Tools for Ferreting-out Fraud: a Book Review of Mark Nigrini's Forensic Analytics

Rodriguez, A.E.

University of New Haven, Pompea College of Business

1 August 2024

A.E. Rodriguez*

University of New Haven

Tools for Ferreting-out Fraud

MARK J. NIGRINI, *Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations*, John Wiley & Sons (Hoboken, NJ: Wiley, 2020, ISBN: 978-0-470-89046-2, 463 pages, $95.00).

A book review of the first edition of Nigrini's book cited a 2010 report on occupational fraud and abuse to highlight that "fraud cost U.S. companies billions in 2009" (Li & Byrnes, 2012).[1] Almost 15 years later after that eye-popping alert we read recently in the *Wall Street Journal*: "Insurer's pocketed $50 billion from Medicare for diseases no doctor treated" (Weaver, McGinty, Mathews, & Maremont, 2024).

To state the obvious, fraud remains ubiquitous and costly. Add to this perennial malady the array of internet-age fraud innovations such as click-fraud, phishing, and ransomware attacks and the stage is set for an epidemic of nervous breakdowns.

The tools to ferret out these nuisances have also been evolving, improving, and importantly, becoming more well known, their presence a measure of relief in this endless whack-a-mole. Nigrini's tome is essential in this combat, indispensable for anticipating and identifying fraud and other ignominious

*Professor, Department of Economics & Business Analytics. Email:arodriguez@newhaven.edu. I am obliged to Scott Lane and Mary Miller for providing helpful comments.

[1] Citing the Association of Certified Fraud Examiners (ACFE), 2010 Report to the Nations on Occupational Fraud and Abuse. The 2024 edition of this study estimates total worldwide losses of approximately $3.1 Billion (Association of Certified Fraud Examiners, 2024).

threats. Nigrini's book is aimed at the forensic accounting world. Fraud actions, however, involve economic, insurance and financial experts, who are likely to find the Nigrini tool-kit highly useful.

Here's a conundrum. Despite eye-popping advances in both supervised, and unsupervised machine learning techniques available for the task of ferreting out fraudulent practices and outcomes, Nigrini overlooks them (Debener, Heinke, & Kriebel, 2023), (Johnson & Khoshgoftaar, 2019), (Baesens, Van Vlasselaer, & Verbeke, 2015).[2]. Rather, his focus emphasizes "meat and potatoes" types of tools. So why would the leading tome in the field emphasize more "conventional" tools? I am going to argue here that Nigrini's tool-kit makes sense, given the practical constraints facing practicing forensic analysts which limit the complexity and breath of the tools available.

What practical constraints? Interrogating machine learning tools appears to have been institutionalized by the General Data Protection Regulation ("GDPR") in the European Union.[3] This legislation codifies people's distrust of opaque, black-box algorithms and the associated "trust-me, I am the expert" culture surrounding it. This skepticism trend will probably be hastened along by data protection legislation in the US, a drumbeat that may have started with the uproar surrounding the *Loomis* decision and merely heighted with the recent AI proliferation anxiety (Bloomberg, 2018). In Loomis, plaintiff was not allowed to examine the inner workings of a proprietary re-offender risk-assessment tool known as COMPAS (Yong, 2018).

What does this mean? Among other things, these social, regulatory, processes encircle and limit the algorithmic toolkit range and enhance the appeal of heuristic, easier-to-

---

[2] To get a feel for the breadth and complexity of the tools available see, for instance, the algorithms underlying the solutions posted in the leaderboards of the following Kaggle fraud competitions: see here for fraudulent transactions, and here for click fraud.

[3] General Data Protection Regulation (GDPR) is a 2008 EU legislation aimed at data protection and privacy. It reaches beyond the European Union and its sphere of influence; it affects firms and organizations outside the EU who do business with EU entities and individuals.

understand algorithms and most-likely place a premium on human-in-the-loop analysis and processes.

A premium on parsimony and transparency has long been evident to forensic analysts. It's not only the trier-of-fact that needs to understand any model underlying economic or accounting analysis in proffered legal proceedings, but counsel as well - on both sides. Novel or newly introduced methodologies or algorithms are Daubert catnip.[4]

The "medium-tech" approach in Nigrini makes sense for it emphasizes two things: interpretability and clarity. These features speak to the complexity of the underlying model or technique and the interpretation of the results. The two objectives are intertwined. Forensic work is conducted with the trier-of-fact in mind. A methodology that is parsimonious, sound and easy to explain will be more appealing than one with unappealing sophistication.

At their essence the underlying methods proffered by Nigrini to scrutinize accounting and financial practices is to set forth a recuring pattern (of normalcy) and then proceed to identify any aberrations or outliers in the process. This is of course, a practical, common-sensical application of the paradigm that underwrites science: null hypothesis significance testing (NHST). In NHST, an assumption of no change is to be met by any contrary evidence. However, the statistical testing in Nigrini's procedures is not for the purpose of inference of statistics representing manipulation or deception towards a broader population. Any seeming instance of impropriety drawn from initial testing raises a presumption against the particular unit or account being examined that may subsequently be more closely scrutinized or audited.

Nigrini is probably best known for his work using Digit Preference Analysis as a fraud detection tool. We would use Digit Preference to examine whether the positions of a number present in any particular data point occurs with a greater frequency that is expected. Expected in what way? Benford's

[4] Daubert is the legal standard establishing the parameters of what is admissible in expert testimony. It requires trial judges to scrutinize the submissions of expert witnesses in legal proceedings.

Law is the foundation underscoring Digit Preference analysis. Benford's law describes a pattern in many naturally-occurring numbers. According to Benford's law, each possible leading digit *d* in a naturally occurring set of numbers occurs with a probability *p(digits)*, where

$$probability(digits) = log10(1 + 1/digit)$$

for digits from 1 to 9. Deviations from the expected pattern in operations data triggers red flags. Benford's Law, fully grown into the "Nigrini Cycle" takes pride of place in the book. The full-fledged Nigrini Cycle comprises a series of eight statistical tests related to Benford's law: a data profile, periodic graph, histogram, first-order test, second-order test, summation test, number duplication, and last-two digits test.

Other statistical tests: descriptive statistics to compare current and prior-period data; correlations to identify entities containing data that deviates from an expected value; and time-series analysis to identify where current-period data diverges from past values.

There are at least three broad areas left off the table worth probing. Nigrini does not tread on clustering methods nor on any other "glitzy" machine-learning tool. As we discussed above, we find this a strength. But what is the tradeoff? Clustering methods are a hugely useful tool for any number of tasks from outlier detection, segmentation, and anomaly detection. In fact, clustering methods are at the core of modern day automatic, continuous, full-population audit scrutiny methods. It is important to know the tradeoff between accuracy and practicality when opting for non-ML tools when favoring Nigrini's over the most recent advances in forensic analysis tools.

Another area that could benefit from some examination entails the relatively recent innovation of full-population testing and where, and how, fraud will rear its ugly head in this context. Importantly, if any changes or supplementary tools are required of forensic analysts.

And last, given the rapid evolution of fraud methods, it is important to continue the search for similar, medium-tech, easily deployed, quantitative methods. For example, the bias ratio for detecting fraudulent investment activity, number bunching, and the recent uses of google searches in detecting revenue management appear promising (King & van Vuuern, 2016) (Simonsohn, 2019) (Chiu, Teoh, Zhang, & Huang, 2023).

The package of resources accompanying the book is an instructor's dream. The book contains screenshots showing how to run the proffered tests in Access, Excel, SAS, and in R. Data sets; case files; code scripts. There are copious references, case studies, data sets, tutorials, and instructive videos. This book is easily the main text in a forensic accounting course, and an invaluable reference for any accounting, insurance, financial or economic forensic expert.

The fraud prevention and management landscape evolves rapidly. It is indispensable for firms and institutions to adopt advanced risk assessment tools, resort to real-time monitoring, turn to full-population auditing and other innovations and emerging technologies. But despite the appeal of readily available, impressive, high-powered advances in AI and machine learning predictive analytics Nigrini's blend of accounting, finance and empirical methods firmly grounded in the accessible and understandable remain essential in the war against fraud.

References

Association of Certified Fraud Examiners. (2024). *Occupational Fraud 2024: A Report to the Nations*. Austin: Association of Certified Fraud Examiners.

Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). *Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques: A Guide to Data Science for Fraud Detection*. Hoboken: John Wiley & Sons.

Bloomberg, J. (2018, 09 16). Don't Trust Artificial Intelligence? Time To Open The AI 'Black Box'. *Forbes*.

Chiu, P.-C., Teoh, S. H., Zhang, Y., & Huang, X. (2023, August). Using Google Searches of Firm Products to Detect Revenue Management. *Accounting, Organizations and Society, 109*. doi:doi.org/10.1016/j.aos.2023.101457

Debener, J., Heinke, V., & Kriebel, J. (2023). Detecting insurance Fraud Using Supervised and Unsupervised Machine Learning. *Journal of Risk and Insurance, 90*, 743-768. doi:10.1111/jori.12427

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Medicare Fraud Detection Using Neural Networks. *Journal of Big Data, 6*, 1-35. doi:doi.org/10.1186/s40537-019-0225-0

King, J., & van Vuuern, G. W. (2016). Flagging Potential Fraudulent Investment Activity. *Journal of Financial Crime, 23*(4), 882-901. doi:doi.org/10.1108/JFC-09-2015-0051

Li, P., & Byrnes, P. E. (2012). Book Reviews. *Journal of Information Systems, 46*(1), 207-212. doi:10.2308/isys-10247

Simonsohn, U. (2019, May 25). *Data Colada*. Retrieved from Number-Bunching: A New Tool for Forensic Data Analysis: https://datacolada.org/77

Weaver, C., McGinty, T., Mathews, A. W., & Maremont, M. (2024, July 8). Insurers Pocketed $50 Billion From Medicare for Diseases No Doctor Treated. *The Wall Street Journal*.

Yong, E. (2018, January 17). A Popular Algorith is No Better at Predicting Crimes Than Random People. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/