Evaluating Noise Infusion for Disclosure Protection for Two Time Periods

Jared Martin

U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233

Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference

Note: This article reports the results of research and analysis undertaken by U.S. Census Bureau staff. It has undergone a more limited review than official U.S. Census Bureau publications. It is released to inform interested parties of research and to encourage discussion. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

Abstract

The U.S. Census Bureau conducts the Economic Census of Island Areas quinquennially. For the 2007 Economic Census of Island Areas, the U.S. Census Bureau provided disclosure protection, a protection detailed in Titles 13 of the U.S. Code, using a noise infusion technique. Noise infusion consists of perturbing values for each contributing establishment prior to table creation by applying a random noise multiplier to the magnitude data (i.e., characteristics such as receipts, payroll, and number of employees). The use of noise factors introduces uncertainty into the published cell values that protects the confidentiality of the contributors to the cell estimates. In prior economic censuses of Island Areas, disclosure protection was accomplished using cell suppression, which reduces the number of published values significantly. Census Bureau staff considered the impact of noise infusion on estimates of change in published cell values from one census cycle to the next when planning for disclosure protection for the 2012 Economic Census of Island Areas. They considered several options for applying noise in the 2012 economic census in an attempt to minimize distortion in cycle-to-cycle changes while providing sufficient uncertainty to protect the confidentiality of the respondents. This paper presents the results of a simulation study used to compare and evaluate the options considered, for estimates of varying size and composition. The paper illustrates each method and shows the tradeoff between bias and low variability in an estimate due to noise.

1. Introduction

Statistical disclosure avoidance methods are procedures designed to protect the confidentiality of individuals, businesses or other entities that report data used to create data products for public use. Protecting confidentiality involves withholding enough information to protect the privacy of individual respondents (Cox and Zayatz, 1995). Disclosure avoidance methods restrict the amount of useable data presented to the public. The goal of any disclosure avoidance method is to maximize the amount of useable information published while ensuring sufficient protection and privacy of individual contributors represented in the data.

The U.S. Census Bureau provided disclosure protection for the 2007 Economic Census of Island Areas using noise infusion. In an effort to publish more data while protecting confidentiality of the respondents in the Economic Census of Island Areas in 2007, the Census Bureau implemented a method that assigns and multiplies micro level data by a multiplicative factor called a noise factor. The use of noise factors introduced uncertainty into an estimate that was sufficient to protect the confidentiality of the respondents and their data. This uncertainty decreased the usability of the estimates, and introduced variability or uncertainty in changes in estimates from cycle to cycle. This paper begins with a brief explanation of the Economic Census of Island Areas (IA), followed by a description of noise infusion and an explanation of how the U.S. Census Bureau applied it to IA in 2007. Then it describes some concerns with the continued use of the current noise methodology, and some suggestions. Finally, we explain the results of a Monte Carlo simulation study replicating each proposed method and find the method that minimizes the uncertainty of estimates between census cycles to the next while providing sufficient protection of the confidentiality of the respondents.

2. The Economic Census of Island Areas

The Economic Census of Island Areas (IA) provides economic data for American Samoa, the Commonwealth of the Northern Mariana Islands (CNMI), Guam, Puerto Rico and the U.S. Virgin Islands. It is a quinquennial census conducted in years ending in 2 and 7, providing information about the structure and functioning of the economy for these islands (U.S. Census Bureau 2013). For purposes of this paper, we focus on the variable sales defined as "total monetary collected for goods and services provided by a business" (U.S. Census Bureau 2013). U.S. Census Bureau staff collects and publishes sales at the four-digit North American Industry Classification System (NAICS) level for each of the islands.

3. Noise Infusion for Disclosure Avoidance

Noise is one method of protecting respondents' data from being disclosed. This method perturbs all the micro data by a small percentage. Each contributor is protected by the amount the perturbation distorts its contribution. A multiplicative noise factor distorts the input data by multiplying each contribution by a factor drawn from a symmetric distribution centered on one. This distortion will guarantee minimum required protection by choosing a factor distribution with no probability within the desired minimum required protection surrounding one (Evans, Zayatz, Slanta 1998)**.** In aggregate, the estimate should have minimal distortion since the factor distribution is symmetric about one. If an estimate is sensitive, with a large dominant contributor, the distortion will be larger to guarantee minimum required protection.

3.1 Balanced Noise as Applied to Economic Census of Island Areas 2007

In 2007, the Economic Census of Island Areas (IA) assigned noise factors generated from a split triangular distribution symmetric about one and that meets the qualifications discussed above (See Figure 1). The expected value of a factor drawn from this distribution is intuitively one, and the expected value of the total noise distortion of each contributor is zero. Since overall distortion is the sum of the distortion of each contributor, and the expected value of the distortion of each contributor is zero, the expected value of overall distortion is zero.

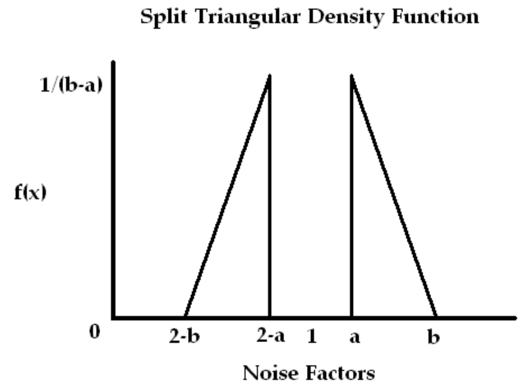
Distortion in either direction is between two values (a-1)% and (b-1)%, while the expected value of the factor is one and expected distortion is zero (See Figure 1). The expected value of noise for any cell is not necessarily zero. If a company reports with more than one establishment, we assign each establishment's noise factor in the same half of the distribution, giving them all the same direction as the company's noise factor, i.e. generated from the same side of the distribution. Each establishment also maintains the same factor throughout all cells in a publication.

Noise factors are 'balanced' using sales at the industry level. Balancing lessens the overall distortion a table cell or estimate receives. The variable and level chosen to balance the noise will have the lowest variance in the distribution of distortion due to noise of all estimates in a table. For IA, we chose reported sales at the industry category level as the most important estimates to have the least distortion.

To 'balance' the noise factors, we sorted all establishments contributing to an industry estimate in descending order of sales and assigned the largest contributor a noise factor. After each assignment, we measured overall distortion, for sales at the industry level. If the overall noise was one direction for the cell, we assigned the opposite direction to the next company. For example, if noise for sales in an industry was positive, then we assigned a negative direction to the noise factor for the next added contributor. A negative direction for noise meant the factor was selected from the part of the distribution below one and a positive direction for noise meant the factor was selected from the part of the distribution above one. We repeated this process until all establishments were assigned a factor corresponding to the appropriate direction.

After we had applied balanced noise, the confidentiality of respondents is sufficiently protected. The estimates may be distorted by noise to varying extents or have data quality issues. To alert the consumer to poor data reliability, we suppress from publication cells with fewer than three contributors, and assign noise flags alerting data consumers to whether the cell had low, medium, or high noise. For IA, cell values changing less than 2 percent due to noise have a low noise flag. Cell values changing by 2 percent or more, but less than 8 percent have a medium noise flag, and cell values changing by more than 8 percent have a high noise flag and are suppressed from publication.

Figure 1



3.2 Concerns About Noise in Future Publications

Each contributing establishment may help balance out each other's noise contribution. A problem arises when a cell has relatively few contributing companies or a cell has a large company driving the estimate. These scenarios may cause the cell estimate to have relatively large distortion because there are fewer opportunities (establishments) to balance out the distortions caused by the largest contributors to the cell. This is what the noise infusion is designed to do. It provides protection by creating a less precise estimate so an estimate can be released instead of suppressed (Evans, Zayatz, Slanta 1998). When a table cell has many contributing companies, the overall noise will be small with small variation for overall noise.

A concern arises when implementing noise in consecutive subsequent cycles. If variation in cell noise is large, an estimate may swing from large positive noise to large negative noise in the next census cycle. Such a large swing in the value of an estimate may mislead the data consumer into thinking something noticeable is happening in the economy when in fact, the noise factors changed from one cycle to the next. For example, suppose a cell has three contributors, companies A, B, and C. Their noise factors are drawn independently from a split triangular distribution with a=1.1 and b=1.15 each cycle (See Figure 1). The first census would report sales of $12.3 million. Next census cycle would report $9.6 million (See Table 1). Since census numbers can be directly compared, it appears there is a 22% drop in sales even though the values were identical.

Table 1

	Year 1			**Year 2**		
Company	**Actual Sales** (millions of dollars)	**Noise Factor**	**Noisy Data** (millions of dollars)	**Actual Sales** (millions of dollars)	**Noise Factor**	**Noisy Data** (millions of dollars)
A	**10.0**	**1.14**	**11.4**	**10.0**	**0.85**	**8.5**
B	**0.5**	**0.90**	**0.45**	**0.5**	**1.10**	**0.55**
C	**0.5**	**0.90**	**0.45**	**0.5**	**1.10**	**0.55**
Total	**11.0**		**12.3**	**11.0**		**9.6**

3.3 New Methods Suggested

Preserving as much of the change between cells from one period to the next is the objective. This means reducing the variability of overall effective noise a cell receives from period to period. We proposed to modify the current noise application method. One way is to fix the noise factor direction and magnitude each time. Each cell still has noise applied, but only once and held constant over time, better preserving differences between censuses. However, completely removing variability of effective noise a cell receives by fixing the noise factor for every company from period to period is data disclosure when the actual change from year to year of a dominant company in a cell can be identified with a certain precision (Evans, Zayatz, Slanta 1996). Some form of noise factor change needs to happen from one period to the next in order to protect these trend statistics. In practice, even if direction and magnitude are fixed, there are companies entering and leaving the census from cycle to cycle which necessitates assigning new factors, adding some change in the amount of effective noise in a cell. We elected to do a simulation study replicating choosing noise factor direction and magnitude many times under differing scenarios of fixing some noise factor directions and magnitudes and not others from period to period.

4. Evaluation Study and Results

To conduct our simulation study, we used data from the 2002 and 2007 IA. For the 2002 Economic Census of the Island Areas, we achieved disclosure avoidance through cell suppression methodology. We changed to using the 'balanced' noise application method mentioned in section 3.1 for the 2007 IA. Since we are concerned with cells that have relatively few contributing companies or cells with a large case that drives the estimate, we decided to begin simulations with sales data from American Samoa, the island with the least number of companies. In total, we simulated 10 different noise application methods on sales data from American Samoa and the Virgin Islands, also an island area that has industries composed of only a few contributors.

Method 1 - Reassigned Noise Direction and Magnitude

First, we simulated applying noise 1500 times to American Samoan companies reporting 2007 data, using the current balanced noise methodology explained in section 3.1, randomly selecting noise factors each time. This created a baseline of current methodology for comparison. The current methodology for the Economic Census of Island Areas applied balanced noise for cells containing sales data at the industry category level. We recorded the overall percent distortion or noise of sales estimates at three digit NAICS industry categories after each replication.

Method 2 - Fixed Noise Direction and Magnitude

Second, we simulated using current methodology on 2007 American Samoa sales data but this time keeping direction and magnitude fixed from replication to replication for companies appearing in both 2002 and 2007 to simulate companies that may be retained from period to period of the census. Fixing the factor for each establishment appearing from cycle to cycle is the most restrictive method, allowing the least amount of variation in overall effective noise. In fact, if every establishment remained in business and reported every cycle there would be no variation in noise factors. We recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time, repeating the process 1500 times.

Method 3 – Fixed Noise Direction

Fixing both direction and magnitude is quite restrictive, especially for estimates that only have a few large contributing establishments. The next simulation involved fixing only the noise factor direction for establishments appearing from cycle to cycle. We simulated using current methodology on 2007 American Samoa sales data, keeping only the direction fixed from replication to replication for companies appearing in both 2002 and 2007. We repeated assigning noise in this way 1500 times, recording the overall percent distortion of sales estimates at three digit NAICS industry categories each time. Fixing the direction of the noise from cycle to cycle reduced variability in the noise distortion of an estimate, at the expense of a biased estimate. Estimates tended to either be positively distorted with noise or negatively distorted with noise depending on the fixed direction of the cases present from cycle to cycle.

Method 4 - Fixed Noise Direction and Magnitude for Large Cases

Next, we simulated using current methodology on 2007 American Samoa data but keeping direction and magnitude fixed for large cases appearing in both 2002 and 2007. We defined large cases as those cases that contributed more than 50% of an estimate or were in the top two contributors that together contributed more than 75% of an estimate. Cases contributing more than 50% to an estimate greatly influence overall noise distortion in an estimate. Even the balancing method we use will rarely counteract and mitigate the distortion or noise in the cell estimate contributed by such a large case. We also chose to fix noise factor magnitude and direction for cases in the top two contributors that together contributed more than 75% of an estimate to control for cases that may together contribute a large amount to an estimate and hence the amount of noise in an estimate. These top two contributors may balance each other's contribution of noise, but by fixing the noise direction and magnitude, we were able to control one of the largest sources of noise variation in an estimate from cycle to cycle. We repeated the simulation 1500 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time.

After analyzing plots of the simulations, we decided that fixing the noise direction of companies, especially large companies, appeared to reduce the overall noise variability in the overall estimate. Fixing the magnitude further reduced variation of the distortion, but raised concerns about not providing enough protection (See Appendix A) since fixing the factors essentially preserved the difference in estimates from cycle to cycle and the estimates would nearly be exact except for a nearly fixed constant added to the overall estimate. We decided to pursue fixing direction of "large" companies and allowing the magnitude of their factors to vary slightly from the original. We also decided to do simulations on the U.S. Virgin Islands data to see if the Virgin Island estimates that also have few contributors behave similarly under different noise conditions as compared to estimates of sales in American Samoa.

After looking more closely at the data, we chose to define a "large" company as any company that contributed more than 50% of an estimate. Companies that contributed more than 50% of an estimate contribute noise that rarely is balanced out by the rest of the contributing companies. In fact, if a company contributed more than 50 + (b-a)/2 % of an estimate, its noise can never be completely balanced and the estimate will always be distorted. Few companies contributed slightly less than 50% of an estimate but together with another company made up the top two contributors and were more than 75% of an estimate. These cases were easily balanced out and did not appear to cause too much variability in the amount of noise an estimate received. We decided to include companies that together contributed more than 75% of an estimate in the definition of large for Method 10, as a continuation of Method 4.

We repeated the first four simulations with 3000 repetitions for American Samoa and Virgin Island sales data to get a more defined picture of the distribution of distortions to each estimate. Similar scenarios and outcomes for noise application appear in the Virgin Island estimates and the American Samoa estimates. The results for select three digit NAICS industry categories appear in Appendix A.

Method 5 – Fixed Noise Direction for Large Cases, but Magnitude Constrained Around an Initial Value

Method 5, the next attempt at controlling the possible large swings in effective noise distortion to an estimate, involved fixing the noise direction for large establishments, but allowing the factor magnitude to vary from cycle to cycle around the initial noise factor assignment. We chose to fix the noise direction for establishments that contributed more than 50% of an estimate and appeared from cycle to cycle. If a contributor was larger than 50% of an estimate, it drove the estimate, and its noise factor drove the overall direction of noise in the estimate. For this method, we simulated using current methodology on 2007 Virgin Island and American Samoan sales data but kept direction fixed at initial values for companies (single establishments, or sum of establishments in a multi-unit company within an estimate) appearing in both 2002 and 2007 and made up more than 50% of an estimate in 2007. We also reassigned the magnitude of the noise factors of companies with fixed direction from a triangular distribution centered on the old previously assigned noise factor. The triangular distribution we first tried had parameters c= old noise factor and b= c + 0.0125 (see Figure 2). We did not allow new factors to exceed the bounds a standard noise factor may take. If a factor chosen from the special distribution exceeded the bounds a standard noise factor may take, we re-selected a new factor until the new factor fell within the bounds of a standard noise factor. We repeated the simulation 3000 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time. See the results for select three digit NAICS industry categories in Appendix A.

Method 6 – Fixed Noise Direction for Large Cases, but Magnitude Constrained Around an Initial Value

Method 6 was similar to method 5. As in method 5, we simulated using current methodology on 2007 Virgin Island and American Samoan sales data. We kept direction fixed at assigned 2002 values for companies (single establishments, or sum of establishments in a multi-unit company within an estimate) that appeared in both 2002 and 2007 and made up more than 50% of an estimate. For these companies with fixed noise direction, we allowed the magnitude of the factors to be reassigned from a triangular distribution centered on the old previously assigned noise factor. For method 6, when reassigning a noise factor, we chose a narrower distribution with parameters c= old noise factor and b= c + 0.005 (see Figure 2). This method provided more restriction than method 5 in the amount a noise factor with fixed direction could change. We did not allow new factors to exceed the bounds a standard noise factor may have taken. If a factor chosen from the special distribution exceeded the bounds a standard noise factor may have taken, we re-selected a new factor until the new factor fell within the bounds of a standard noise factor. We repeated the simulation 3000 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time. See the results for select three digit NAICS industry categories in Appendix A.

Method 7 – Fixed Noise Direction for Large Cases

Method 7 used current methodology on 2007 American Samoa and Virgin Island sales data but kept direction fixed for companies (single establishments, or sum of establishments in a multi-unit company within an estimate) that appeared in both 2002 and 2007 and that made up more than 50% of an estimate in method 7. Cases contributing more than 50% to an estimate greatly influence overall noise distortion in an estimate. Even the balancing method we use will rarely counteract and mitigate the distortion or noise in the cell estimate contributed by such a large case. We also chose to fix the direction for companies with several establishments contributing to an estimate that together contributed more than 50% of an estimate to control for cases that have the same direction such as establishments in a company and may have together contributed a large amount to an estimate and hence the amount of noise an estimate had. We repeated the simulation 3000 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time. Results for select three digit NAICS industry categories appear in Appendix A.

Method 8 – Fixed Noise Direction for Very Large Cases

Method 8 fixed the noise direction for establishments appearing in both 2002 and 2007 and that made up more than 75% of an estimate and randomly assigned all other cases. We simulated using current methodology on 2007 American Samoa data but kept the noise direction fixed for establishments that appeared in both 2002 and 2007 and made up more than 75% of an estimate and used current methodology to assign direction and magnitude for all other cases. We chose to fix the direction of establishments that contributed more than 75% of an estimate because establishments this large contribute so much noise that the rest of the contributors cannot balance by assigning a noise factor in the opposite direction. Even if we assigned the large establishment the smallest amount of noise, and all the other establishments the largest amount of noise in the other direction, the overall estimate would always be biased. The change in an estimate from cycle to cycle may be overshadowed by whatever noise factor and direction the large establishment may have unless we at least fix its noise direction. See Appendix A for results for select three digit NAIS industry categories.

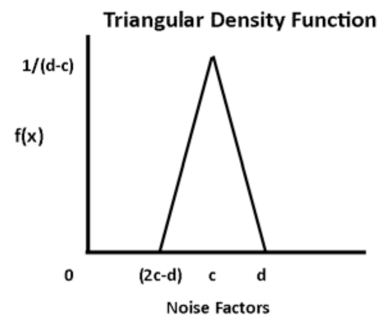
Method 9 - Fixed Noise Direction for Large Cases or Companies with Multiple Establishments, but Magnitude Constrained Around an Initial Value

For method 9, we simulated using current methodology on 2007 Virgin Island and American Samoan sales data but kept noise direction fixed at assigned 2002 values for establishments that appeared in both 2002 and 2007 and made up more than 50% of an estimate or were a company with more than one establishment. Since each company within a multi-unit establishment had the same direction, they could potentially severely bias the noise in the overall estimate if there were many companies from a multi-unit establishment in any one estimate. We allowed the factors of companies with their direction fixed to be reassigned from a triangular distribution centered on the old previously assigned noise factor, with parameters c = old noise factor and $b = c + 0.0125$ (see Figure 2). We did not allow new factors to exceed the bounds a standard noise factor may take. If a factor chosen from the special distribution exceeded the bounds a standard noise factor may take, we chose a new factor until the chosen factor fell within the bounds of a standard noise factor. We repeated the simulation 3000 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time. See Appendix A for histograms of the results for select three digit industry categories.

Method 10 – Fixed Noise Direction for Large Cases

In method 10, we simulated using current methodology on 2007 American Samoa and Virgin Island sales data but kept noise direction fixed for companies that appeared in both 2002 and 2007 and were single establishments or multiple establishments in a company contributing to an estimate that together made up more than 50% of an estimate or were in the top two contributors that together made up 75% of an estimate. Cases that contributed more than 50% to an estimate greatly influenced overall noise distortion in an estimate. Even the balancing method we used will rarely counteract and mitigate the distortion or noise in the cell estimate contributed by such a large case. We also chose to fix the noise factor direction for cases in the top two contributors that together made up more than 75% of an estimate to control for cases that may have together contributed a large amount to an estimate and hence the amount of noise an estimate had. These top two contributors may have balanced each other's contribution of noise, but by fixing the noise direction for them, we were able to control the largest sources of noise variation in an estimate from cycle to cycle. We repeated the simulation 3000 times and recorded the overall percent distortion of sales estimates at three digit NAICS industry categories each time. Appendix A contains histograms of the results for select three digit industry categories.

Figure 2



5. Discussion

After reviewing methods 5, 6, and 7 we noticed that, for companies with fixed direction, the amount of constraints placed on the factor size seemed to matter little. The histograms showing the distribution of potential noise in an estimate for American Samoa and the Virgin Islands appeared almost identical in methods 5, 6, and 7. All of these methods, for the most part, reduced variability due to noise, and produced distributions only slightly biased or centered on zero distortion, which is a desirable attribute.

Method 8 produced distributions of overall percent noise very similar to method 1. This result is understandable because very few establishments contributed more than 75% of an estimate, and hence, we fixed the noise direction of very few establishments. This meant that most estimates would not have any establishments fixed, which was method one. One notable exception was industry section 443 *Electronics and Appliance Stores* in American Samoa. In this instance, the distribution of overall percent noise resembled a split triangular distribution in method 1, but in method 8, the distribution was only the lower side of what appeared to be the same split triangular distribution.

Methods 5 and 9 were nearly identical except in method 9 the direction of the noise factor in all multi-unit establishments was fixed instead of only the establishments that contributed more than 50% of an estimate. Thus, method 9 produced similar results as method 5, with a few exceptions. For example, industry section 339 *Miscellaneous Manufacturing* in the Virgin Islands displayed slight negative bias in method 9 compared to method 5. See Appendix A.

Methods 3 and 10 produced distributions of overall percent noise that were similar. Neither method stood out as superior. Sometimes the variance was larger for percent noise under method 3 (such as in industry section 524 *Insurance Carriers and Related Activities* in American Samoa) than under method 10, and sometimes the variance was smaller under method 3 (such as in industry section 551 Management of Companies and Enterprises in American Samoa) than under method 10.

Reducing variability in an estimate due to noise was one goal. To assess this goal, we calculated the inner quartile range (IQR) of each estimate for each method simulated in section 4. We calculated the IQR by subtracting the 25th percentile from the 75th percentile of the distribution created from the noise simulations explained in section 4, at the three-digit industry category. There are a total of 42 estimates for American Samoa, and 61 estimates for the Virgin Islands for each method. Table 2 summarizes the number and percentage of estimates that had a reduced IQR compared to method 1, the base line method. For American Samoa, method 3 had the largest number of estimates with a reduced IQR, except for method 2, where direction and magnitude are fixed. Method 2 is not advisable since it could potentially lead to a data disclosure as mentioned earlier. Methods 3 and 9 reduced the IQR almost as much as Method 2 for the Virgin Islands. Based on the reduced IQR criterion, Method 3 is the most desirable after method 2.

Table 2

	American Samoa			**The Virgin Islands**		
Method	**Number of estimates with IQR < IQR Method 1**	**Number of estimates**	**Percent of cells with reduced IQR**	**Number of estimates with IQR < IQR Method 1**	**Number of estimates**	**Percent of cells with reduced IQR**
2	18	42	42.9	13	61	21.3
3	15	42	35.7	11	61	18.0
4	13	42	31.0	10	61	16.4
5	12	42	28.6	9	61	14.8
6	12	42	28.6	9	61	14.8
7	12	42	28.6	9	61	14.8
8	11	42	26.2	8	61	13.1
9	14	42	33.3	12	61	19.7
10	14	42	33.3	8	61	13.1

Another goal was to preserve as much of the change between estimates from one period to the next without creating a data disclosure. Therefore, we needed to reduce the variability of overall effective noise a cell received from period to period. In an ideal world, we also wanted the estimator (noisy estimate) to be unbiased with respect to the actual estimate before we added noise. Based on these criteria, we rejected methods 5 through 7 because of the exceptions where the distribution of overall distortion was biased or had too much variation. These three methods had relatively less bias overall, but had relatively large variance in overall noise distortion and were similar to the method we currently use. We also rejected method 8 because too few establishments had fixed noise factors and the result was nearly identical to the current method. Methods 3, 9, and 10 produced similar results. We recommend method 3, which fixes the direction of cases appearing from last cycle to the current cycle because it most consistently reduced variance and bias in the distribution of overall percent noise of an estimate and is the simplest of methods 3, 9, and 10.

References

Cox, L. H., and Zayatz, L. (1995), "An agenda for Research in Statistical Disclosure Limitation," Journal of Official Statistics, Vol. 11, No. 2, pp.205-220.

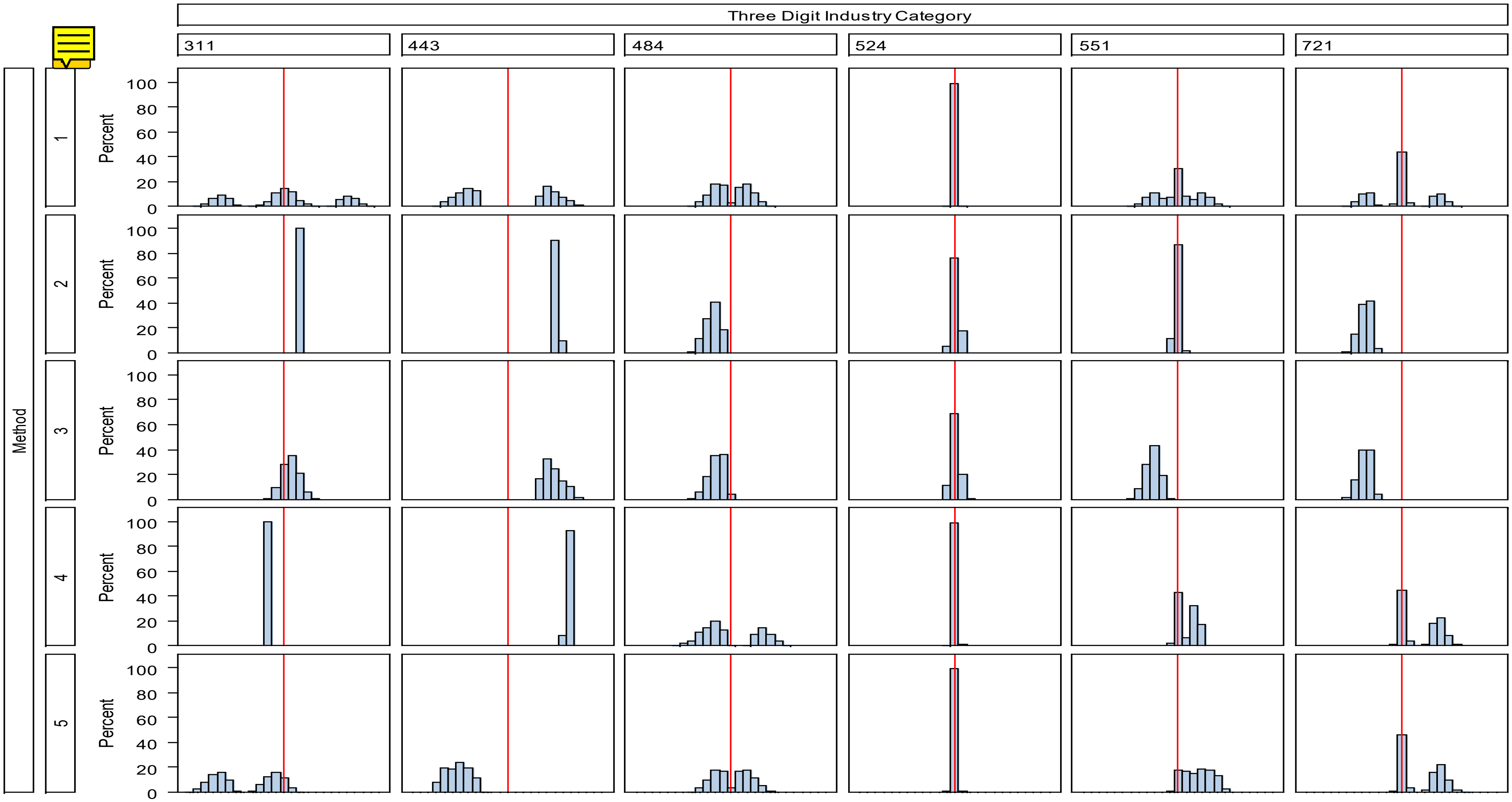
Evans, Timothy, Laura, Zayatz, and John Slanta (1998), "Using Noise for Disclosure Limitation of Establishment Tabular Data," Journal of Official Statistics, Vol.14, No.4, pp.537-551
URL http://www.jos.nu/Articles/abstract.asp?article=144537

Evans, B.T., Zayatz, L., and Slanta, J. (1996). "Using Noise for Disclosure Limitation of Establishment Tabular Data," Proceedings of the Annual Research Conference, U.S. Bureau of the Census, Washington, DC 20233, pp.65-86.

U.S. Census Bureau, "Economic Census of Island Areas. U.S. Census Bureau", Web. 7 Aug. 2013. URL http://www.census.gov/econ/islandareas

Appendix A

Distribution of Distortion as a Percent of Non-Noisy Total Sales for American Samoa



Method 1 – Reassigned noise direction and magnitude

Method 2 – Fixed noise direction and magnitude

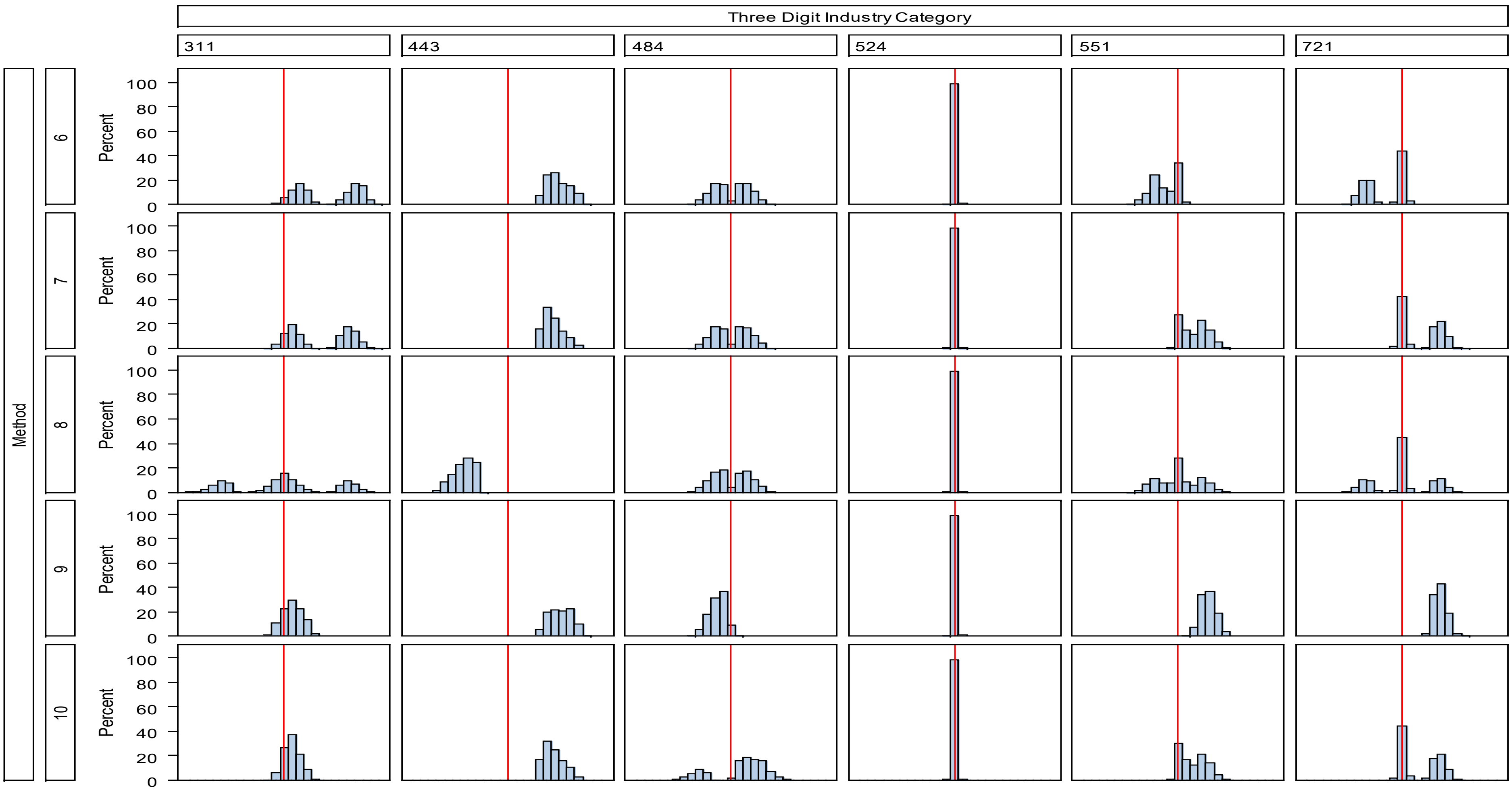
Method 3 – Fixed noise direction

Method 4- Fixed noise direction and magnitude for cases that are > 50% of an estimate or in top two contributors that together are > 75% of an estimate

Method 5 – Fixed noise direction and magnitude varies by 0.0125 for cases that are > 50% of an estimate

Appendix A

Distribution of Distortion as a Percent of Non-Noisy Total Sales for American Samoa



Method 6 – Fixed noise direction and magnitude varies by 0.005 for cases that are > 50% of an estimate

Method 7 – Fixed noise direction for companies that are > 50% of an estimate

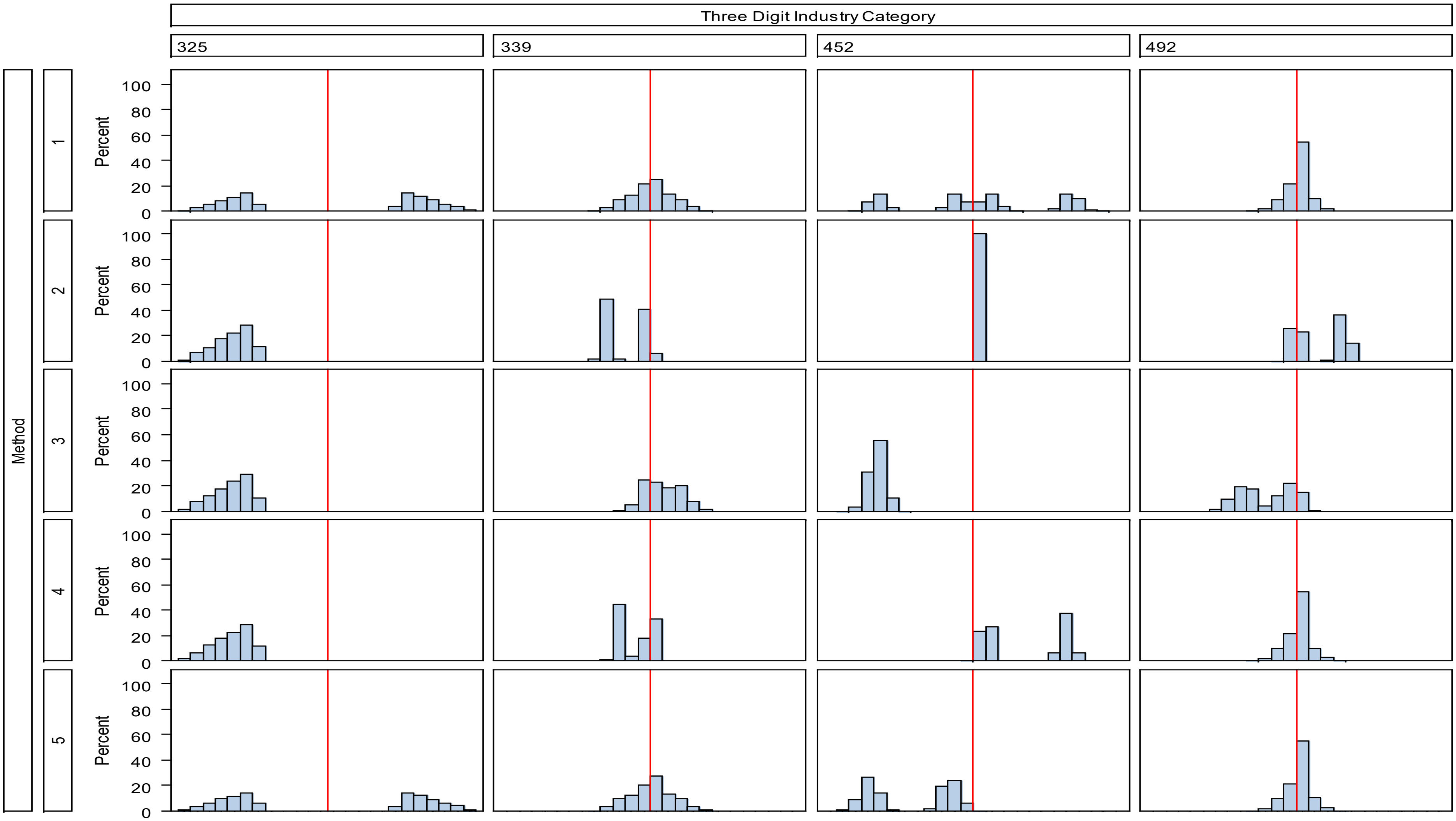
Method 8 – Fixed noise direction for companies that are > 75% of an estimate

Method 9 - Fixed noise direction and magnitude varies by 0.0125 for cases that are > 50% of an estimate or are a company with multiple establishments.

Method 10 – Fixed noise direction for cases that are > 50% of an estimate or in top two contributors that together are > 75% of an estimate

Appendix A

Distortion as a Percent of Non-Noisy Total Sales for The Virgin Islands



Method 1 – Reassigned noise direction and magnitude

Method 2 – Fixed noise direction and magnitude

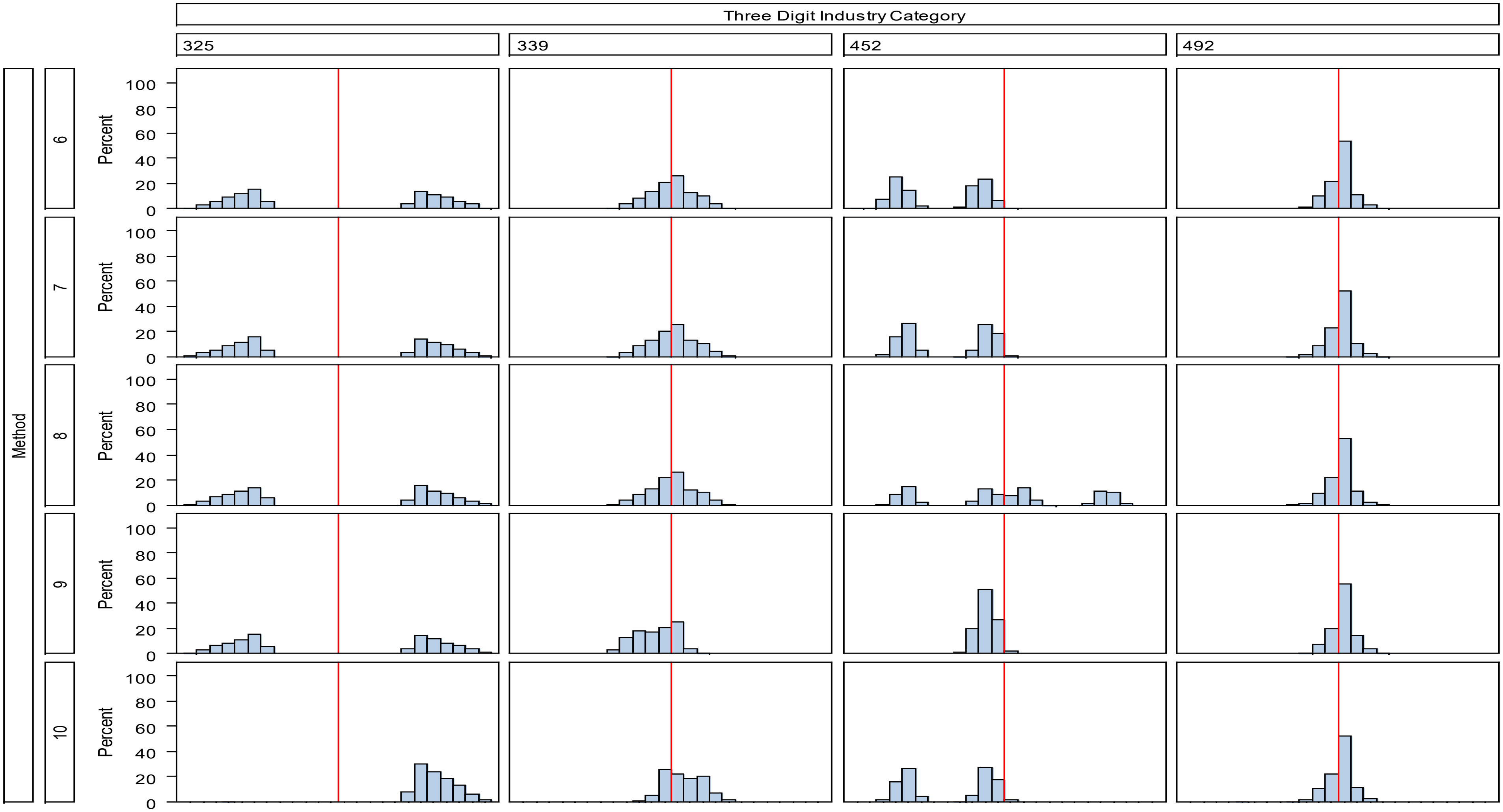
Method 3 – Fixed noise direction

Method 4 - Fixed noise direction and magnitude for cases that are > 50% of an estimate or in top two contributors that together are > 75% of an estimate

Method 5 – Fixed noise direction and magnitude varies by 0.0125 for cases that are > 50% of an estimate

Appendix A

Distortion as a Percent of Non-Noisy Total Sales for The Virgin Islands



Method 6 – Fixed noise direction and magnitude varies by 0.005 for cases that are > 50% of an estimate

Method 7 – Fixed noise direction for companies that are > 50% of an estimate

Method 8 – Fixed noise direction for companies that are > 75% of an estimate

Method 9 - Fixed noise direction and magnitude varies by 0.0125 for cases that are > 50% of an estimate or are a company with multiple establishments.

Method 10 – Fixed noise direction for cases that are > 50% of an estimate or in top two contributors that together are > 75% of an estimate