Appraising Auditor Error in Medicaid Audits

A. E. Rodriguez (Corresponding author)
Department of Economics & Business Analytics
Pompea College of Business, University of New Haven
West Haven, CT 06514, USA
Tel: 203-932-7372 E-mail: arodriguez@newhaven.edu

Kristin Kucsma
Sobel Tinari Economics Group
293 Eisenhower Parkway, Livingston, NJ 07039, USA
Tel: 973-992-1800 E-mail: kkucsma@SobelTinariEconomics.com

Received: August 30, 2023 Accepted: September 16, 2023 Published: September 18, 2023
doi: 10.5296/ijafr.v13i3.21281 URL: https://doi.org/10.5296/ijafr.v13i3.21281

Abstract

Medicaid conducts statistical audits to determine whether providers appropriately billed for Medicaid services. The basis for any claimed liability is an audit based on statistical extrapolation derived from a sample drawn from the totality of claims filed by the provider. Auditor error may be present in nearly every Medicaid audit, yet it is rarely explicitly taken into account in favor of a focus on survey error. It is demonstrated that for commonplace, plausible, audit conditions the presence of audit-error may result in coverage percentages that fall short of those derived via the classical Central Limit Theorem-based method. This data artifact may impeach the Medicaid auditor's examination of the claimant. The methodology here is not necessarily specific to Medicaid audits; rather, it may generalize to every statistical sampling-based health-care forensic audit.

Keywords: Misclassification, Auditor-error, Forensic audit, Medicaid fraud, Sampling-based audit

JEL Codes: C83, I13, I18

Sometimes a mistake is like wearing white after Labor Day, and sometimes a mistake is invading Russia in winter!

Alan Alda (Nothing but the Truth, 2008)

1. Introduction

Medicaid Fraud Control Unit (MFCU) auditors determine whether a Medicare provider claim complies with applicable state and federal regulations. Practically all Medicaid audits rely on conventional sampling procedures. (Note 1) When a paid claim fails to be in compliance, the recorded (paid) amount represents an error, a non-zero entry representing the difference between the amount claimed and the auditor's value; this non-zero difference could be positive or negative. It is negative when it is determined that the provider was underpaid for a particular service. When a claim does conform, there is no error, a zero entry. The end result of this review process is the presence in the sample of a large number of zero errors, a data artifact known as a zero-inflated population ("ZIP") (Kvanli & Schauer, 2018).

Zero-inflated populations are characterized by a frequency distribution with an observed spike at zero value and high positive skewness. The zero-inflated and highly skewed population weaken the theoretical statistical rationale underscoring the construction of the confidence interval of the population mean (Paneru, Padgett, & Chen, Estimation of Zero-Inflated Population Mean with Highly Skewed Nonzero Component: a Boostrapping Approach, 2022) (Satter & Zhao, 2018) (Paneru, Padgett, & Chen, Estimation of Zero-Inflated Population Mean: a Boostrapping Approach, 2018) (Kvanli, Shen, & Deng, Construction of Confidence Intervals for the Mean of a Population Containing Many Zero Values, 1998).

Claimants under examination are not privy to auditors' procedures and protocols, the level of experience and training of the audit staff nor, importantly, are claimants aware of the quality control measures designed to assure audit quality although it is well understood that auditors are susceptible to cognitive or procedure-based auditing-error (Anderson & Kraushaar, Measurement Error and Statistical Sampling in Auditing:The Potential Effects, 1986) (Bazerman, Loewenstein, & Moore, 2002) (Brody, DeZoort, Gupta, & Hood, 2022). A Medicaid auditor reviewing an individual claim cannot avoid the possibility of incurring one of two errors: the auditor can incorrectly flag as fraudulent or inappropriate a truly honest claim; this type of misclassification is known as a false-positive. A second type of audit-error occurs when the auditor incorrectly fails to flag a truly fraudulent or inappropriate claim. This latter type of misclassification error is known as a false-negative (Lohr, 2022) (Kvanli & Schauer, 2018) (Bonnett & Clute, 1990).

Auditor-error is a well-known and serious concern in audit sampling (Brody, DeZoort, Gupta, & Hood, 2022) (Caster, Massey, & Wright, 2000). Here from the *Handbook For Healthcare Auditors*:

> *For example, where there are many items to review, the auditor may, in the interest of practicality, tend to shorten the time devoted to reviewing each*

> *individual unit. Although a detail audit may be thought of as being better than using a sample because no sampling error is introduced into the examination, overall accuracy can be nevertheless severely impacted. This is because the audit scrutiny given to each unit is less than what is necessary. Therefore, the presence of audit error may become greater when compared to sampling error introduced by a review using a sample.* (Kvanli & Schauer, 2018)

The presence of error leads to considerable variation of the purported statistical accuracy of any overbilling estimate.

> *Further, a random sample taken from a population of tax errors will have all the problems found with a population of examined amounts and more. In Figure 1, the injection of numerous zero values (low error rate population) will make the population even more skewed than the examined amounts and cause the traditional methods to produce even less reliable lower confidence intervals* (Kvanli & Schauer, Is Your Agency Too Conservative? Deriving More Reliable Confidence Intervals, 2005).

In this paper, objectives are two: (i) to demonstrate the applicability of a misclassification decision-making algorithm in scrutinizing for audit-error in the Medicaid claims review process; and, (ii) determine whether correcting for audit error questions the soundness of the Medicaid audit processes. Left unaddressed, a showing of a statistical likelihood of auditor-error in an open proceeding may constitute prima facie evidence of the invalidity of the Medicaid audit. Methodology, applicable literature and results are provided in this paper – as follows: the next section discusses non-sampling error. Section three details the misclassification model and results. The section also illustrates the impact of zero-inflated populations on coverage proportions. The last section concludes.

This paper contributes to the literature and practice of audit sampling (Elder, Akresh, Higgs, & Liljegran, 2013), the furthering of error studies (Caster, Massey, & Wright, 2000) and forensic accounting – specifically related to healthcare audits and associated litigation (Hill, Hunter, Johnson, & Coustasse, 2014).

2. On the Possibility of Audit Error

A survey-based audit of health-insurance claims is fraught with two possible types of error: sampling error and auditor (non-sampling) error. Sampling error is the difference between a value inferred for a population and the actual value of the population that arise because of reliance on a sample. Assuming certain base conditions as to the drawing of the examination sample, the statistical procedure used by Medicaid audits takes into account sampling error. (Note 2) The MFCU has no such immunization to ward off cognitive or procedure-based auditing-error; more to the point, knowledge of any protocols salutary or otherwise is not available to claimants on any specific investigation. Claimants are not privy to the auditor's procedures and protocols, the level of experience and training of the audit staff nor, importantly, are they aware of the quality control measures designed to assure audit quality. The inability to scrutinize auditor practice and methodology prevents either the outright showing of auditor error or the documenting of indications that may raise a presumption of auditor error.

Non-sampling errors encompass a wide variety of possible auditor mistakes. Cognitive biases are of great – and more recent – concern. Non-sampling error can occur at any point in the audit process and include inter alia: unrepresentative (biased) sampling, inadequate frame, processing errors, misinterpretation of claims, time period biases (Lohr, 2022).

An auditor reviewing an individual claim cannot avoid the possibility of incurring one of two-errors: the auditor can incorrectly flag as fraudulent a truly honest claim; this type of misclassification is known as a false-positive. A second type of audit-error occurs when the auditor incorrectly fails to flag a truly fraudulent claim. This latter type of misclassification error is known as a false-negative.

A showing of auditor error compounds an already difficult statistical issue present in MFCU audits– a class imbalance: a disproportionately large number of zeroes. The low error rate, the class imbalance, may result in significant deterioration in the precision of the overbilling estimate.

> *"Specifically, a sample taken from this type of population usually consists of a very large number of zero values, plus a small number of non-zero values that follow some continuous distribution.* ***In this situation, the traditional confidence interval constructed for the population mean is known to be unreliable****."* (Kvanli, Shen, & Deng, Construction of Confidence Intervals for the Mean of a Population Containing Many Zero Values, 1998) (Note 3)

It is possible to scrutinize a sample for auditor misclassification error.

3. Methodology

3.1 A Mixture Model

Medicare audits are designed to establish the total amount of ineligible claims. If at fault on a particular claim, the amount of overpayment (or underpayment) is booked. For the most part, compliance violations are a small proportion of the total. The result is a sample consisting of a very large percentage of zero values because of the presumption that most claims comply with extant regulations.

A zero-inflated Medicaid audit sample can be viewed as a two-component mixture model. One consists entirely of a proportion of zero values; a second component of non-zero values characterized by some unknown data-generating-process (DGP). As follows:

$$H(y;\ \alpha,\mu,\sigma) = \alpha I_{(y\neq 1)} + (1\text{-}\alpha) f(y;\ \mu,\ \sigma) I_{(y=0)} \tag{1}$$

Where α is the proportion of non-zeros, μ is the mean, σ the nuisance parameter of non-zero components, and I is the indicator function with value 1 if true and 0 if false. The parameter of interest is the mean of the mixture distribution:

$$\theta = \alpha\mu \tag{2}$$

We assume the occurrence of the nonzero values follow a Bernoulli data generating process underpinned by audit-error. Audit error can occur at any state of the audit process. The term encompasses all other errors in the claims examination and may include any of the following: cognitive mistakes, inadequate sample, processing errors, misinterpretation of claim attributes, time-period bias. The non-zero component representing contested claims follows either a normal distribution with mean μ and standard deviation σ or an exponential function with rate, ρ.

For each of the nonzero value distributions generated, the coverage percentage for (1) the traditional method (classical method based on CLT), (2) a normal DGP, and an exponential DGP is estimated. The coverage percentage is the percentage of times (out of 1000) that the population mean is inside the confidence limits.

This mixture characterization of the audit survey DGP can be seen in Figure 1 below for the resulting mixture as well as the two constituent distributions: a Bernoulli distribution with a success ("1") probability of 0.10 and a Normal (5,1) distribution characterizing the overpayments. These distribution parameters are plausible values in Medicaid audits.

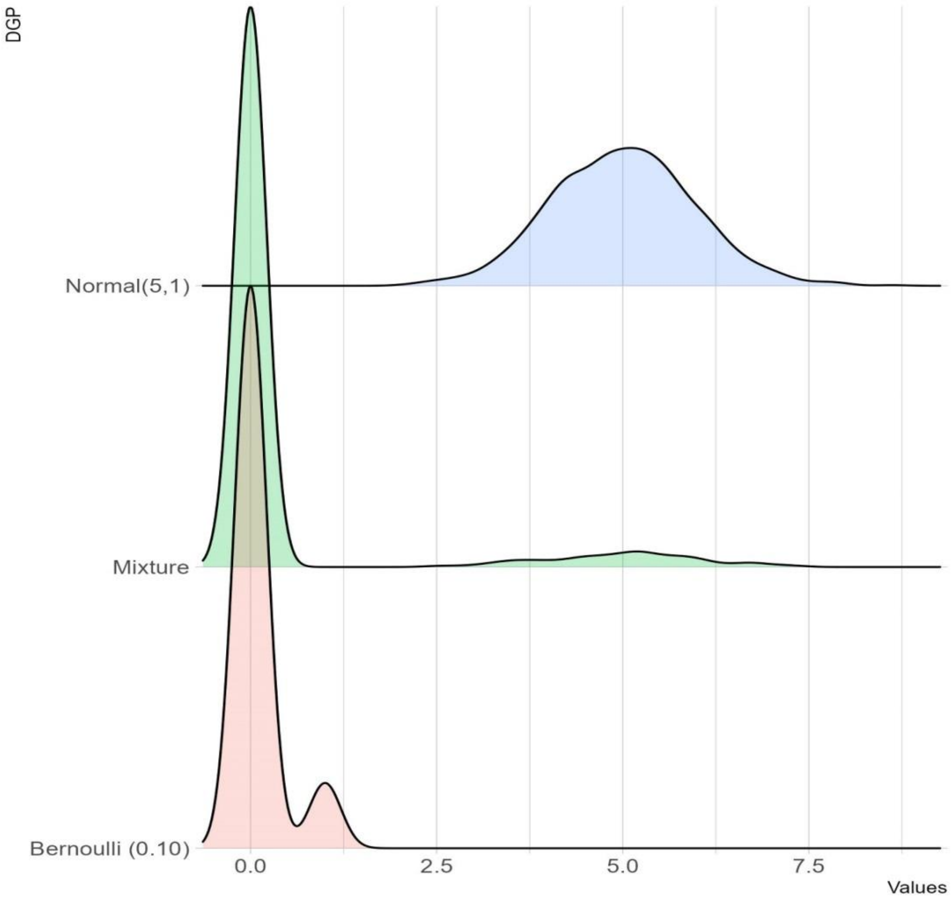

Figure 1

3.2 An Audit Misclassification Model

To identify the likelihood of audit-error we use the well-established binary choice misclassification model introduced by Hausman, Abrevaya and Scott-Morton (Hausman, Abrevaya, & Scott-Morton, 1998) (Liu & Zhang, 2017) (Feinstein, 1990). This model has found wide application in business applications including statistical-sample-based audit practice (Artis, Ayuso, & Guillen, 2002) (Caudill, Ayuso, & Guillen, 2005).

Let $\{(y_i, x_i), i = 1 \ldots n\}$ constitute the audit sample assembled by the MFCU, and the basis of the overpayment allegation. In the absence of auditor error, the recorded binary data y_i's are the true estimate of the audit inspection. The audit sample values differ from the true status $\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n\}$, which are unknown. The chance of misclassification is related to the true status, through the transition probability distribution function as follows:

$$Pr(y_i = 1 | \tilde{y}_i = 0) = \rho_0 \tag{3}$$

$$Pr(y_i = 0 | \tilde{y}_i = 0) = 1 - \rho_0 \tag{4}$$

$$Pr(y_i = 0 | \tilde{y}_i = 1) = \rho_1 \tag{5}$$

$$Pr(y_i = 1 | \tilde{y}_i = 1) = 1 - \rho_1 \tag{6}$$

Where ρ_0 and ρ_1 are false positive and false negative rates and represent the extent of misclassification (Liu & Zhang, 2017). In order to account for misclassification, we need to find the true distribution of the y_i's. For any y_i there are two possible outcomes. First, the underlying $y_i = 1$ and the audit result is correct. Second, the underlying $y_i = 0$ and the audit result is flagged, a misclassification. It follows that the probability π_i of an outcome $y_i = 1$ is the expected value:

$$
\begin{aligned}
\pi_i &= Pr(y_i = 1 | x_i) \\
&= Pr(y_i = 1 | \tilde{y}_i = 1 | x_i) Pr(y_i = 1 | x_i) + Pr(y_i = 1 | \tilde{y}_i = 0 | x_i) Pr(y_i = 0 | x_i) \\
&= (1-\rho_1) Pr(y_i = 1 | x_i) + \rho_0 [1 - Pr(y_i = 1 | x_i)] \\
&= \rho_0 + (1 - \rho_0 - \rho_1) [1 - Pr(\tilde{y}_i = 1 | x_i)] \\
&= \rho_0 + (1 - \rho_0 - \rho_1) F_i
\end{aligned}
$$

The regular logistic model can be extended to include false negative and false positive classification parameters as follows:

$y_i \sim \text{Bernoulli}(\pi_i)$

$\pi_i = \rho_0 + (1 - \rho_0 - \rho_1) F_i$

$F_i = 1/[1 + \exp(-\beta_0 + \beta_1 X)]$

Where Xi is the observed realization of the ith predictor. Thus, the formal model is:

$$\rho_0 + (1 - \rho_0 - \rho_1) F_i$$

$$f = y_i \sim \rho_0 + (1 - \rho_0 - \rho_1) exp(\alpha + \beta_0 + \beta_1 X)/(1 + exp(\alpha + \beta_0 + \beta_1 X))$$

and once the parameters have been estimated the resulting misclassification correction can be obtained as follows:

$$E(\tilde{y}_i) = (E(y_i) - \rho_0)/(1 - \rho_0 - \rho_1)$$

The model is amenable to solution via nonlinear least squares or via maximum likelihood. The R-package GSLNLS is used to estimate the misclassification parameters. (Note 4)

4. Results

To estimate the extent of misclassification of claims - we estimate a measure of both false positives and false negatives using nonlinear least squares. We adjust the sample to account for misclassification and proceed to gauge the coverage proportion.

4.1 Simulating an Audit Sample and Testing for Misclassification

Table 1 below provides the estimated misclassification for the randomly drawn data set. The proportion of non-zeroes, that is to say, the proportion of flagged claims, is in the 4th column, labeled Sample Proportion. The measure of false positives, in Column 2 indicates the proportion of claims that are likely to be erroneously classified as fraudulent. The third Column, “True Proportion” is the corrected proportion of zeros as a result the False Positive identification. Note the negative results, for instances where the misclassification rate exceeds the proportion of non-zero’s.

Table 1

	False Positive Misclassification Rate (FN Rate = 0)	**True Proportion**	**Sample Proportion**
1	0.05	0.0656	0.113
2	0.08	0.0379	0.113
3	0.10	0.0169	0.113
4	0.05	0.022	0.059
5	0.08	-0.016	0.059
6	0.10	-0.044	0.059

Table 2 contains similar results for the case where the hypothesized proportion of False Positives and False Negatives is identical. In turn, the “True Proportion” is the sample adjusted for the identified False Positives and False Negatives. Note the negative results, for instances where the misclassification rate exceeds the proportion of non-zero’s.

Table 2

	False Positive, False Negative Misclassification Rate (FN = FP)	True Proportion	Sample Proportion
1	0.05	0.081	0.113
2	0.08	0.047	0.113
3	0.10	0.026	0.113
4	0.05	0.017	0.059
5	0.08	-0.021	0.059
6	0.10	-0.048	0.059

4.2 Impact of False-Positives on Estimated Overbilling Amount

What is the impact of any identified false positive claims on the overbilling estimate? It is hypothesized that instances of misclassification negatively impact the accuracy and the precision of the MFCU's overbilling estimate.

4.3 Coverage Proportions: A Simulation

Because the error population consists of two different types of difference values we model the process underlying the appraisal of a provider's liability as a mixed distribution reflecting dual data-generating-processes: a Bernoulli distribution and a non-zero process drawn from (i) a normal and (ii) an exponential distribution (Kvanli, Shen, & Deng, Construction of Confidence Intervals for the Mean of a Population Containing Many Zero Values, 1998) (Chen, Chen, & Rao, 2003). The Bernoulli variables x_i has a probability p of being equal to one. The non-zero proportion represents the "misclassification adjusted" sample identified in the previous section. We examine outcomes reflecting commonplace Medicaid audit sample characteristics; specifically, we use the same audit sample mean ($37.38) and standard deviation ($11.48) used as a working example in Woodard (Woodard, 2015). For each of the nonzero value distributions generated, we compute the coverage percentage for the normal, and an exponential DGP. The coverage percentage is the percentage of times (out of 1000) that the population mean is inside the confidence limits. A reliable and accurate two-sided 95 percent confidence interval for the population mean should contain the actual mean 95 percent of the time.

We assume that the Bernoulli error rate p varies from 0.02 to 0.30 in increments of 0.05, the means and standard deviation of the error distribution was also specified; the sample size was varied from 100 to 300 in increments of 100. Due to space limitations, we give a summary finding for our simulations and display graphs representing the coverage percentages for the two nonzero value distributions.

The result in the figures below illustrates the coverage proportion of a misclassification adjusted error rate for the specified parameters and mixture distributions.

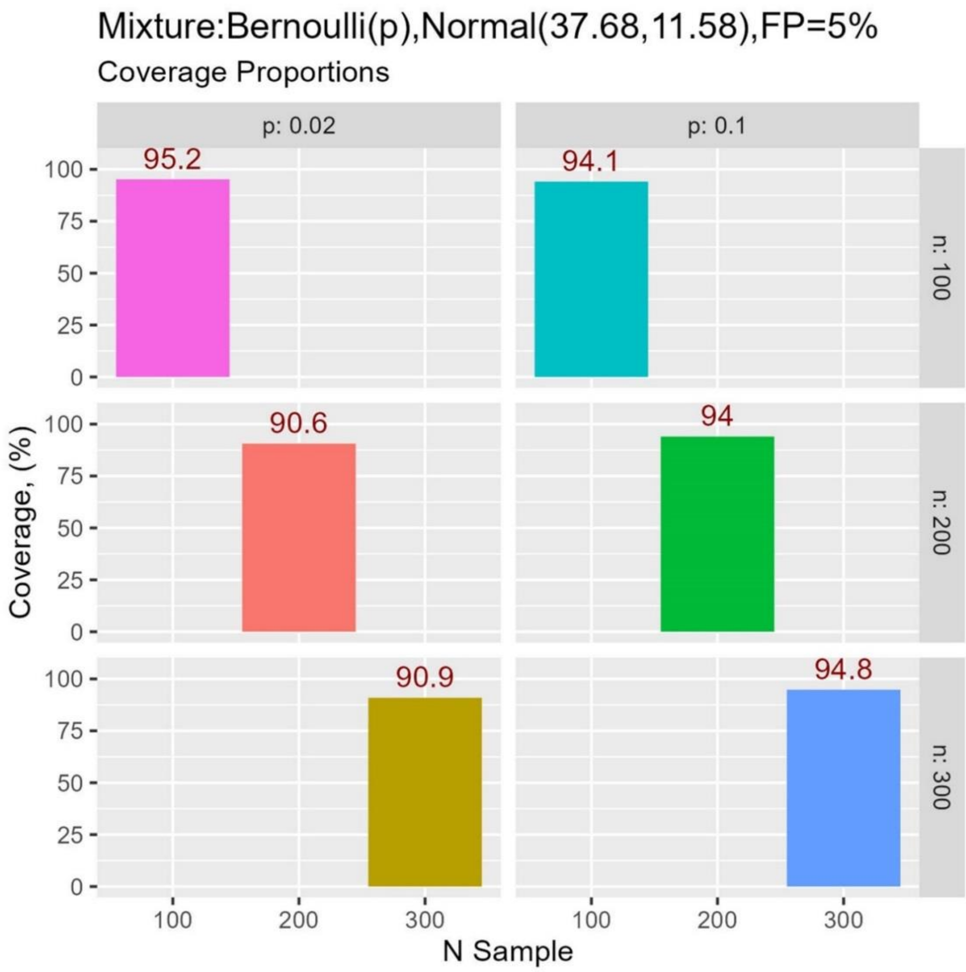

Figure 2

Not surprisingly, adjusting for the presence of misclassified claims reduces the non-zero proportion in the audit sample. In turn, this results in reduced coverage proportions. These results hold for simulations of two different mixture distributions representing the audit outcome results.

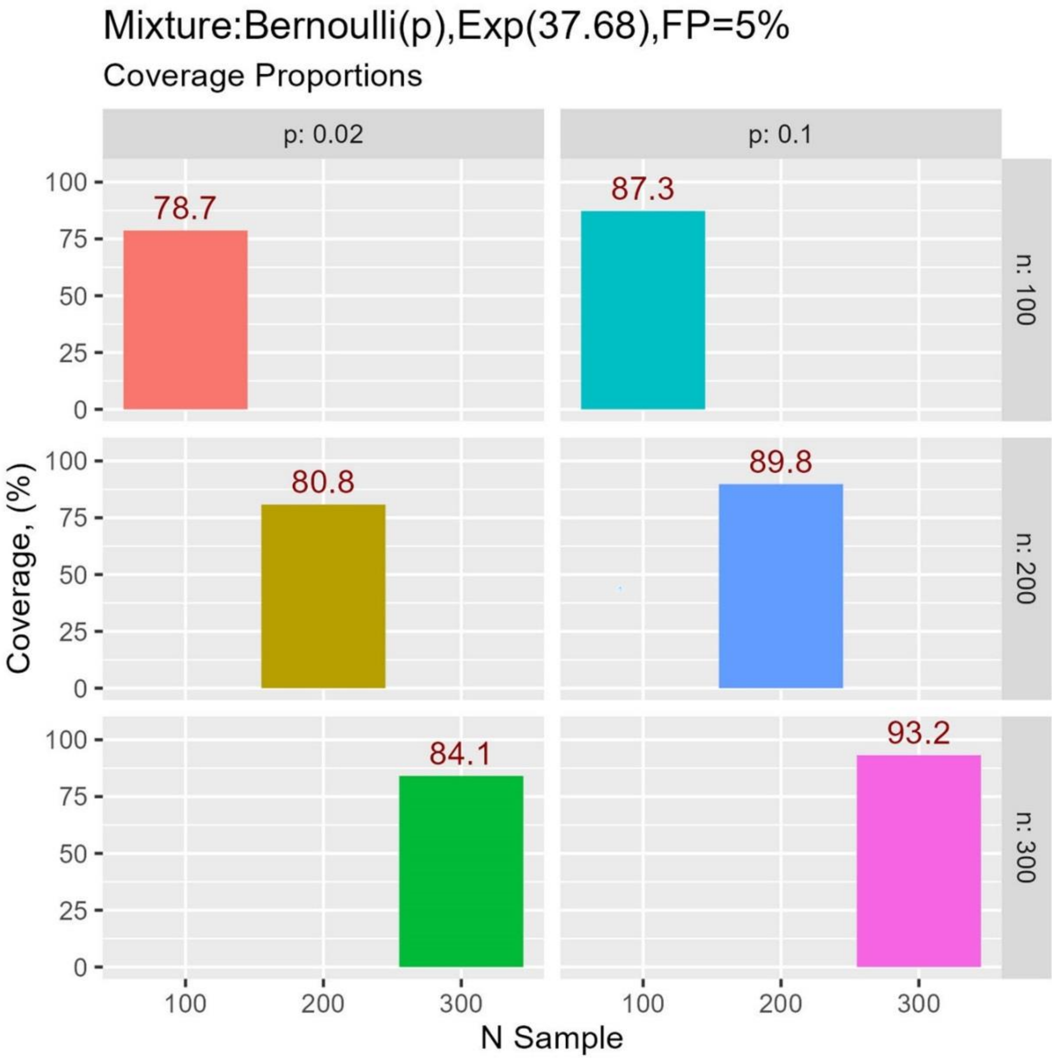

Figure 3

The coverage percentage of the exponential varies widely – and is heavily dependent on the various sample parameters. For adjusted error rates of approximately 2 percent and a small sample the coverage percentage is grossly distorted.

4.4 Discussion

It is well known that samples with a disproportionate number of zeroes will render the calculated confidence intervals unreliable. Auditors can overcome these limitations to some extent by altering their sampling methodology, by stratifying across any number of plausible claim and provider attributes and other possible fixes. These tools however, are all ex-ante fixes. A showing of the presence of audit misclassification in a Medicaid sample will compound the zero-inflation distortion worsening the purported accuracy.

5. Concluding Comments

The methodology underscoring random-sample-based surveys typically suggests that the sample overpayment estimate can serve as proxy for the universe estimate. This is known as accuracy (Lohr, 2022). Accuracy is how close a given set of measurements (observations or readings) are to their true value, while precision is how close the measurements are to each other. Classical statistics (and implicitly and explicitly the auditor) suggest that the amount proffered is an accurate measure of the true value; sampling error suggests that the best they can do (95 percent confidence) is that the true value will be within the proffered, associated confidence interval. Error in survey estimates is due to two types of errors: sampling error and non-sampling error. Measuring non-sampling error can be difficult.

The process does not take into account the likelihood of auditor error. The likelihood of such error is measurable, as shown here. If the auditor error is corrected, the confidence interval can vary as a result, and it can do so significantly. This variation may vitiate any overbilling claim because it impugns the accuracy and the precision of overpayment estimates. The results may cast doubt on the Medicaid audit in its entirety.

More generally, the ability to estimate the probabilities of false positives and false-negatives representing the likelihood of audit-error during the audit process, establishes a more robust burden on an auditor. The model proposed here provides a flexible approach to account for misclassification, applicable to a range of situations encountered in sampling-based forensic audits.

The approach here is necessarily limited. It does not examine prospective results in situations where the estimated misclassification parameters are constrained. Neither does the work here establish the tradeoffs between the solvers that could be used to estimate the misclassification parameters: for example, the limitations and advantages of non-linear solvers vs maximum likelihood – especially in forensic environment where interpretability and clarity are at a premium. Also, maximum likelihood is known to be sensitive to the choice of starting value in the optimization; this sensitivity was not explored here. These questions are left for further work.

Acknowledgements

A previous version of this paper was presented at the 2023 Eastern Economic Association Annual Meetings, sponsored by the National Association of Forensic Economists ("NAFE"). Rob Baumann, Victor Matheson of NAFE offered helpful comments. Colleagues Jim Mohs at the University of New Haven, Scott Lane at Auburn-Montgomery, and Joe Hearl of the Internal Revenue Service took time to review our work. We are grateful to all.

References

Anderson, J. C., & Kraushaar, J. M. (1986). Measurement Error and Statistical Sampling in Auditing:The Potential Effects. *The Acccounting Review, 61*(3), 379-399.

Anderson, J. C., & Kraushaar, J. M. (1993). Auditing Measurement Error and Statistical Sampling: the Dependently Occurring Case. *Journal of Accounting, Auditing & Finance, 8*(1), 53-75.

Artis, M., Ayuso, M., & Guillen, M. (2002). Detection of Automobile Insurance Fraud With Discrete Choice Models and Misclassified Claims. *The Journal of Risk and Insurance, 69*(3), 325-340. https://doi.org/10.1111/1539-6975.00022

Bazerman, M. H., Loewenstein, G., & Moore, D. A. (2002). Why Good Accountants Do Bad Audits. *Harvard Business Review*, 3-8.

Bonnett, D. G., & Clute, R. C. (1990). Audit Sampling with Nonsampling Errors of the First Type. *Contemporary Accounting Research*, *6*(2), 432-445. https://doi.org/10.1111/j.1911-3846.1990.tb00768.x

Brody, R. G., DeZoort, F. T., Gupta, G., & Hood, M. B. (2022). The Effects of Cognitive Bias on Fraud Examiner Judgments and Decisions. *Journal of Forensic Accounting Research*, *7*(1), 50-63. https://doi.org/10.2308/JFAR-2020-030

Caster, P., Massey, D., & Wright, A. (2000). Research on the Nature, Characteristics, and Causes of Accounting Errors: the Need for a Multi-method Approach. *Journal of Accounting Literature*, *19*, 60-92.

Chan, K. H. (1996). Estimating Accounting Errors in Audit Sampling: Extensions and Empirical Tests of a Decomposition Approach. *Journal of Accounting, Auditing & Finance*. https://doi.org/10.1177/0148558X9601100201

Chen, J., Chen, S.-Y., & Rao, J. N. (2003). Empirical Likelihood Confidence Intervals for the Mean of a Popuylation Containing Many Zero Values. *The Canadian Journal of Statistics*, *31*(1), 53-68. https://doi.org/10.2307/3315903

Efron, B. (1992). Bootstrap Methods: Another Look at the Jackknife. In S. Kotz, & N. L. Johnson (Eds.), *Breakthroughs in Statistics. Springer Series in Statistics* (Vol. 7). New York/NY: Springer. https://doi.org/10.1007/978-1-4612-4380-9_41

Elder, R. J., Akresh, A. D., Higgs, J. L., & Liljegran, J. (2013). Audit Sampling Research: a Synthesis and Implications for Future Research. *Auding: A Journal of Practice & Theory*, *32*(1), 99-129. https://doi.org/10.2308/ajpt-50394

Hausman , J. A., Abrevaya, J., & Scott-Morton, F. M. (1998). Misclassification of the Dependent Variable in a Discrete-Response Setting. *Journal of Econometrics*, *87*, 239-269. https://doi.org/10.1016/S0304-4076(98)00015-3

Hill, C., Hunter, A., Johnson, L., & Coustasse, A. (2014). Medicare Fraud in the United States: Can it Ever be Stopped?. *Health Care Manager*, *33*(3), 254-260. https://doi.org/10.1097/HCM.0000000000000019

Karim, K. E., & Siegel, P. H. (1998). A signal detection theory approach to analyzing the efficiency and effectiveness of auditing to detect management fraud. *Managerial Auditing Journal*, *13*(6), 367-375. https://doi.org/10.1108/02686909810222384

Kennedy, K., & Bishop, J. (2014). Random Sampling Isssues in a Federal Court Case, a case study. *Case Studies in Business, Industry and Government Statistics*, 111-114.

Kvanli, A. H., & Schauer, R. (2002). The Bootstrap: What the Government Auditor Should KNow. *The Journal of Government Financial Management*, *51*(3), 24-33.

Kvanli, A. H., & Schauer, R. (2005). Is Your Agency Too Conservative? Deriving More Reliable Confidence Intervals. *The Journal of Government Financial Management*, *54*(2), 30-37.

Kvanli, A. H., Shen, Y. K., & Deng, L. Y. (1998). Construction of Confidence Intervals for the Mean of a Population Containing Many Zero Values. *Journal of Business and Economic Statistics*, *16*(3), 362-368. https://doi.org/10.1080/07350015.1998.10524776

Kvanli, A., & Schauer, R. (2018). *Handbook for HealthCare Auditors.* Alan Kvanli.

Liu, H., & Zhang, Z. (2017). Logistic Regression with Misclassification in Binary Outcome Variables: a Method and Software. *Behaviormetrika*, *44*(2), 447-476. https://doi.org/10.1007/s41237-017-0031-y

Lohr, S. L. (2022). *Sampling: Design and Analysis* (3rd ed.). Boca Raton, Fla: CRC Press.

Office of Inspector General. (2018). *Statistical Sampling: a Toolkit for MFCUs.* US Department of Health & Services.

Paneru, K., Padgett, R. N., & Chen, H. (2018). Estimation of Zero-Inflated Population Mean: a Boostrapping Approach. *Journal of Modern Applied Statistical Methods*, *17*(1), 2-14. https://doi.org/10.22237/jmasm/1525133460

Paneru, K., Padgett, R. N., & Chen, H. (2022, September). Estimation of Zero-Inflated Population Mean with Highly Skewed Nonzero Component: a Boostrapping Approach. *Statistics, Optimization and Information Computing*, *10*, 1044-1055. https://doi.org/10.19139/soic-2310-5070-1491

Ramsay, R. J., & Tubbs, R. M. (2005). Analysis of Diagnostic Tasks in Accounting Research Using Signal Detection Theory. *Behavioral Research in Accounting*, *17*(1), 149-173. https://doi.org/10.2308/bria.2005.17.1.149

Rotenstein, A., Epstein, S., & Wilamowsky, Y. (2012, January/February). Auditing the Auditors In Medicare and Medicaid. *Journal of Business Case Studies*, 67-71. https://doi.org/10.19030/jbcs.v8i1.6739

Satter, F., & Zhao, Y. (2018). Nonparametric Interval Estimation for the Mean of a Zero-Inflated Population. *Communications in Statistics - Simulation and Computation*, 1-10. https://doi.org/10.1080/03610918.2018.1510527

Notes

Note 1. Medicaid is an American public insurance program that pays for health care claims for services rendered from health care providers and health plans. It is a federal program administered by the states. The Medicaid Fraud Control Units (MFCUs) is the enforcement unit typically housed within a State Attorney General's office.

Note 2. The MFCU audit process relies on RAT-STATS – an open source statistical package (Woodard, 2015). RAT-STATS has conventional CLT-based statistical procedures baked-in.

Note 3. Letters in bold are ours. See also (Anderson & Kraushaar, Auditing Measurement Error and Statistical Sampling: the Dependently Occurring Case, 1993) (Anderson & Kraushaar, Measurement Error and Statistical Sampling in Auditing:The Potential Effects, 1986), and (Chan, 1996).

Note 4. The error parameters can also can also be estimated via maximum likelihood. The HASM misclassification algorithm is available in the R package McSpatial. Estimates of the errors using this alternative were consistent with those obtained via a nonlinear solver. Results are not reported here.