EMBRACING FRAGILITY IN OUR DATA: A CAUTIONARY EXAMPLE FROM RESEARCH ON THE FCPA AND VOLUNTARY DISCLOSURE

Peter Leasure*

I. INTRODUCTION

Enacted by the United States in 1977 as part of the Security Exchange Act of 1934,¹ the Foreign Corrupt Practices Act (FCPA) was one of the first of its kind in the area of foreign official anti-bribery legislation.² Despite increasing levels of enforcement³ and a litany of scholarly critiques, statistical data and analysis on the FCPA is hard to find. In fact, the Organization for Economic Cooperation and Development (OECD) has noted this gap. In a Phase 2 Report, the OECD stated:

[T]here are no clear, documented, formal processes between agencies to underpin the vital exchange of information and reporting of suspected violations, and a corresponding absence of statistics. This results in a lack of transparency and of data, which, if captured, could serve useful analytical purposes in reviewing the workings of the FCPA.⁴

To fill this gap, some authors have begun to address various FCPA

---

¹ Peter Leasure is a Ph.D. candidate in the Department of Criminology and Criminal Justice at the University of South Carolina. His research focuses on corporate compliance and collateral consequences of conviction. His work has appeared in journals such as the Journal of Financial Crime, Journal of Experimental Criminology, Journal of Money Laundering Control, and Yale Law and Policy Review Inter Alia.


Embracing Fragility in Our Data

empirical questions with statistical analysis by building their own datasets. Unfortunately, some of these studies suffer from severe methodological flaws.

The current paper aims to critique one of these studies which explored whether companies with FCPA transgressions received a benefit to voluntary disclosure. The goal of this paper is to show that those involved in FCPA research, especially those publishing legal journals, could greatly benefit from a better understanding of statistics. Though this paper will demonstrate that the previous study possesses several flaws, this author does not argue for the abandonment of statistical research in FCPA research. On the contrary, the present study shows that a better understanding of statistics is indeed vital for FCPA research to progress.

II. THE FCPA AND THE BENEFITS OF VOLUNTARY DISCLOSURE

There were promises of benefits for voluntary disclosure of FCPA misconduct from the FCPA’s inception. These early promises have been commonly repeated and are most prominent in the Organizational Sentencing Guidelines created by the Sentencing Reform Act in 1984. The Organizational Sentencing Guidelines offer mitigation for companies that voluntarily disclose conduct (Organizational Sentencing Guidelines). Recent promises of mitigation for voluntary disclosure have also come from the Department of Justice ("DOJ"), Acting Deputy Attorney General Gary

---


6 See generally Hinche... note 4.


9 Memorandum from Larry Thompson, Deputy Attorney Gen., to Heads of Dep't
Embracing Fragility in our Data

G. Grindler, and Assistant Attorney General of the Criminal Division, Lanny Breuer. Despite these repeated assurances of mitigating benefits for voluntary disclosure, many question whether a tangible benefit for voluntarily disclosing actually exists.

In the only empirical study looking at the benefits to voluntarily disclosing FCPA misconduct, the author examined 40 cases from 2002 through 2009, separated cases into voluntary and involuntary disclosure groups, and then utilized bivariate linear regression in an effort to compare the two groups’ coefficients. The dependent variable was the total fines and forfeitures received by the company and the independent variable was the amount of bribe paid. The author hypothesized that all things being equal, a company that voluntarily disclosed should receive a lower fine than those that did not voluntarily disclose. If the coefficient for the voluntary group was larger than or equal to the involuntary group, then the author would determine that there was no benefit to voluntary disclosure. If the coefficient for the involuntary group was higher than the voluntary group, then the author would determine that there was a benefit to voluntary disclosure.

The study’s results were as follows: [The Voluntary group’s (N=15) coefficient was 4.259917, (p < .05). The Involuntary group’s (N=19) coefficient was 1.649584, (p < .05).] "The coefficients listed above suggest that voluntarily disclosing companies tend to face stiffer fines and forfeitures than those that do not voluntarily disclose. The linear regression model above shows that the difference between these two groups is nearly 2.5:1." Interpreting the coefficients can make things a bit clearer. According to the previous study, a one-dollar increase in bribe resulted in a $4.26 increase in the total fine for the voluntary group. A one-dollar increase in bribe for the involuntary group resulted in a $1.65 increase in total fine. Therefore, the previous study determined that companies who voluntarily


Gary Grindler, Acting Deputy Attorney General, Address at the 2010 Compliance Week Conference (May 25, 2010).


See Hinchey, supra note 4, at 398 (citing Lucinda Low et al., the Uncertain Calculus of FCPA Voluntary Disclosures 24, 16th National American Conference Institute on Foreign Corrupt Practices Act (Nov. 9, 2006); Jessica Tillipman, Foreign Corrupt Practices Act Fundamentals, BRIEFING PAPERS, 15 (2008).

Hinchey, supra note 4, at 399–404.

Hinchey, supra note 4, at 404 n.59.
disclose actually face higher penalties. In fact, voluntarily disclosing companies received approximately two and a half times higher fines than those that did not disclose.

III. CURRENT STUDY

The present paper revisits the Hinchey study in an attempt to determine whether the methods and statistics employed in that analysis were appropriate. To achieve this goal, methodological approaches of the Hinchey study will first be reviewed. After that review, findings of the Hinchey study will be replicated with the inclusion of confidence intervals.

A. Methodological Review of Hinchey

The most serious methodological error from the Hinchey study comes in the form of sample selection bias. When you attempt to collect data for an analysis, you must decide whether you are going to study the entire population (census approach) or draw a sample. If you use a sample approach, you must ensure that the sample is representative of the population; otherwise, your results may not generalize to the entire group you are studying. The most effective method to avoid sample selection bias is to utilize random sampling. In random sampling, representativeness is ensured because each case has an equal chance of being selected for the analysis. In the Hinchey study, the author selected FCPA cases that arose from 2002 to 2009. Such a process does not ensure representativeness, as perhaps FCPA cases in that era were unique. Therefore, the methodological error is not that the author chose that sample, it is that the author generalized the results of that sample to the entire population of FCPA cases.

15 An example of the census approach in this area would be a study that looked at every FCPA case ever prosecuted.
16 An example of a sample would be choosing 30 FCPA cases from a population of FCPA cases.
17 Further, many analyses are conducted to predict future outcomes. With the present research question, one would hope that any findings on whether there is a benefit to voluntary disclosure would be used to guide companies on the decision to disclose future transgressions. Sample selection bias could therefore also negatively impact these companies seeking to use the Hinchey findings as definitive guidance.
18 WILLIAM R. SHADISH, THOMAS D. COOK, & DONALD T. CAMPBELL, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR GENERALIZED CAUSAL INFERENCE 56 (2d ed. 2002). To produce a random sample in the previous study, the author would have needed to collect every FCPA case ever prosecuted, number the cases and then randomly choose a sample of those cases.
19 Id.
Embracing Fragility in our Data

Second, Hinchey utilized only bivariate analysis. One of the main benefits of regression is that it can estimate the effect of a key independent variable on an outcome while also accounting for other potentially influential variables (control variables). Some potential control variables left out of the Hinchey study include industry type, previous transgressions, foreign official nationality, and whether the company received a compliance monitor. Failing to control for these and other variables creates a possibility of biased results.

A third methodological concern deals again with analytic strategy. In the Hinchey study, the author used linear regression to determine whether there is a benefit to voluntary disclosure. However, before one can conduct linear regression analysis, several assumptions need to be addressed because violations can lead to biased results. Using the Stata software, this current critique detected several possible violations (normality of the dependent and independent variables, linearity, no autocorrelation between residuals, normally distributed residuals, and lack of influential observations) which were not expressly dealt with in the Hinchey paper. While violations do not always mean that bias will result, an analysis should note any diagnostic or corrective actions.

A fourth methodological concern comes from the decision to separate the voluntary and involuntary groups. A better approach would have been to create another independent variable that looked at whether the company voluntarily disclosed. However, when two or more independent variables are involved in a research design, there is more to consider than simply the effect of each of the independent variables (bribe and the new variable of whether the company voluntarily disclosed) on the dependent variable (total fine). When the effect of one independent variable may depend on the level of the other independent variable, we term that as an interaction. Failing to identify an interaction can lead to incorrect results.

Finally, one should always exercise caution with results from studies utilizing small sample sizes. Using such small sample sizes means that results can be sensitive. For example, with the sample sizes used in the Hinchey paper, including another one or two FCPA cases (not even influential observations) in the regression could have significantly altered the

---

21 StataCorp LP, Stata Statistical Software (Release 13), http://www.stata.com/products/ (last visited July 11, 2017) (at the time of this writing the most recent version is Release 15).
22 An overlay of these diagnostic tests and the data used is available from the author upon request.
conclusions drawn in that study.

**B. Replication of Hinchey Findings with Confidence Intervals**

When a regression is calculated, it may only produce one value for the coefficient(s). This can be referred to as a point estimate as it is only a single value. The Hinchey study produced only point estimates. The problem is that researchers can also include a confidence interval which produces a range (a lower value and an upper value) for the coefficients. The confidence interval is important (and is now generally standard practice;\(^{24}\)) because it attempts to account for uncertainty in a research design. Further, reporting confidence intervals is especially important when using a sample and also when one wishes to predict future outcomes. What is vital to understand about confidence intervals is that the actual value can lie anywhere within that range. The most common interval used in statistics is the 95% confidence interval, meaning that there is only a 5% chance that values could fall outside of the stated range.

Therefore, to definitively answer whether voluntarily disclosing companies receive a benefit, we must get results where the 95% confidence intervals for the voluntary and involuntary disclosure groups do not overlap. The 95% confidence interval shows that the voluntary group’s coefficient can range from 2.05 and 6.52.\(^{25}\) The 95% confidence interval shows that the involuntary group’s coefficient can range from .87 and 2.27. Therefore, the replication shows that the coefficients overlap from 2.05 to 2.27. While this amount of overlap is slight, the fact remains that drawing a definitive conclusion from these results is simply not possible.

**IV. DISCUSSION**

What do we know for sure about the benefits of voluntary disclosure for companies who violate the FCPA from the Hinchey analysis? The answer is very little. And this author is perfectly comfortable with that answer. After conducting the analysis presented in this paper, even the previous study’s title, “Punishing the Penitent: Disproportionate Fines in Recent FCPA Enforcement Actions and Suggested Improvements,” is a misstatement as this analysis shows that it is impossible to definitively conclude that the fines

---


\(^{25}\) The regression output is available upon request from the author.
are disproportionate for voluntarily disclosing companies.\textsuperscript{26}

I believe the previous study fell victim to an unfortunately too common error in policy analysis in reporting strong conclusions from weak data.\textsuperscript{27} Too often certain conclusions are drawn from weak data and strong assumptions. Why do researchers sacrifice credibility in favor of strong conclusions and certain answers? Charles F. Manski answers as follows:

The scientific community rewards those who produce strong novel findings. The public, impatient for solutions to its pressing concerns, rewards those who offer simple analyses leading to unequivocal policy recommendations. These incentives make it tempting for researchers to maintain assumptions far stronger than they can persuasively defend, in order to draw strong conclusions.\textsuperscript{28}

\textbf{V. CONCLUSION}

Hinchey took a big step in seeking to apply statistical analysis to this area. This author cannot stress enough that this current paper should not be used as evidence against using statistical analysis in FCPA research. On the contrary, the present paper shows the absolute necessity for those of us in this field to seek a better understanding of statistics. Only with that better understanding can we as a field both recognize the limits of statistical analysis and also utilize its beneficial power. We as a field need to accept fragile answers. If responsible and appropriate methods yield a fragile answer, then that answer should be embraced.

\textsuperscript{26} In fact, a recent study that addressed the limitations in the Hinchey analysis found that voluntarily disclosing companies actually do receive a benefit in their total fine calculation. See Peter Leasure, \textit{The Foreign Corrupt Practices Act and the Benefits of Voluntary Disclosure}, 23 JOURNAL OF FINANCIAL CRIME, 4, 916-931 (2016).


\textsuperscript{28} CHARLES MANSKI, \textit{Identification for Prediction and Decision} 7-8 (Harvard Univ. Press 2007).