, the distributions of the variables are similar—this fact comes from the literature on propensity scores for matching in observational studies—so that the altered data have high utility.

Because global measures are only tangentially tied to specific estimands, we consider fidelity measures based on specific models. We use the confidence interval overlap measure of Karr *et al.* (2006). First, the server computes the 95% confidence intervals for the estimand from the masked data, $Q(M) = (L_r, U_r)$, and from the collected data, $Q(D) = (L_o, U_o)$. Then, the server computes the intersection of these two intervals, $(L_i, U_i)$. The fidelity measure is

$$FM(Q(D), Q(M)) = \frac{U_i - L_i}{2(U_o - L_o)} + \frac{U_i - L_i}{2(U_r - L_r)}. \quad (1)$$

When the intervals are nearly identical, corresponding to high utility, the $FM \approx 1$. When the intervals do not overlap, corresponding to low utility, the $FM = 0$. The second term in (1) is included to differentiate between intervals with $\frac{U_i - L_i}{(U_o - L_o)} = 1$ but different lengths. For example, for two masked data intervals that fully contain the collected data interval, the measure (1) favors the shorter interval. Other fidelity measures are possible and could be released as part of the verification server's output.

## 4 Risks associated with measures

Values of $FM(Q(D), Q(M))$ provide intruders with additional information to attempt disclosure attacks. In this section, we describe examples of how intruders might utilize this information when $M$ is constructed with common disclosure limitation strategies.

In what follows, let $X_j$ represent the vector of quasi-identifiers for unit $j$, such as age, race, sex, marital status, and geography. We assume that these are known without error by the intruder for selected records in the database. Let $Y_j$ be the vector of sensitive attributes collected in the survey for unit $j$, such as health, monetary, or other personal variables. We assume that these are not known by the intruder.

### 4.1 Data swapping

Consider a scenario where a small percentage of $X_j$ are swapped, but $Y_j$ is left alone. Whole vectors of $X_j$ are swapped together. The agency does not reveal which or how many records were swapped. As we shall show, given infinitely precise fidelity measure values the intruder can undo many of the swapping protections.

#### 4.1.1 Determining which records underwent swapping

The intruder can use a trial and error approach to determine which records underwent swapping. The intruder submits a query for the confidence interval for the slope in the regression of one variable in $Y$ on some of the variables in $X$. When $FM(Q(D), Q(M)) = 1$, with high probability the values of $X_j$ for the records submitted with the query are the original values. (It is possible that $FM(Q(D), Q(M)) = 1$ by random chance, but with sufficiently large sample size in the query this is a small probability.) When $FM(Q(D), Q(M)) \neq 1$, at least one record has undergone swapping. The intruder can isolate the swapped records by submitting many queries based on different subsets of records. This process can be repeated many times—an efficient algorithm can be devised to reduce the number of queries—to uncover all swapped records.

#### 4.1.2 Determining original values for swapped records

Once the set of swapped records is determined, intruders can determine the original values for those records. To illustrate, suppose the intruder seeks the actual value of marital status for a target record with swapped value of marital status equal to "married." First, the intruder selects a set of unswapped records with sufficient numbers in each marital status category. Second, the intruder appends the target record to this unswapped set. Third, using the appended data in the query, the intruder asks for the fidelity measures for the proportion of people in each marital status category. Any marital status category for which $FM(Q(D), Q(M)) = 1$ is eliminated as a candidate for the target's true marital status. Only two categories have $FM(Q(D), Q(M)) \neq 1$. The target's true marital status is the remaining marital status not equal to the swapped value, e.g., the one other than "married" in our example. The intruder can repeat this process for all swapped values in $X_j$ for all $j$, thus uncovering the true values in $D$.

As another attack strategy, the intruder could construct all possible datasets by permuting $X_j$. For each dataset, the intruder estimates a regression involving one component of $Y$ on $X$, including all records in the regression. With high probability, the dataset for which $FM(Q(D), Q(M)) = 1$ is the actual $D$. This is a computationally expensive strategy for large datasets, particularly when the intruder does not know any details about the swapping strategy.

### 4.2 Top-coding

Suppose now that one component of $Y_j$ for some records $j$ is protected by top-coding; that is, large values of this variable are reported only as exceeding a certain threshold $t$. As we shall show, given infinitely precise fidelity measure values the intruder can undo many of the top-coding protections. In our examples, we assume that the intruder computes $Q(M)$ by setting the top-coded values equal to $t$.

### 4.2.1 Rank order top-coded values

The intruder can utilize the fidelity measure to order the top-coded records from smallest to largest values of the unobserved $Y$. First, the intruder obtains a subset of $n$ records not subject to the top-coding. Second, for some record $j$ subject to top-coding, the intruder appends the record to this subset. Third, the intruder asks for fidelity measures for the mean based on the $n+1$ records in this query. The intruder repeats the second and third steps of this process for each top-coded record, each time using the same $n$ records not subject to top-coding. Finally, the intruder orders the values of $FM(Q(D), Q(M))$ from smallest to largest. This ordering matches the ranking of the unobserved $Y$, since $Q(M)$ successively worsens in quality as the true values deviate from $t$.

### 4.2.2 Determine values of top-coded records

The strategy used in Section 4.2.1 can be modified to learn exact values of top-coded records. As before, the intruder obtains $FM(Q(D), Q(M))$ for a subset of $n$ records with $Y < t$ and one top-coded record $j$. Then, the intruder guesses a plausible value of the true $Y_j$ for that record, using that guess to make a proposed true dataset, $D^*$, for those $n + 1$ records. The intruder computes $FM(Q(D^*), Q(M))$. The intruder repeats this process many times using different initial guesses of the true $Y_j$. The guess that results in $FM(Q(D^*), Q(M)) = FM(Q(D), Q(M))$ and exceeds $t$ is the true value for that record.

## 4.3 Added noise

With clever transformations, the intruder can estimate the actual values of variables protected by additive noise. Specifically, the intruder submits a query based on $n$ records in which all values but the one for record $j$ are transformed to equal the same number. The intruder obtains $FM(Q(D), Q(M))$ for that query. Then, the intruder guesses a plausible value of the true $Y_j$, using that guess to make a proposed true dataset, $D^*$, for those $n$ records. The intruder computes $FM(Q(D^*), Q(M))$. The intruder repeats this process many times using different initial guesses of the true $Y_j$. The guesses that result in $FM(Q(D^*), Q(M)) = FM(Q(D), Q(M))$ are candidates for the true value. Some of these values may be more plausible than the others, given other information released in the data.

## 4.4 Synthetic data

The attack strategies described for the other disclosure limitation methods can be applied on partially synthetic data, for which there is a one-to-one correspondence between $D$ and the (multiple sets of) released $M$. As an example, to learn the original value $Y_j$ for a record with $Y$ synthesized, the analyst appends this record to a set of $n$ records whose values of $Y$ are not synthesized. The intruder guesses a plausible value of the true $Y_j$, using that guess to make a proposed true dataset, $D^*$, for those

$n+1$ records. The intruder computes $FM(Q(D^*), Q(M))$. The intruder repeats this process many times using different initial guesses of the true $Y_j$. The guesses that result in $FM(Q(D^*), Q(M)) = FM(Q(D), Q(M))$ are candidates for the true value. This attack is not possible when all values of the variable are synthesized.

# 5 Reducing risks for measures

The key factors driving the risks outlined in Section 4 include (i) the ability of intruders to submit queries based on subsets of records and transformations of variables and (ii) the availability of infinitely precise fidelity measures. Thus, to reduce these risks, it makes sense to limit what queries are answered with fidelity measures or to coarsen the reported fidelity measures.

When limiting queries, we should preserve as much as possible the ability of the legitimate user to obtain fidelity measures for complex models. Otherwise, the verification server has limited usefulness. When coarsening fidelity measures, we should provide the analyst with enough information to decide whether or not $Q(M)$ is "good enough" compared to $Q(D)$ for publication. The agency should not make that decision, as different analysts will have different quality requirements.

## 5.1 Limiting the query space

The verification server need not report fidelity measures for all possible queries. Certain attacks can be made much more difficult with query restrictions that may have minimal impact on legitimate data users. This section describes some query restrictions.

### 5.1.1 Require well-defined target populations

The verification server might require that records in any queries comprise well-defined sub-populations rather than arbitrary collections. For example, the server could require selection criteria to be functions with limited complexity, such as a maximum of three-way interactions among variables (e.g., all women under age 25 living in a city). The verification server could be programmed to force the analyst to enter the subset selection criteria rather than specify a collection of record identification numbers. Such restrictions go a long way towards defeating attacks based on subsetting records. They are not foolproof; a clever intruder might be able to cobble together desired records from legal subsetting requests.

### 5.1.2 Disallow certain transformations

Some transformations, such as those described in Section 4.3, are not useful for many legitimate applications but are useful for disclosure attacks. It may be possible to prevent some of these transformations. One approach is to allow a set of standard transformations, such as polynomials and low powers, and disallow all others. This is tricky, however, because the restrictions may rule out legitimate uses.

### 5.2 Coarsening the fidelity measure

Coarsening fidelity measures creates uncertainty in all of the attacks, because the true fidelity measure is not available for backsolving. This section describes some approaches to coarsen fidelity measures.

#### 5.2.1 Interval reporting

Rather than report precise measures, the verification server can report measures in intervals; for example, between 90% and 100% overlap, between 80% and 90% overlap, etc. Such reporting may be sufficient to enable the user to evaluate quality, yet provide enough uncertainty in the fidelity measures to mask true values. For example, if the server never reports $FM(Q(D), Q(M)) = 1$ precisely, then the attacks on swapped data described in Section 4.1 can be prevented. However, there is no guarantee that interval reporting prevents all attacks. For example, when seeking to order top-coded records, the intruder can determine which record owns the largest value when the interval with the least overlap (e.g., 10% to 20% has smaller overlap than 30% to 40%) is unique.

Intruders can use trial and error attacks like those in Section 4.2.2 – 4.4 to obtain ranges of possible values for unknown $Y$. For example, to get a range for a particular top-coded value $Y_j$, the intruder appends that record to a collection of $n$ other records that are not top-coded. The intruder submits queries about the mean of $Y$ based on those $n+1$ records, obtaining the reported interval $FM(Q(D), Q(M))$. The intruder then proceeds as follows.

A1. Set $f$ equal to the lower bound of the reported interval for $Q(M)$.

A2. Find the value $Y_j^*$ such that, when used to make a plausible true dataset $D^*$, produces $FM(Q(D^*), Q(M))^* = f$, where $FM(Q(D^*), Q(M))^*$ is a single number measurement of the fidelity measure. Store this value of $Y_j^*$.

A3. Set $f = f + c$, where $c$ is some small number like 0.5.

A4. Repeat Step 2 and 3 until $f$ equals the upper bound of the reported interval for $Q(M)$.

The stored values $Y_j^*$ are the plausible values for the original $Y_j$. When this distribution does not have sufficient variance, an inferential disclosure occurs.

To illustrate these attacks, we use data from the 1995 U.S. Current Population Survey. The variables include adjusted gross income (AGI) and amount of interest income (INTVAL). The intruder submits a query for the fidelity measure for the intercept in a regression of AGI on INTVAL. The regression is based on a set of 31 records, one of which has AGI top-coded. The intruder seeks to estimate the original value of the top-coded AGI. Figure 1 graphically displays the output from the attack protocol described above. The curved line connects the values of

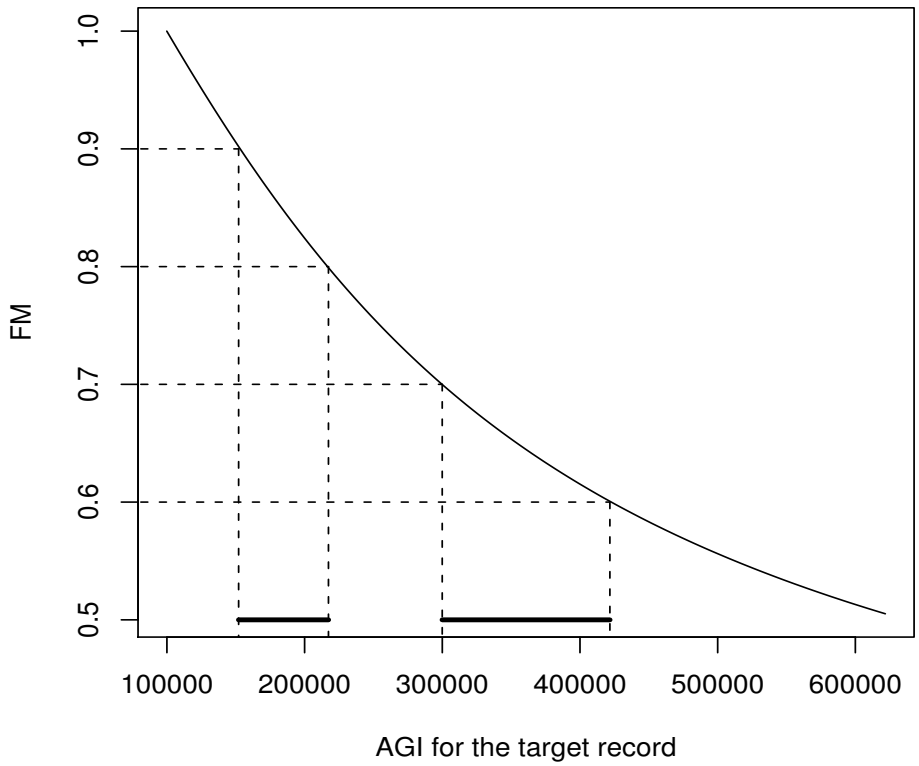


Figure 1: Illustration of intruder's bounding procedure when server reports interval fidelity measures. The solid horizontal lines are ranges of plausible values for original AGIs when the reported interval is (0.6, 0.7) or (0.8, 0.9).

$FM(Q(D^*), Q(M))^*$ for several guesses at the original AGI. If the top-coded record is such that the server reports an interval of (0.8, 0.9), the intruder knows that the original AGI is between \$152,082 and \$217,318. If the top-coded record is such that the server reports (0.6, 0.7), the original AGI is between \$299,950 and \$421,723.

To reduce confidentiality risks when providing interval measures, the server should report as wide an interval as necessary to generate sufficient variability in the prediction distribution, but not so wide as to destroy the usefulness of the interval measure. To do so, the agency can define acceptable bounds of uncertainty for at risk values, such as large incomes. Then, for any given query, the server determines the maximum and minimum values of the fidelity measures that yield these acceptable bounds, and it reports an interval that contains these bounds. This approach still is risky if, for example, the intruder can get close to the true $Y_j$ by backsolving with the midpoint of the reported interval. Additionally, the intruder may be able to sharpen bounds by submitting multiple queries involving the same target value.

A related approach is to add random noise to the fidelity measures before re-

porting, again with the goal of obscuring the true fidelity measure. The amount and distribution of the noise should be hidden from the analyst to provide less information for backsolving attacks. The noise should be the same for all requests for the same $Q(M)$ to eliminate intruders' ability to sharpen estimates by averaging the results of multiple queries for the same analysis. But, the noise must differ by analysis to reduce the chance that the intruder can guess the value of the added noise for some queries, for example if the intruder knows $Q(D)$ and submits $Q(M)$ anyway. These two criteria can be met by tieing the random seed to some function of $Q(D)$ that provides unique seeds for almost all different queries.

The intruder could treat the reported, noisy fidelity measure as a true value, and attempt trial and error attacks. This suggests that the noise distribution should be based on acceptable bounds for risk, as described for the interval measures.

### 5.2.2 Computing on different datasets

As another approach, the verification server can report fidelity measures based on datasets that differ slightly from $D$ and $M$. The server does not tell analysts how the datasets differ. As an example, the verification server can do the following.

B1. Delete $k$ randomly sampled records from the data used to compute $Q(D)$. Let $r_d$ and $r_n$ be the row numbers of the deleted records ($r_d$) and the not deleted records ($r_n$). Let $D_{r_n}$ and $M_{r_n}$ be the data in $D$ and $M$ for the records in $r_n$.

B2. Sample $k$ row numbers from $r_d$, with replacement. Let $r_s$ be the sampled row numbers. Let $D_{r_s}$ and $M_{r_s}$ be the data in $D$ and $M$ for the records in $r_s$.

B3. Construct $D' = (D_{r_n}, D_{r_s})$ and $M' = (M_{r_n}, M_{r_s})$.

B4. Report $FM(Q(D'), Q(M'))$ to the analyst.

With large $k$, this creates a combinatorial explosion of possible true values for the records in $D$. A related approach for model servers was suggested by Steel and Reznek (2006).

This approach is not immune to disclosure risks in verification servers, although the intruder must work hard to guess original values sensibly. First, the intruder proposes a possible value of $D$, say $D^*$. Second, the intruder proposes a possible value of $D'$ and $M'$, say $D'^*$ and $M'^*$, obtained by mimicking steps B1 – B3 on $D^*$ and $M$. Third, the intruder computes $FM(Q(D'^*), Q(M'^*))$. The intruder repeats this process many times, collecting all $D'^*$ for which $FM(Q(D'^*), Q(M'^*)) = FM(Q(D'), Q(M'))$. The values of the targeted $Y_j$ among the $D'^*$ meeting this criterion are plausible values of the original $Y_j$. If the distribution does not have sufficient variability, or if some obvious function like the mean/median of the plausible values is too close to the truth, there may be a disclosure risk.

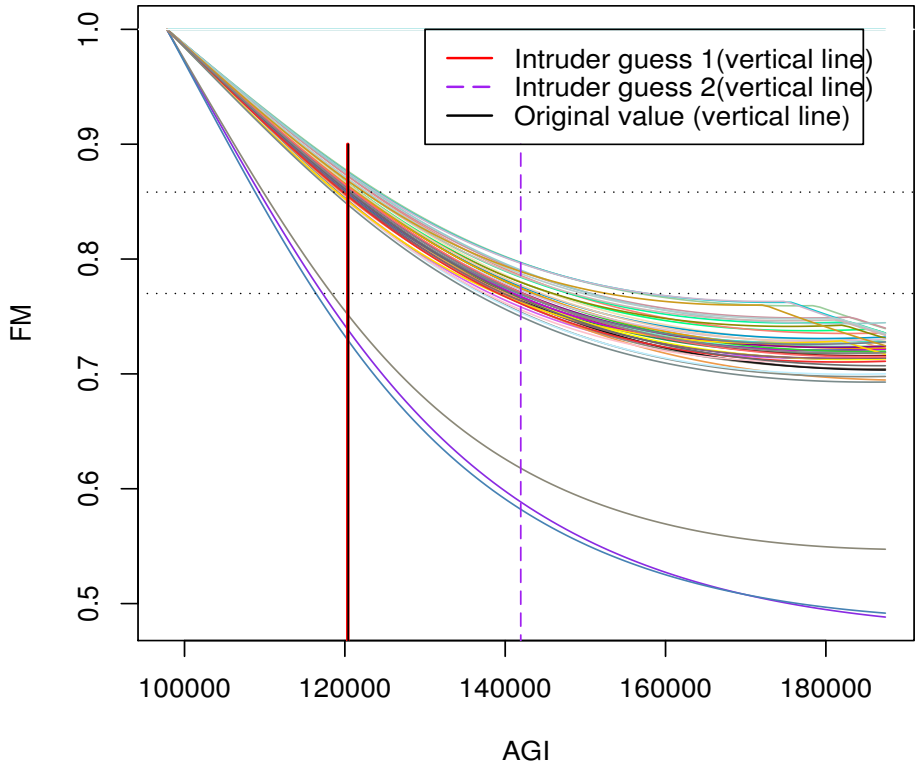


Figure 2: $FM$ curves constructed by the intruder to defeat the *delete-add* strategy of the verification server.

Figure 2 illustrates this attack strategy for one top-coded value of AGI. The intruder's query is based on $n$ records with $Y < t$ and one record subject to top-coding. To draw any one curve, we assume that the intruder knows $k$ and follows steps B1 – B3 only with $M$ to get $M'^*$. For all top-coded records in $M'^*$, we replace $t$ with a guess of the original AGI, thus obtaining a $D'^*$. We then compute $FM(Q(D'^*), Q(M'^*))$. We repeat the last two steps for different guesses, and connect values of $FM(Q(D'^*), Q(M'^*))$ to make the curve. The figure shows 50 such curves, each derived from different records in $M'^*$. The intruder draws a horizontal line at the reported $FM(Q(D'), Q(M'))$. Its intersections with the curves gives the distribution of plausible values for the original datum. For one realization of $D'$ we obtained $FM(Q(D'), Q(M')) = 0.87$. For this $D'$, the intruder can average the plausible AGI values to get very close to the original value. However, for another realization of $D'$ we obtained $FM(Q(D'), Q(M')) = 0.77$, for which the average of the plausible AGI values is not a close estimate of the original value. Although 0.87 and 0.77 are not that different in terms of data quality, Figure 2 indicates that protection is sensitive to which units are in $D'$.

## 6 Concluding remarks

Verification servers could have enormous benefit for statistical agencies and consumers of their data. However, releasing precise quality measures could threaten confidentiality. The examples in this article suggest that both restricting queries and coarsening fidelity measures show promise for providing sufficient protection.

## References


Abowd, J. M. and Woodcock, S. D. (2004). Multiply-imputing confidential characteristics and file links in longitudinal linked data. In J. Domingo-Ferrer and V. Torra, eds., *Privacy in Statistical Databases*, 290–297. New York: Springer-Verlag.

Gomatam, S., Karr, A. F., Reiter, J. P., and Sanil, A. P. (2005). Data dissemination and disclosure limitation in a world without microdata: A risk-utility framework for remote access servers. *Statistical Science* **20**, 163–177.

Karr, A. F., Kohnen, C. N., Oganian, A., Reiter, J. P., and Sanil, A. P. (2006). A framework for evaluating the utility of data altered to protect confidentiality. *The American Statistician* **60**, 224–232.

Little, R. J. A. (1993). Statistical analysis of masked data. *Journal of Official Statistics* **9**, 407–426.

Reiter, J. P. (2003). Inference for partially synthetic, public use microdata sets. *Survey Methodology* **29**, 181–189.

Reiter, J. P. (2004). Simultaneous use of multiple imputation for missing data and disclosure limitation. *Survey Methodology* **30**, 235–242.

Rubin, D. B. (1993). Discussion: Statistical disclosure limitation. *Journal of Official Statistics* **9**, 462–468.

Steel, P. and Reznek, A. (2006). Issues in designing a confidentiality preserving model server. In P. D. Munoz and H. Brungger, eds., *Monographs of Official Statistics: Work Session on Statistical Data Confidentiality*, 29–36. Eurostat.

Willenborg, L. and de Waal, T. (2001). *Elements of Statistical Disclosure Control*. New York: Springer-Verlag.

Woo, M. J., Reiter, J. P., Oganian, A., and Karr, A. F. (2007). Global measures of data utility for microdata masked for disclosure limitation. *Journal of Privacy and Confidentiality* forthcoming.