Financial Crime and Punishment: A Meta-Analysis

Abstract

This version: August 5, 2021

A substantial amount of research has been performed on how financial markets react to the disclosure of financial crimes committed by listed firms. This research is important because an accurate quantification of those consequences in terms of drops in stock price returns is imperative since the violation of securities laws is one of the major causes of corporate failure. While there is consensus that financial crime is followed by negative stock price returns, the exact size of the effect is far from clear. We are the first to provide a quantitative synthesis of the relevant literature and relate heterogeneity in the results to heterogeneity in the type of financial crime, study design, and an array of other relevant characteristics. We survey 111 studies published between 1978 and 2020 from which we collect a total of 479 estimates from event studies. Then, we perform a thorough meta-analysis based on the most recent available techniques. We show that the negative abnormal returns found in the literature seem to be exaggerated by more than three times and, as such, the "punishment" effect suffers from a serious publication bias. After controlling for this bias, our meta-analysis indicates that the disclosure of financial crimes is followed by statistically significant negative abnormal returns, which suggests the existence of an informational effect. We also document that crimes committed in common-law countries such as the U.S. and accounting fraud carry particularly weighty information for market participants. Conversely, regulatory procedures and convicted crimes do not trigger significant abnormal reactions. The results indicate a need for more transparency about enforcement procedures in order to foster timely and proportionate market reactions and to support efficient markets.

JEL Classification: C83, G14, G18, K42, N24
Keywords: Meta-Analysis; Event study; Financial Misconduct; Information and Market Efficiency; Listed Companies.

Declarations of interest: none

1. Introduction

How do financial markets react to the disclosure of financial crimes committed by listed firms and how large is the reaction in quantitative terms? Most of the literature we analyze is in some agreement that financial crime is punished by a drop in stock price, i.e. a negative return, but researchers are less in agreement on its quantification. As we show, the negative abnormal returns found in the literature seem to be exaggerated by more than three times and, as such, the "punishment" effect suffers from a serious publication bias (Doucouliagos and Stanley, 2013; Ioannidis et al., 2017; Bajzík et al., 2020; Blanco-Perez and Brodeur, 2020; Brodeur et al., 2020; Havránek and Sokolova, 2020; Gechert et al., 2021; Sokolova and Sorensen, 2021; Zigraiova et al., 2021). The publication bias also echoes the notorious search for statistically significant results within the academic community to maximize publication probability, as emphasized by Brodeur et al. (2016). High negative abnormal returns documented in the literature are disturbing because understanding the true underlying dynamics of the punishment of financial crimes is key to better enforcement of the financial regulation of securities markets and protecting investors (La Porta et al., 2000). The reason is that every investor should have access to quality information about listed firms prior to and after investment (Black, 2000). This arrangement forms a basis for the trust on which the existence and efficiency of capital markets depend (Amiram et al., 2018). Trust is formed by the ex-ante belief that one's counterpart will suffer consequences for opportunistic or fraudulent behaviors (Dupont and Karpoff, 2019). An accurate quantification of those consequences in terms of drops in stock price returns is imperative since the violation of securities laws is one of the major causes of corporate failure (Soltani, 2014). Because of the non-marginal heterogeneity in studies dealing with market reactions to disclosed financial misconduct, we analyze the extant literature both to uncover sources of displayed heterogeneity and also to provide the most accurate estimation of the extent of the punishment of financial crime (Stanley and Doucouliagos, 2019). To the best of our knowledge and confirmed by a recent review of meta-analysis in finance (Geyer-Klingeberg et al., 2020), this is the first meta-analytic assessment of this pressing issue.

Further motivation for our analysis comes from the fact that recent in-depth reviews by Amiram et al. (2018) and Liu and Yawson (2020) document a substantial growth of empirical literature assessing the adverse link between financial crimes and corporate financial performance. When there is corporate misconduct, financial crimes trigger the strongest market reactions when disclosed and shareholders can be harmed by the misconduct itself. But the crime can further severely damage corporate reputation (Engelen, 2011; Engelen and van Essen, 2011; Haslem et al., 2017; Karpoff, 2012 and 2020). What are the consequences? The disclosed

financial misconduct of listed firms becomes publicly available information and, based on the semi-strong efficient market hypothesis, it should be reflected immediately, fully, and in an unbiased manner in the stock prices (Fama, 1970). Hence, the market sanction should reflect the forecasted cumulative costs of long-term impacts including fines, legal fees, compensations, cost of doing business, etc. (Dechow et al., 1996; Palmrose et al., 2004). In this sense, financial markets are an enforcement channel inducing companies to behave responsibly (Engelen, 2011).

Another perspective that motivates our analysis comes from the regulatory side. One segment of the extant literature quite often concludes that the negative impact on returns from disclosed financial crimes is, to some extent, due to imposed legal penalties (Karpoff and Lott, 1993; Alexander, 1999; Karpoff et al., 2005; Murphy et al., 2009; Engelen, 2011; Haslem et al., 2017; Karpoff et al., 2017; Armour et al., 2017). Conversely, Morris et al. (2019; p. 318) emphasize that "theory suggests that regulator action may result in limited or no benefits, and the empirical evidence to this effect is mixed. A hint of financial crime can dampen investors' expectations, as a reflection of corporate shortcomings or degraded future prospects, and market quality should deteriorate. However, the SEC's investigation can be an opportunity for the firm to correct internal problems and bad behaviors. Market participants may then respond positively during the investigation thereby revising forecasts to the upside." Christensen et al. (2016) empirically validate the "no-effect" hypothesis of U.S. Securities and Exchange Commission (SEC) enforcement actions on market quality, presented by Stigler (1964) and Peltzman (1976). Amiram et al. (2018) even challenge the rationale for levying fines. Despite the above dissent, finding a good balance in enforcement is important since financial markets can complement enforcement of securities laws. A bold example is that acting legally can become an economic disadvantage if the benefits from cheating the law (e.g. higher returns on assets, lower costs of financing and doing business) exceed the expected costs for being sanctioned (Becker, 1968; Aupperle et al., 1985; Hawley, 1991). In addition, the deterrent value of enforcement might rise with offenders' wealth (Garoupa, 2001).

Despite the richness of the literature, no consensus can be identified in terms of the presence, direction, and magnitude of the stock price reaction following the disclosure of an intentional financial crime; the evidence is often mixed or less than fully observed (Karpoff et al., 2017). This imperfect observability makes the meta-analysis a relevant tool to test the robustness of the results by aggregating conclusions of individual studies as indicated by Geyer-Klingeberg et al. (2020). Until now, no meta-analysis has consolidated, synthesized, and evaluated the empirical findings from studies assessing whether and to what extent stock

markets react to the disclosure of financial crimes committed by listed firms. The purpose of our study is to synthesize previous empirical results systematically and quantitatively regarding market reactions subsequent to the disclosure of intentional financial crimes, specifically when a listed firm (through its managers or employees) deliberately cheats on investors. We surveyed all available literature until May 1, 2020, and identified 862 articles published from 1978 to 2020 that use event-study methodology. We selected 111 articles and collected a large sample of 479 estimates of abnormal returns following financial crimes disclosed between 1965 and 2018 (see Appendix D, Table A.2 for stylized facts). This way, we are able to synthesize market reactions after the disclosure of 31,800 crimes in total.

Based on the large number of studies in hand, we ask if the differences are due to heterogeneity among studies. To do so, we employ a meta-regression analysis and investigate the extent and nature of the effect that materialize on a stock market after crimes are disclosed. Our meta-analysis is unique in that it covers the impacts of the disclosure of financial crimes to the widest possible extent since we cover both alleged and sanctioned (convicted) financial fraud as well as different enforcement procedures, information channels, geographical coverage, and methodological specifications. Given the market size and the high regulatory transparency, the majority of the studies investigate crimes committed in the U.S. Still, it is of great interest to put these results into perspective with a wider geographical scope, so we also cover Asian and European countries.[1]

Four important features make our analysis objective. First, we aim to cover the largest possible extent of disclosed crimes since the knowledge about financial crimes is constrained by the low share of detected crimes in the first place (Becker, 1968).[2] Alarmingly, Alawadhi et al. (2020) assess that only 3.5% of financial misrepresentations are eventually caught and sanctioned. Consequently, "our knowledge of financial misconduct comes almost exclusively from firms that were caught, and the characteristics of those firms may differ from firms that commit fraud without detection" (Amiram et al., 2018; p. 738). In line with the academic, practitioner, and policy literature, we cover the following types of financial crime committed by listed firms: insider dealing,[3] price manipulation,[4] breaches of public disclosure

[1] We cover 17 countries (ordered alphabetically): Australia, Belgium, Canada, China, France, Germany, Japan, Luxembourg, Malaysia, the Netherlands, South Korea, Spain, Sweden, Thailand, Turkey, the UK, and the U.S.
[2] Becker (1968) models the choice to engage in misbehavior like any other decision involving cost-benefit tradeoffs, in light of the expected profits from fraud, the probability of being caught, and the subsequent sanction.
[3] Use and/or divulgence of insider information for investment decisions.
[4] Deliberate misconduct to influence securities prices and fair price formation.

requirements,[5] and any other breach of financial regulations.[6] The nature of financial crimes implies that the sample contains only "bad" news, which are priced-in by the market more rapidly than "good" news (Taffler et al., 2004). Plus, the expectation of negative abnormal returns, following "bad" news, suggests the existence of potential publication bias and supports the relevance of meta-analyzing this literature. Further, we analyze market reactions to intentional financial crimes since unintentional errors likely provoke a different market response (Lev et al., 2008; Hennes et al., 2008).[7] Figures A.1 and A.2 in Appendix B provide a comprehensive account of the types of crime and steps taken by regulators.

Second, in accordance with a general introduction and overview of meta-analysis applications in financial economics (Geyer-Klingeberg et al., 2020), our dataset represents a unique international sample of market reactions to financial crimes, even though international evidence is not available at the primary study level. This is even more vital due to the existence of country-specific features and the non-existence of global datasets.[8]

Third, our dataset provides directly available and comparable results since it is derived from studies that employ an event-study approach (see Appendix A for methodological details). The event study methodology, originally outlined in Ball and Brown (1968) and Fama et al. (1969), is widely recognized in the finance and empirical economic literature as an efficient tool to analyze abnormal market reactions to unanticipated news (MacKinlay, 1997; Kothari and Warner, 2008). This methodology has proven to be particularly adequate in policy analysis (Fama, 1990; Bhagat and Romano, 2002a, b) as well as in financial analysis (Geyer-Klingeberg et al., 2020). It isolates and quantifies firm-specific abnormal movements in security prices (so-called "abnormal returns") over a specific time interval (the event window) around the event. It avoids the issue of endogeneity and is quite unambiguous with regard to the causal direction of the relationship (Endrikat, 2016). The event study methodology is particularly relevant for the scope of this meta-analysis on financial crime as the event dates are precisely known and are most often communicated via official channels. This also facilitates the search for confounding events and their avoidance. Additionally, we limit the scope of the surveyed studies to short-term event windows because Kothari and Warner (1997) and Bhagat and Romano (2002a), among others, raised serious concerns about the specification and explanatory

[5] Failure to comply with financial reporting laws and regulations, most frequently accounting fraud.
[6] Failure to meet professional obligations.
[7] Unintentional financial crimes are mostly accounting restatements due to changes in accounting standards or in consolidation perimeters. Hence, they do not signal corporate misconduct but rather accounting changes that have to be considered.
[8] It is worth emphasizing the excellent initiative by the European Securities and Markets Authority (ESMA) in creating a European repository of sanctions applied and published in the EU member states.

power of event studies with long-term event windows because they are plagued by a higher noise-to-signal ratio and a higher occurrence of other confounding events interfering with the investigated event.

Fourth, we capitalize on recent methodological innovations in an emerging literature on meta-analysis in economics and finance. We first analyze the entire sample of estimates, graphically and statistically, using linear and non-linear models to investigate for publication bias (Ioannidis et al., 2017; Andrews and Kasy, 2019) and for the true effect beyond bias (Stanley and Doucouliagos, 2012; Bajzík et al., 2020). Then, we focus on more homogeneous subsets of estimates, depending on the investigated countries (the U.S. versus the rest of the world) and on the financial crimes (accounting crimes versus general breaches of securities laws). Finally, since no (or little) theory supports the choice of variables in the primary studies, we circumvent the inherent model uncertainty by using Bayesian and frequentist methods of model averaging to choose the most important factors as in Bajzík et al. (2020), Havránek and Sokolova (2020), Gechert et al. (2021), Kočenda and Iwasaki (2021), Matousek et al. (2021), Sokolova and Sorensen (2021), and Zigraiova et al. (2021). Our contribution to the literature can be summarized in the number of findings that represent the true state of reality assessed via meta-analysis. The evidence we survey suggests that disclosing the involvement of a public firm in a financial crime substantially dampens the wealth of shareholders, quantified as negative abnormal returns over the few days around the event (an initial naïve average of -1.82% per day). However, our results hint that this evidence suffers from serious publication bias towards negative abnormal market reactions, leading to biased estimates and distorted inferences. After controlling for the bias that large and negative results are more likely to be published than others, our meta-analysis shows an average loss in returns of -0.53% per day over the event window following the disclosure of financial crimes (or -2.12% in cumulative returns).

Based on the results from complementary analyses, we show that, beyond standard errors, some factors contribute to the reported abnormal returns – in particular, the sampled countries and the types of crime. Crimes committed in the U.S. (and more generally in common-law countries, where enforcement is more transparent) and accounting fraud drive market corrections. Our findings contribute to a regulatory debate on how to come closer to an optimal level of regulation to deter future crimes. The recent shift towards the "name and shame" mechanism adopted for accounting standards enforcement in the U.S., Germany, and the UK corresponds to the evidence on negative abnormal returns as well. It implicitly assumes that investors will react negatively to published findings of erroneous accounting treatments. Their

behavior will penalize the firms as a substitute for financial fines and provide incentives for peer firms to not break the law. In terms of policy implications, our analysis demonstrates how transparent enforcement actions are priced-in by market participants. Hence, if an enforcer's goal is that markets react to their decisions and communications, then enforcement actions serve as a regulatory tool per se.

The rest of the article is structured as follows. We first detail in section 2 how the data was collected and present the big picture of the information extracted from the studies. The assessment of the extent of the publication selection bias is detailed in section 3, followed by a heterogeneity analysis of why market reactions vary between studies (section 4). Finally, section 5 concludes and proposes policy-related interpretations.

2. Data selection and stylized facts

2.1 Selection of the data

Following the recent guidelines for meta-analytic research (Havránek et al., 2020), we selected an initial set of studies by systematic keyword searches performed in Google Scholar, which has the advantage of going through the full texts of studies and not only titles, abstracts, or keywords. We searched for specific topics related to financial crime and punishment via combinations of two groups of relevant keywords. The first group included *financial crime*, *regulatory breach*, *misconduct*, *fraud*, *sanction*, *penalty*, *class action*, *restatement*, and *lawsuit*. The second group included *firms*, *financial market*, *event study*, *return*, and *abnormal*. We examined the first 500 papers returned by the searches in Google Scholar. The search was complemented through other major economic databases such as JSTOR, Econlit, Science Direct, RepEc (IDEAS), NBER, CEPR, and SSRN. After this first selection of papers relevant to our study, we systematically inspected the lists of references in these studies, and the Google Scholar citations, to check if we could find usable studies not captured by our baseline search. No a priori filter was used concerning the date or type of publication. This procedure further increased the number of potential studies. We terminated the search on May 1, 2020 and did not add any new studies beyond that date. In total, 862 articles were reviewed and analyzed.[9]

[9] We tried to circumvent the fact that language issues can act as a constraint on the scope of meta-analyses. We extended searches to the following languages: English, French, German, Portuguese, and Spanish. Some articles in Chinese, Japanese, and Turkish could not be included in the literature review, although they appeared relevant in view of their references. Still, as stressed by Reurink (2018), the representativeness of the presented findings remains skewed heavily toward the Anglo-Saxon world.
The iterative process of selecting articles is graphically illustrated by the PRISMA statement displayed in Appendix C in Figure A.4, as recommended by Havránek et al. (2020).

We form our dataset from studies that strictly satisfy the following six conditions in that they must: 1) use a daily event study methodology; 2) analyze market reactions to the disclosure of intentional or alleged financial crimes (see Appendix C, Figure A.3, for a graphical illustration of the scope of the sample); 3) specify the first public reporting of the crime, whatever the source of information (newspaper, regulatory or corporate communication); 4) report (Cumulative) Average Abnormal Returns ($(C)AARs$) and an explicit indication of statistical significance (*t*-statistics, *p*-values, *z*-statistic, and/or a significance level (1%, 5%, or 10%)), to calculate (or proxy) standard errors;[10] 5) use short-term event windows, defined as two business weeks before and after the event; and 6) not be Masters or PhD theses (working papers are included). Studies not satisfying all six conditions were excluded.

At the end of our selection process, we had a set of 111 studies. Out of these studies, 90 were published in academic journals (81%) and the rest are working papers, colloquium proceedings, or chapters of collective publications. Our aim is to analyze how, and to what extent, the disclosure of intentional financial crimes committed by listed firms impacts their stock price returns. For that, we follow Stanley and Doucouliagos (2012) and Hubler et al. (2019) and extract all short-term $AARs$ and $CAARs$ included in the 111 articles, with their respective event windows, ranging from -10 to +10 trading days around the event occurring in $t = 0$.[11] We obtain 479 effect estimates from the disclosure of 31,800 intentional financial crimes committed by listed firms (i.e. the events).

Including event windows preceding the events controls for market anticipations of the news, resulting from potential corporate or regulatory leaks of information. Including 10 trading days after the event controls for the time persistency of the impact and some market inefficiencies, if the reaction is not full and immediate (Fama, 1970). The aim to use all $(C)AARs$ from the primary studies is to get as many estimates as possible to account for the variability found across studies and between estimates, without introducing potential selection bias and to properly weight the reported findings. However, this approach results in potential interdependence between studies that we accommodate for by systematically clustering the dataset by studies. For each study, the complete reference can be found in Appendix I, and

[10] We made the choice to only include studies containing both $(C)AAR$s and information on statistical significance, which are the natural output of event studies. In this sense, our approach is stricter than that of Lane (2016) who sent data requests to about half of the authors of primary studies when he could not construct the effect sizes using the information provided in the primary studies themselves.

[11] For $AARs$, we only included the results for the following days: $AAR[-1]$, $AAR[0]$, and $AAR[+1]$. These are not only – by far – the most frequent, but also more importantly the most meaningful, capturing possible anticipation by the market and some market inefficiencies. Some studies published 21 *AARs* for the 21-day event window, with hardly any being significant.

descriptive statistics in the meta dataset in Appendix D, Table A.2.

2.2 Descriptive statistics

Reported abnormal returns are comparable between articles as, by construction, they all use an event study methodology (see Appendix A for methodological details). Still, there is no standardized event window, depending on the authors' ad hoc decisions, even though the event day ($t = 0$) is at least included in the reported event windows (see Appendix E, Table A.3). Hence, we follow Veld et al. (2018) and normalize all reported abnormal returns by the length of their respective event windows to create the variable Average Abnormal Return per Day ($AARD$) that equals the $CAAR$ divided by the length of the event window (in days) or the AAR for one-day event windows.[12] $AARDs$ are winsorized at the 1% level, to ensure that the presence of a few outliers does not result from mistakes in the original articles.[13]

Event studies typically use hypothesis tests for the statistical significance of abnormal returns around the event day and, conventionally, the null hypothesis is that $AARDs$ equal zero. The great majority of studies in the sample (80% of the sample) reports a statistical significance level, usually complemented with some statistics (Student's *t*-test statistics, *z*-statistics, *p*-values, and on rare occasions, non-parametric test statistics). Often, no (or little) information on how the test was run is given. The parametric *t*-tests (or statistical significance levels) are provided by the primary studies themselves, under the assumption that the underlying source population is normally distributed. This assumption is never discussed in the literature, given the large sample sizes. As done by Frooman (1997), when the *t*-statistics are not published, conservative *t*-statistics are estimated as follows: 1) the statistical significance levels are converted into conservative levels of significance;[14] 2) the *z*-statistics are directly changed into *t*-statistics, with the assumption that, as sample size increases, Student's *t* distribution approaches a normal distribution (Marascuilo and Serlin, 1988); and 3) the *p*-values are converted into *t*-statistics by using a *t*-table and the appropriate degrees of freedom. Finally, three studies report that the abnormal returns are significant, without including *t*-statistics or the statistical significance.[15] We make the conservative assumption that the statistical

[12] The $(C)AARs$ could not be standardized by their standard deviations (Frooman, 1997) as only few event studies report them. Complementarily to normalizing $CAARs$, Veld et al. (2018) included dummy variables for observations with different event windows.

[13] The means of $AARDs$ and of standard errors are slightly impacted by the winsorization, from -1.827% to -1.819% and from 0.01968 to 0.01904, respectively. The results hold with different levels of winsorization (2.5% and 5%).

[14] 10% to t = -1.645, 5% to t = -1.96, 1% to t = -2.576, etc.

[15] The three studies are Desai et al. (2006), Nelson et al. (2009), and Goldman et al. (2012), representing seven estimates.

significance level was at least 10% for each. Conservative standard errors are then calculated from the conservative *t*-statistics and the $(C)AARs$ when not included in the study.[16] Standard errors are also winsorized at the 1% level, given the few dramatic outliers with limited impact on the mean.[13]

Overall, 479 estimates compiled from the sample of 111 primary studies provide a diverse set. Figure 1 depicts the average abnormal returns per day ($AARD$) by study. Most frequently, studies investigating the spillovers of the disclosure of financial crimes report negative and statistically significant abnormal returns, with a naïve average of -1.82% per day of the event window (-2.44% when weighting by the number of estimates reported per study).[17] The large average sample size (264 on average), and hence the large degree of freedom, supports the significance of the results.

In addition to the estimated $(C)AARs$ and their statistical significance, we build a set of variables to account for the existing heterogeneity among studies, the heterogeneity introduced by the choices of authors of the primary studies, as well as the typical dimensions of research coded in meta-analysis (Stanley and Doucouliagos, 2012). In our approach, we follow the latest guidelines for conducting a meta-analysis (Havránek et al., 2020) and the best practices of other meta-analyses. We cover the data characteristics to account for structural variations, the event study estimation, and the publication of the study. A detailed definition of these variables and descriptive statistics are displayed in Appendix E (Table A.3).

The following initial observations can be made. Table 1 and Figure 2 give the first indications of the potential causes of heterogeneity in $AARDs$ by comparing sub-samples. Markets react more to financial crimes (i.e. greater negative $AARDs$ than average and higher variance) with the following characteristics: exclusively accounting fraud (-2.86%), crimes committed in the U.S. (-2.47%, this also holds more generally in common-law countries), alleged crimes (-2.21%), and crimes directly disclosed by the firms (-2.08%). Second, given the long span of financial crimes, from 1965 to 2018, Figure 3 depicts graphically an upward trend of $AARDs$ in time. Less negative abnormal returns over time hints at less responsive markets to financial crimes along time. This trend could reflect fundamental changes in market perceptions leading to less responsive financial markets (for example echoing consecutive regulatory tightening), financial crises, or an information overload of market participants). It

[16] Only two studies published standard errors of $(C)AARs$, standing for 10 estimates or 2% of the sample of estimates.

[17] Weighting by the inverse of the number of estimates reported per study assigns the same weight to each study, hence accounting for the unbalanced nature of the dataset.

could as well result from quality improvements in the data and techniques across time. Third, articles on convicted crimes tend to be published in better journals and get more attention; these features are measured, for example, by the Scopus cite score, by the RePec impact factor, and by the number of Google citations. The sampled articles were published between 1984 and 2020, thereby covering close to four decades of research. Most frequently, articles are published in refereed and cross-disciplinary journals,[18] and co-authored by more than two researchers. A third of the latter authored more than one article out of the 111-article sample, indicating expertise in the domain of financial crime.

These observations call for deeper analyses to confirm any potential systematic difference between subsamples of reported $AARDs$, to potentially correct for publication bias, and to account for potential correlations between explanatory variables.

3. Testing for publication bias

3.1. Publication bias and funnel plots

A publication bias means that published manuscripts are biased in the direction or strength of the findings as the result of the combined actions of researchers, reviewers, and editors (Stanley, 2005). This bias distorts empirical evidence and policy recommendations (Bom and Rachinger, 2019). It is also worth stressing that event studies can be easily subjected to *p*-hacking, a theme that has been receiving increased attention recently (Brodeur et al., 2020 and Bruns and Ioannidis, 2016, among others). Authors can play with the event windows to get results with the "expected" sign and statistical significance or they can ignore statistically insignificant estimates or estimates with the "wrong" sign. A publication bias towards negative abnormal returns would demonstrate a tendency to search for negative abnormal returns in response to misconduct, in line with the hypotheses of efficient markets and rational investors.

We construct funnel plots to graphically analyze the distribution of the reported estimates of the impact of financial crimes on returns, which could illustrate a potential publication selection bias (Egger et al., 1997; Stanley and Doucouliagos, 2010). We plot the size of the estimated effect ($AARDs$) on the horizontal axis against a measure of the estimate's precision (the inverse of the standard errors of the $AARDs$) on the vertical axis. Without a publication selection bias, the effect sizes reported by independent studies vary randomly and symmetrically around the "true" value of the effect (Stanley and Doucouliagos, 2012). They

[18] This confirms Amiram et al. (2018)'s observation that studies on financial misconduct belong to three perspectives: law, accounting, and finance (for our sample, by declining order of importance: finance, accounting, business, and law).

should form an inverted funnel, with the most precise estimates being closer to the true mean, and less precise estimates being more dispersed. Additionally, the dispersion of effect sizes should be negatively correlated with the precision of the estimate. Figure 4 compares the distributions of the $AARDs$ against their precisions depending on the countries under review, with a split between the U.S. and other countries (Panel A),[19] and on the types of financial crime, with exclusively accounting fraud and more generally violations of securities laws (Panel B), to compare and possibly discriminate by groups similarly to Lane (2016). In both cases, the funnel plot is skewed to the left, confirming average negative $AARDs$. What is more puzzling is that the distribution of $AARDs$ is clearly asymmetrical to the left (more negative $AARDs$). This suggests a publication selection bias, under the assumption of a "true" effect holding for the whole sample regardless of the studies' specificities. This skew could indicate a preference in the literature for reporting more negative abnormal returns in the aftermath of disclosed intentional financial crimes committed by listed firms. This is particularly acute for articles focusing on the U.S. (see additional graphical illustrations in Appendix F, Figure A.5) and, to a lesser extent, on accounting fraud. Complementarily, and echoing the literature on *p*-hacking, the histograms of the distribution of the *t*-statistics (frequency and Kernel densities displayed in Figure 5) indicate jumps in the distributions at the critical significance levels (5% and 1%). They also suggest that the main source of publication bias is the underreporting of positive $AARDs$ in the literature, even if the true effect of the disclosure of financial crimes is negative.

3.2. Quantification of the publication bias and the true effect of the disclosure of financial crimes

The publication selection bias is further investigated with the Funnel-Asymmetry Test (FAT). In addition, we use a Precision-Effect Test (PET) to assess the true (i.e. beyond bias) impact of the disclosure of financial crimes on returns. Equation (1) is specified to test the correlation between the reported estimates and standard errors:

$$AARD_{i,j} = \beta_0 + \beta_1 SE_{i,j} + \varepsilon_{i,j}, \quad (1)$$

where $AARDs$ are the average abnormal returns per day (i.e. the reported effect), $SE_{i,j}$ are the standard errors of the $AARDs$, β_0 and β_1 are the parameters to be estimated, i and j denote the

[19] Another possible split, with similar results, is by types of commercial law enforced in a country: common or code law. As in Leuz et al. (2003) and Liang and Renneboog (2017), we assume that the type of commercial law is predetermined and exogenous to our analysis as the legal frameworks were set centuries ago via complex interactions (wars, occupations, and colonization, amongst others). It is noteworthy that common-law countries (and in particular the U.S.) are more transparent along enforcement or legal procedures. Therefore, they have a higher share of alleged than convicted crimes in the literature.

i^{th} estimate from the jth study ($j \in [\![1; 111]\!]$), and ε are the residuals. A publication selection bias (FAT) is demonstrated by a statistically significant correlation between the reported effects and their standard errors ($\beta_1 \neq 0$), resulting in an asymmetrical funnel plot as previously described (see Figure 4). The estimated intercept between the $AARDs$ and their standard errors β_0 (PET) stands for an unconditional measure of the genuine empirical effect of the disclosure of financial crimes on the returns of the involved listed firms, corrected for any publication selection bias (Stanley and Doucouliagos, 2012).

The results of the estimation of Eq. (1) are presented in Table 2 for the full set of $AARDs$ (column [1]) and the four sub-samples, echoing the previous observations of heterogeneity in the sample: crimes committed in the U.S. or in other countries (columns [2] and [3]) and exclusively accounting fraud or violations of securities laws (columns [4] and [5]). To support the robustness of the results, we compare three types of estimation technique following recent testing innovations. First, Panel A is based on unweighted data using the following approaches: a baseline OLS regression; an OLS regression adding study-level fixed effects, to account for unobserved study-specific characteristics (such as quality, but also to some extent for the country specificities as most of the studies focus on one country); a regression using between-study variance; a hierarchical Bayes (Bajzík et al., 2020); and instrumenting for the standard error with the number of observations reported by study (Havránek and Sokolova, 2020).[20] Second, Panel B uses a weighted-least-squares model of panel A with weighting 1) by the precision (i.e. the inverse of the standard errors) to adjust for the apparent heteroskedasticity in the regression (Stanley and Doucouliagos, 2017)[21] and 2) by the inverse of the number of estimates reported by the study, to give equal weight to every study whatever the number of estimates. Third, Panel C uses recent non-linear estimation techniques. These techniques relax the implicit assumption made in Panels A and B that the publication bias is a linear function of standard errors. This panel is comprised of 1) the weighted average of adequately powered estimates (WAAP) designed by Ioannidis et al. (2017), which focusses only on estimates with adequate statistical power; 2) the selection model by Andrews and Kasy (2019), which corrects

[20] Estimated $AARDs$ and their standard errors could potentially be jointly determined. As Havránek and Sokolova (2020) emphasize in related methodological approach, we account for this possible endogeneity by using the number of financial crimes of the event study as an instrument, which is correlated with the standard errors by construction but not - a priori - with the event study methodology.

[21] Beyond the advantage of giving more weight to more precise results, Havránek and Sokolova (2020) summarize the limits of weighting by the precision: in economics, and contrary to medicine, the estimation of standard errors is an important feature of the model and, if the study underestimates the standard error, weighting by precision can create a bias by itself. More generally, Lewis and Linzer (2005) show that, in estimated-dependent-variable models, weighted-least-squares usually leads to inefficient estimates and underestimated standard errors, and that OLS with robust standard errors yields better results.

the publication bias by estimating the probability of publication of each estimate in the literature depending on its *p*-value, based on the observation that conventional cut-offs for the *p*-value (0.01, 0.05, and 0.10) are associated with jumps in the distribution of reported estimates. This model is based on the observation that the conditional publication probability (depending on the results of the study) can be non-parametrically identified and corrected for in light of the jumps in *p*-value cut-offs; and 3) the stem-based bias correction method (Furukawa, 2019), which focusses on the most precise estimates (median values from each study, as in Gechert et al., 2021) to minimize the tradeoff between variance and bias.[22] Studies with the highest precision are called the "stem" of the funnel plot. They are used to estimate a bias-corrected average effect, under the assumption that precise studies suffer less from publication bias than imprecise studies. The model is optimized over a bias-variance trade-off (as the most precise studies suffer from high variance) and the results are generally more conservative, with wide confidence intervals. We systematically cluster standard errors by study to control for the data dependence within studies (Stanley and Doucouliagos, 2012), as the dataset is comprised on average of 4.3 (unlikely independent) estimates per study.

Table 2 confirms the significant publication bias in the analyzed literature towards negative estimates of $AARDs$ hinted at by the funnel plots. Whatever the estimation methods (Panels A and B), all samples have highly statistically significant and negative coefficients for standard errors clustered by studies. The publication bias is particularly high when weighting by precision (panel B.2), when adding study-level between effects (Panel A.3), when instrumenting by the number of observations reported by the study (Panel A.5), and when weighting by the inverse of the number of reported estimates (panel B.2). Additionally, the genuine underlying empirical effect beyond the distortion due to publication bias is negative and statistically significant, but much more limited than the averaged estimates shown in Table 1 (in particular for non-linear techniques).[23] Most of the $AARDs$ reported in the original studies are accounted for by publication bias. This indicates that markets would be much less responsive to the disclosure of financial crimes than initially thought. For the whole sample, the effect beyond bias on $AARDs$ (column [1]) is nearly three times lower than the naïve simple mean of the reported estimates (-0.53% per day on average across panels compared to -1.82% on average). Over the average four-day event window [-1.6; +1.4], returns would lose an

[22] The results for the Hedges test are detailed in Appendix H, Table A.4, with similar conclusions.

[23] Complementary results for the selection model of Andrews and Kasy (2019) are displayed in Appendix F, Figure A.7, with funnel plots and histograms of Z-statistics for the full sample of disclosed financial crimes and sub-samples.

accumulated -2.12% after the disclosure of financial crimes. The estimated effect beyond bias is higher for linear estimators than for non-linear estimators (-0.67% and -0.22%, respectively).

When we explore more detailed results grounded in specific sub-samples (columns 2 to 6 of Table 1), the following three conclusions can be drawn. First, the publication bias is more pronounced for crimes committed in the U.S. than in other countries (with no publication bias estimated from the simple OLS and OLS weighted by the inverse of the number of estimates). The result holds when enlarging the sample to common-law versus non-common-law countries, although to a lesser extent. Second, the observation that financial markets would be more responsive to the disclosure of financial crimes committed in the U.S. than in other countries is confirmed across all estimators beyond publication bias (by three times on average); $AARDs$ on average contract by -0.76% per day in the U.S. compared with -0.25% in other countries. This difference may be accounted for by structural differences between common- (typically the U.S.) and code-law countries in terms of disclosure, liability standards, and public enforcement. La Porta et al. (2006) conclude that common-law systems are more favorable to stock market development, as they accentuate private contracting and standardized disclosure and rely on private dispute resolution using market-friendly standards of liability. Third, differences in publication bias and the effect beyond bias are more limited between financial crimes. Accounting fraud suffers from a similar significant publication bias as violations of securities laws. Corrected for the bias, returns following accounting fraud would contract two times more than violations of securities laws (by -0.84% and -0.44% per day, respectively). Stronger reactions to (intentional) accounting fraud can be explained by the direct impacts on the Profit & Loss statements subsequent to accounting restatements.[24]

All in all, we find robust evidence of publication bias in the literature towards reporting negative abnormal returns, and of genuine empirical evidence in the collected estimates: markets penalize listed firms for engaging in intentional financial crimes, though less than initially estimated. However, some of the apparent correlations between the estimated abnormal returns following the disclosure of financial crimes and their standard errors could be driven by heterogeneity in the data and/or in the event study methodology. We investigate this issue in

[24] As a robustness check, we added to the original sample twelve additional studies. These studies were initially excluded because they either published statistical significance between samples (four articles), or did not include any information regarding the statistical significance of the results (eight articles). We made a reasonable assumption that all estimates reported in these studies were significant at the 10% level (granting a *t*-statistic of 1.645 across the board to estimated $(C)AARs$). Consequently, this compounded sample covers 123 studies, with 499 $AARDs$ estimated from 34,550 intentional financial crimes. The sample extension did not alter our findings as all conclusions were confirmed with the larger sample. Detailed results are not reported for the sake of brevity but are available on request.

the next section.

4. Why do market reactions vary among studies?

4.1. Potential factors explaining heterogeneity

This section is the first attempt to statistically explain sources of heterogeneity as, to date, the literature on the spillovers of financial crimes is constrained by the low share of detected financial crimes and by the limited information publicly available; in addition, the scope of the studies is almost systematically limited to one country. Amid all the coded variables, we selected 38 characteristics of study design as potential sources of variability in the $AARDs$. Our choice is grounded in the existing literature on the enforcement of financial regulations, on financial crimes, and on event studies. All these variables, their definitions and descriptive statistics can be found in Appendix E (Table A.3). For ease of exposition, we sort the variables into the following categories: structural characteristics, estimation characteristics, and publication characteristics potentially related to quality, which are not captured by data and estimation characteristics.

Structural characteristics

The enforcement of financial regulations is always country-specific, evolves along time, and can be characterized by various dimensions (see Appendix B for some stylized facts).

Geographical scope. Reactions to financial crimes can differ between countries, regions, and even legal origins (Djankov et al., 2008). Types of commercial law system can be split between common law (typically in the U.S. or the UK) and code law. Studies on the U.S. represent the majority of studies, despite the fact that we have 17 countries in our sample; the proportion correlates with the long history of enforcement, the size and liquidity of the financial markets, greater regulatory transparency resulting in more data availability, and tougher verdicts. By using the largest possible scope of results, a meta-analysis can challenge whether patterns observed in the U.S. can be generalized to other regions as differences in market reactions were documented to materialize due to various factors as social attitudes (Parsons et al., 2018), level of democracy (Shleifer, 2005), and sources of data (Karpoff et al., 2017). Therefore, we create a dummy variable *only U.S.*, along with another for *emerging economies*, where the enforcement standards might be lighter. Further, as in Hubler et al. (2019) and Rusnák et al. (2013), we control for some country-specific factors and complement the dataset with several measures of economic and institutional development in specific geographical locations: 1) the level of financial development proxied with the World Bank *Market liquidity* index, 2)

the economic freedom in the country under review proxied with the World Bank *Rule of law* index, and 3) the trust in national institutions proxied with the *Regulation* indicator of the Fraser Institute and the *Confidence in the government* indicator from the World Value Survey.[25]

Period under review. Since market reactions might have tamed across time (Figure 3), we investigate market reactions across time. This is supported by the long timespan of the dataset and the global trend towards regulatory tightening. Between 1965 and 2018, the number and type of information channels and the quantity of news dramatically increased, to the point that more and more research investigates the consequences of information overload (Ripken, 2006). We control for the age of the data by including a variable that reflects the mid-point year of the sample (*mid-point year*). The variable *length of the period under review* reflects how estimates differ when obtained over longer time periods.

Types of crime. The heterogeneity within financial crimes also needs to be controlled for. First, part of the literature on financial crimes surveys any violations of securities laws, for example by investigating all the sanctions made by a given authority over a period of time. Others focus on one specific type of misconduct, most frequently accounting fraud (in the U.S.) or insider trading. Accounting restatement subsequent to accounting fraud will directly impact shareholder wealth, contrary to the disclosure of other violations of securities laws, except for regulatory fines. Still, these other violations can translate into a reputational penalty regarding the lack of professionalism and business ethics of the top management of a firm (or some of its employees) by means of sharing or using insider information or by manipulating others' shares. Based on the above background, we create a dummy variable for studies investigating *exclusively accounting fraud*. Second, the literature investigates market reactions to alleged or condemned financial crimes, along the consecutive steps of enforcement (see Appendix B, Figure A.1). Fraud can be alleged by newspaper articles or by official corporate or regulatory statements (see Appendix B, Figure A.2 for a graphical illustration). In fact, in the U.S., enforcers and defendants can communicate during enforcement procedures whereas, outside the U.S., enforcement procedures are most frequently confidential until the publication of the verdict, as a way to guarantee the presumption of innocence. Under the semi-strong efficient market hypothesis (Fama, 1970), the very first hint of financial crimes (including alleged crimes) triggers the most important and significant abnormal market reaction, even when compared to the sanction publication itself (Feroz et al., 1991; Pritchard and Ferris, 2001). Solomon and Soltes (2019; p. 1) even underline the persistent-in-time stigma that arises after a

[25] https://www.fraserinstitute.org/economic-freedom/dataset?geozone=world&page=dataset&min-year=2&max-year=0&filter=0 and http://www.worldvaluessurvey.org/wvs.jsp.

fraud allegation by stressing the difference between "not guilty" and "innocent": "even when no charges are ultimately brought [after SEC financial fraud investigations], firms that voluntarily disclose an investigation have significant negative returns, underperforming non-sanctioned firms that stayed silent". We codify the *alleged fraud* as a dummy set to one when the crimes were not yet sanctioned (i.e. alleged or investigated) and zero otherwise. Third, each country has its own enforcement mix, with different weights given to public (higher in code-law countries) and private (higher in common-law countries, typically the U.S.) enforcement, and to self-regulation of the market (Djankov et al., 2008). Enforcement can also rely more on informal discussions and administrative guidance (such as in the UK, Japan, and France) or on formal legal actions against wrongdoers (like in the U.S.). Depending on the enforcement mix, the level of disclosure (during the procedure or not) and liability standards might also differ. Consequently, we control for the type of enforcement with a dummy set to one for *regulatory procedures* and zero otherwise. And fourth, the specific channel of the disclosure of financial crimes (by the media, the firm, or the enforcer) may impact the subsequent spillovers of the fraud. The media coverage of financial crimes is typically correlated with market reactions: the more articles, the stronger markets react (Feroz et al., 1991; Karpoff and Lot, 1993; Nourayi, 1994; Miller, 2006; Choi and Kahan, 2007; Barber and Odean, 2008; Fang and Peress, 2009; Tibbs et al., 2011; Fang et al., 2014; Peress, 2014). The business media can even be perceived by investors as a watchdog (Miller, 2006), whose credibility is supported by more independent sources of information than analysts and corporations (Kothari et al., 2009). Despite the above, it is acknowledged that all empirical proxies of securities fraud grounded in media coverage have some shortages when compared to broader proxies based on public or regulatory datasets that merge information on all financial reporting errors, securities litigations, or enforcement procedures (Karpoff et al., 2017). Two dummy variables control the origin of the source: *crimes disclosed by the press* and *crimes disclosed by the firm* (the other alternative being crimes disclosed by enforcers).[26]

Estimation characteristics.

The estimation characteristics cover the main possible divergences in event study methodology application (see Appendix A for details). First, the heterogeneity between samples is controlled for by three variables: 1) a dummy for *papers specifying the initial sample size* (before cleaning the data), 2) a dummy for papers explicitly explaining how they *excluded confounding events*,

[26] Some articles exploit regulatory information, which can be confidential (if a regulator shares data) or not (when datasets are built from a repository of all enforcement decisions made by a regulatory authority).

and 3) the *number of observations* (sampled crimes in the paper).

Additionally, the following six characteristics of the model are included: 1) whether the model is used to estimate abnormal returns is a *market model*, 2) whether the market index used is an *equally weighted market index*, 3) whether the source of the data is the U.S. dataset *CRSP*,[27] 4) the *number of estimates of* $(C)AARs$ *reported in the dataset per study*, 5) whether the *estimation window is specified*, and 6) whether the reported estimations are done over a *long-term event window* (beyond [-10;+10]).

Four variables control for the event windows of the reported estimates: 1) the *length of the event window* of the estimated $(C)AAR$ (1 day for AAR and 2 to 21 days for $CAAR$, given the limit put on the reported short-term estimates), as longer event windows might curb the $AARD$; whether the *event window* is 2) *strictly before* or 3) *on the event day* when the financial crime is revealed (i.e. $AAR(0)$), as in the meta-analysis on event studies on rating agencies' decisions by Hubler et al. (2019)); and 4) whether the *event window* is *"exotic"*, echoing the literature on *p*-hacking.[28] Markets can anticipate the news (Bhagat et al., 2002b) through leaks of information over the days preceding the public announcement by the firm or the regulator (Bhagat et al., 1994; Pritchard and Ferris, 2001; Djama, 2013; Gande and Lewis, 2009; Dyck et al., 2010; Nainar et al., 2014; Armour et al., 2017; Haslem et al., 2017; de Batz, 2020). Appendix F (Figure A.6) compares the funnels plots of the seven most frequently used event windows complemented with "exotic" event windows, for which the publication bias appears to be particularly strong.

Another key estimation characteristic is the statistical significance characterization, with a mere *significance level* (stars), *t-statistics*, *p-values*, and/or *z-statistics*.

Finally, some event studies are complemented with *cross-sectional regressions* and/or *reputational penalty estimations*. These characteristics of primary studies control for quality and depth. Shareholder wealth can be harmed during the fraud period and then again when the fraud is revealed due to direct and indirect subsequent costs. The cost of cumulated indirect spillovers can be called a "reputational penalty", as described by Engelen and van Essen (2011). The reputational penalty can be proxied by deducting direct costs from the abnormal market reactions following the publication of the financial crime (Karpoff and Lott, 1993; Cummins et al., 2006; Karpoff et al., 2008; Armour et al., 2017). It reflects downgraded investor expectations regarding the firm due to compounded factors (future cash flows, costs of doing

[27] The Center for Research in Security Prices (http://www.crsp.org/) for the U.S. market.
[28] Event windows are qualified as "exotic" when they stand for less than 5% of the compounded event windows (i.e. less than 24 estimates).

business, top management and human resources, etc.), as stressed by Karpoff et al. (2008) and Armour et al. (2017). In that sense, financial markets are an enforcement channel inducing companies to behave responsibly (Engelen, 2011). Reputational penalties complement enforcement as a tool to deter financial crime, contrary to, for example, foreign bribery or environmental violations (Karpoff, 2012, 2020).

Publication characteristics.

Since the sample is comprised of articles published in a peer-reviewed journals and working papers, we also investigate the publication selection bias and the sensitiveness of reported effects (abnormal returns) with respect to research quality. Veld et al. (2018) stress that, for seasoned equity offerings, articles published in top journals document higher abnormal market reactions than working papers. The four following publication characteristics are relevant for a meta-analysis, as highlighted by Geyer-Klingeberg et al. (2020): 1) the *number of authors* of the article; 2) if authors authored more than one article in the sample (*multiple authorships*), as a way to assess the level of expertise of the authors of the article; 3) the *year of publication*; and 4) whether the article was published in a *cross-disciplinary journal*, as this field of research is at the intersection between accounting, finance, law, and economics and being in a cross-disciplinary journal could increase the visibility of the findings. Finally, as is typical for meta-analyses, two variables control for the quality of the article: 1) the number of citations of the article recorded in Google Scholar (*number of Google quotes*) and 2) the *Scopus cite score* of the journal.

4.2. Endogeneity of the potential factors explaining heterogeneity

As stressed in the previous section, numerous factors could jointly influence the estimated $AARDs$ and their standard errors. That would violate the exogeneity assumption, supporting the asymmetrical funnel plot previously described, even without any publication bias. To circumvent this possible endogeneity problem, we run explicit controls for the most likely causes of endogeneity. Based on the previous description, we augment Eq. (1) with interaction variables between the standard errors and five study characteristics, along with a study fixed effect. The five study characteristics are: for structural characteristics, 1) crimes committed in the U.S., 2) exclusively accounting frauds, and 3) crimes disclosed by the firms; for publication characteristics, 4) the number of Google Scholar quotes and 5) the Scopus cite score. If a variable is associated with a strong publication bias, or simply with systematically different standard errors that might falsely indicate a publication bias, the interaction should prove strong

and statistically significant, which is not the case in the results displayed in Table 3. Out of the 10 estimated coefficients, one is significant at the 5% level and another at the 10% level for the model with all 10 variables, which could be due to chance. Moreover, the coefficients on the non-interacted standard errors are statistically significant in all cases but with the crimes committed in the U.S., and the means beyond bias are relatively similar to the baseline estimates (Table 2). Hence, we fail to model a violation of exogeneity explicitly.

4.3. Estimations

In this step, we estimate a regression to quantify the sources of heterogeneity in the surveyed studies. Such a regression is specified to explain the magnitude of the $AARDs$ with the 38 control variables $X_{k,i,j}$ detailed in section 4.1. Hence, the specification quantifying the sources of heterogeneity in the surveyed studies is defined as:

$$AARD_{i,j} = \beta_0 + \beta_1 SE_{i,j} + \sum_{k=1}^{n} \gamma_k X_{k,i,j} + \varepsilon_{i,j}. \quad (2)$$

However, the number of factors described in Sections 4.1-2 that potentially affect heterogeneity among studies is large and their inclusion in a regression might be problematic. Some variables are supported by strong rationale or theory, but others are not. They were chosen as controls in line with their relevance in the surveyed literature and with respect to the methodologies employed in the primary studies. Still, there is a great deal of uncertainty regarding which variables are truly relevant controls and the inclusion of those variables per se would disregard the problem of model uncertainty in the absence of a theoretical model. We deal with this important issue in the following way. Eq. (2) is estimated with the set of 38 variables $X_{k,i,j}$ by the three methods outlined below that minimize the model uncertainty and guarantee the interpretation with a truly relevant set of controls to explain the heterogeneity of results.

First, as recommended in Havránek et al. (2020), model uncertainty is circumvented by using Bayesian model averaging (BMA). BMA runs numerous regression models with different subsets of the 2^{38} possible combinations of explanatory variables to detect the most likely models.[29] The likelihood of a model is represented by its Posterior Inclusion Probability (PIP) calculated across models (Raftery et al., 1997; Eicher et al., 2011), whose interpretation is similar to statistical significance. The estimated BMA coefficients for every variable represent posterior means and are weighted across all models by the posterior probabilities. That way,

[29] We use the Metropolis-Hastings algorithm of the BMS R package (Zeugner and Feldkirhcer, 2015). The latter employs a Markov Monte Carlo chain.

each coefficient is assigned a PIP that reflects the probability of the variable being included in the underlying model. It is calculated as the sum of posterior model probabilities across all the models in which the variable is included. Given our lack of knowledge regarding the probability of individual parameter values, we follow the recommendation of Eicher et al. (2011) in our baseline specification by employing the unit information *g*-prior. Hence, the prior that all regression coefficients are null has the same weight as one observation in the data. Additionally, as in Bajzík et al. (2020), we use the dilution prior (George, 2010) to alleviate potential collinearity between the 38 explanatory variables. We use unweighted data in our baseline estimate, and complementary results with weighted alternatives are displayed in Appendix G. Robustness checks also include the beta-binomial random model prior (Ley and Steel, 2009), with an equal weight to each model size and the BRIC *g*-prior (Fernandez et al., 2001) as depicted in the sensitivity analysis of Figure 7.

Second, as in Gechert et al (2021), we run a hybrid frequentist-Bayesian model as a robustness check. The variables deemed unimportant based on PIPs from a BMA are excluded (Eicher et al, 2011).[30] The resulting model is estimated by an OLS with clustered standard errors.

Third, to avoid using priors entirely, we employ Frequentist Model Averaging (FMA). We use Mallow's criteria as weights for model averaging since they prove asymptotically optimal (Hansen, 2007) and orthogonalize the covariate space (Amini and Parmeter, 2012), as it is unfeasible to estimate 2^{38} potential models.

4.3. Results

Figure 6 illustrates graphically the results of the Bayesian Model Averaging. The vertical axis depicts the standard error and the 38 explanatory variables, sorted by their posterior inclusion probabilities (PIP), from top to bottom in descending order. The horizontal axis displays individual regression models, also sorted by the posterior model probabilities. A blank cell means that the variable is not included in the model. Otherwise, a blue cell indicates a positive sign for the estimated parameter in the model and conversely a red cell indicates a negative sign. Figure 7 and Appendix G depict the sensitivity of our results to the priors, by comparing our baseline set of priors to three other sets used recently in the literature. We can see that our baseline priors are relatively conservative though all the priors point to similar results, as the

[30] Eicher et al. (2011) classify the variables according to the following scale of posterior inclusion probabilities: 1) decisive between 0.99 and 1, 2) strong between 0.95 and 0.99, 3) substantial between 0.75 and 0.95, and 4) weak between 0.5 and 0.75.

PIPs rankings of the variables are broadly preserved with higher PIPs. In total, the four sets of priors tend to converge in the 10 control variables with significant PIPs.

The estimation results are reported in Table 4. All models we run confirm the prevalence of a publication bias in the literature on the spillovers of financial crimes, with a posterior probability of inclusion of the standard error in the BMA of 100%. This demonstrates that the reported abnormal returns following the disclosure of a financial crime are systematically exaggerated, even after controlling for numerous specific factors of individual primary studies. When accounted for, the magnitude of the estimated publication bias from Table 2 is only slightly lowered, from -1.4 to -1.28.

Structural characteristics. The evidence on the systematic importance of structural characteristics is rather mixed. In line with the results reported in the previous sections, the hypothesis of a more direct and stronger (more negative) effect on returns of *exclusively accounting frauds* than of other violations of securities laws is strongly corroborated by the results of the BMA and the FMA. Similarly, and with a greater economic importance, our results suggest that the very first disclosure of a financial crime triggers the strongest market reaction, in line with the literature: abnormal market reactions following disclosed *alleged financial crimes* are higher than for convicted ones, when controlling for publication bias. Surprisingly, the rest of the structural characteristics of the sample does not corroborate our hypotheses: these variables do not influence statistically and economically the magnitude of the estimated average abnormal returns per day. Still, our data sample suffers from a strong lack of cross-country variation as the great majority of the studies and of the estimates investigate the U.S. (70% and 64%, respectively). Consequently, the conclusion concerning the country-level variables should be analyzed with caution. Bearing this in mind, it is interesting to note that the level of confidence in the government would, to some extent (only for the FMA), mitigate abnormal market reactions. That would emphasize the fact that a financial crime committed in a country characterized by a trustworthy government providing credible enforcement would receive a lower response in the market. Conversely, disclosing a financial crime in countries where the government is less credible would trigger more negative abnormal market reactions. The market would compensate for shortages in enforcement and in the credibility of the government by self-regulating wrongdoers (Djankov et al., 2008). It is also worth stressing that the initial observation in the literature that the time dimension (controlled for the *mid-point year* of the data and the *publication year*) would curb market reactions is not confirmed by the BMA nor the FMA, when controlling for study design.

Estimation characteristics. Studies with a lower *number of sampled financial crimes* produce lower estimates of $AARDs$, which might reflect a small-sample bias. Similarly, our results also suggest that the event windows *strictly before the event* are associated with lower reported $AARDs$, despite some observations of anticipation in the literature. Such anticipation might be to some extent country- and event-specific. As expected, the FMA (and, to a much lower extent, the BMA) stresses that $AARDs$ estimated on the event day will exhibit more negative results; the opposite is found for studies using *"exotic" event windows*. Out of the seven characteristics of the estimation controlling for the quality and the rigor of the event study methodology application in every study,[31] only two turned out to be economically and statistically significant: the qualification of the *statistical significance with stars* and the complementary *cross-sectional regressions*. This observation might reflect a search for brevity and the limited attention brought to the description of the data in the original studies, despite the fact that the data was duly collected and cleaned, a prerequisite to any quality event study. Counterintuitively, the fact that an event study is complemented by further analysis in terms of *reputational penalty estimation* is not statistically significant, even though the most prominent articles do encourage such estimation (such as Karpoff and Lott, 1993, 2008; Armour et al, 2017). The baseline specification suggests that higher (or less negative) estimated $AARDs$ are characterized with statistical significance levels ("*stars*") and/or *z-statistics*.

Publication characteristics. Our results suggest that the *number of Google Scholar quotes* is robustly and negatively correlated with the reported $AARDs$, implying that studies reporting more negative abnormal returns are likely to be more quoted. Under the assumption that the number of citations is a good proxy for the unobserved study quality, this negative correlation would hint at the fact that better studies publish higher $AARDs$ (all else equal). It could also illustrate the negative publication bias in the sense that studies exhibiting more negative $AARDs$ get more quoted, as benchmarks of latest results compared with the existing literature. This interpretation is supported by the fact that articles published in journals with higher *Scopus cite scores* report lower $AARDs$, though with a lower posterior inclusion probability. Finally, our results suggest that co-authored articles (nb *authors*) exhibit higher

[31] The seven following variables characterize the quality of the application of the event study methodology: details regarding the data selection process with 1) the publication of the initial sample size; 2) explicit information on the exclusion of confounding events; 3) details on the estimation window over which the parameters of the model are estimated; 4) the limited precision of statistical significance with levels ("stars"); 5) the use of so-called "exotic" event windows, as a strategy to publish statistically significant $AARDs$; and 6) complementary analyses are undergone with a cross-sectional regression; and 7) a reputational penalty estimation.

$AARDs$ to some extent, even though we do not find any significant effect of *multiple authorship* nor of publishing in *cross-disciplinary journals* on the estimated $AARDs$.

4.3. Implied $AARDs$

The main takeaways of the previous investigations are that the reported average abnormal returns following the disclosure of intentional financial crimes are exaggerated by publication bias and vary systematically depending on the types of financial crime, on the countries, and on the publication outlet. Even though the scope of the studies covers 17 countries, the great majority investigate the U.S. and these studies tend to be better published. Additionally, the BMA results are based on the dilution prior, to address collinearity, which complicates the interpretation of individual estimates of partial derivatives of variables. Consequently, as in Bajzík et al. (2020), we create a dummy variable equal to zero for estimates in which we have a higher confidence, and to one for lower confidence. We use four different proxies for confidence by combining the following two parameters: (i) the quality of the outlet (either with a RePec impact factor above the *International Review of Law and Economics* or published in a peer-reviewed journal) and (ii) the rigor in the application of the event study methodology (disclosed estimation window and publication of the *z*-statistics). We then regress the reported $AARDs$ on a confidence dummy and on the standard error of the estimate. Results are shown in Table 5 for the full sample comparing higher and lower confidences (Panel A) and for subsamples depending on geography (Panel B), the type of crime (Panel C), and the implementation of the event study methodology (Panel D). Standard errors assess publication bias, while the constant stands for the mean $AARDs$ conditional on higher confidence and corrected for publication bias.

We observe great variations in Panel A depending on the definition of confidence. The resulting mean $AARDs$ range from -1.55% for the strictest definition up to -0.58% for Confidence 4, which is larger than the simple mean $AARDs$ corrected for the publication bias (Table 2). Given the substantial effect of the variable “lower confidence”, Panel A stresses that better estimates – according to our definitions – tend to be significantly more negative than the less reliable ones. Digging into the details of the subsamples, the following conclusions can be made. Differences between countries are confirmed, with much more responsive financial markets in the U.S. (-1.98%) than in Europe (-0.76%) and, to a lesser extent, in emerging economies (-0.51%). Interestingly, studies publishing estimates for “exotic” event windows and with limited statistical information tend to conclude with less responsive financial markets (-0.62% and -0.95%, respectively).

5. Concluding remarks

We present the first quantitative synthesis of the rich literature on the spillovers of intentional financial crimes committed by listed firms. Based on event study methodology, such spillovers are estimated as average abnormal returns around the disclosure of such crimes. Such market reactions are key in terms of enforcement as the market could complement or even substitute for enforcers by imposing reputational penalties, possibly at lower cost. This reflects a regulatory shift towards the "naming and shaming" of financial misconduct, for less serious but still weighty misconduct.

We examine a total of 479 estimates of abnormal returns following the publication of intentional financial crimes committed by listed firms that reported in one of the 111 research studies. We perform a meta-analysis to examine the relationship between these abnormal returns and the features of the sample of misconduct under review, the estimations, and the publication.

The results of the meta-analysis reveal a strongly negative publication selection bias in this literature, which is in line with the a priori hypothesis of efficient markets and rational investors: markets should react negatively to the disclosure of financial crimes (bad news). Our results (FAT complemented with Bayesian and frequentist model averaging, to address the model uncertainty inherent to every meta-analysis) stress that standard errors are the most prominent explanatory variable to variations in the reported abnormal returns, when they should be statistically independent. This is also supported by the results of the non-linear estimations (Ioannidis et al, 2017; Andrews and Kasy, 2019; Furukawa, 2019). The correlation between $AARDs$ and standard errors might be caused by a preference in the literature for larger $AARDs$, to compensate for larger standard errors. The publication selection bias overestimates by close to three times the abnormal market reactions subsequent to the disclosure of a financial crime. Beyond this bias, our results confirm the existence of an informational effect of the disclosure of intentional financial crimes. Crime like this are bad news regarding the firm, and it potentially leads to substantial costs for listed firms, justifying a negative market reaction.

Complementary analyses also demonstrate that some structural characteristics contribute to the materialization of the negative market reactions, with exclusively accounting and/or alleged fraud leading to more negative market reactions. The very first hint of misconduct typically triggers the strongest correction. Conversely, regulatory enforcement procedures do not significantly impact market reactions. The U.S., and more generally

common-law countries, appear to be more responsive markets to news of misdeeds, with stronger negative market reactions to the news of (possibly alleged) financial crimes.

We also assess the quality of the estimates characterized by the publication of primary studies in a peer-reviewed journal, Google Scholar quotes, journal ranking, Scopus and RePec impact factors, and methodological rigor. We find the existence of a robust correlation between the quality of the estimates and the magnitude: more quality studies report more negative $AARDs$, ranging between -0.58% to -1.55% when corrected for publication bias and the quality of the study.

The takeaways of this meta-analysis for policy recommendations depend on the regulatory goals. The intentions of enforcers and regulators may be that market participants be afraid of being associated with alleged or condemned financial crimes. Consequently, the threat of sanctions and subsequent reputational penalties should encourage compliance with regulations. The magnitude of market reactions to regulatory transparency will also depend on the ex-ante financial health of listed firms as shown on a theoretical level by Chakravarty et al. (2021) for financial institutions. Regulators may alternatively choose a lighter touch, possibly with anonymized sanction decisions or with confidential bilateral procedures. If enforcers intend markets to complement their enforcement actions by setting reputational sanctions, our results stress how regulatory transparency can be an efficient regulatory tool per se. Significant negative abnormal returns follow the publication of alleged crimes, in particular when committed in the U.S. and other common-law countries. Conversely, regulatory procedures and condemned crimes do not trigger significant abnormal reactions. Enforcers could (for example) communicate during enforcement procedures and substitute sanctions with “name and shame” strategies at lower cost. That way, market participants could better price financial crimes, should the enforcers’ objective be that markets account for their work in terms of market supervision and the detection and sanction of financial misconduct. Conversely, if regulators reckon that the regulatory sanction is sufficient (and that markets do not have to double-sentence wrongdoers), anonymization could protect listed firms, and such a decision would still stand as an educational tool.

References

Alawadhi, A., Karpoff, J. M., Koski, J. L., & Martin, G. S. (2020). The Prevalence and Costs of Financial Misrepresentation. *Working paper*, available at SSRN 3532053.

Alexander, C. R. (1999). On the Nature of the Reputational Penalty for Corporate Crime: Evidence. *The Journal of Law and Economics*, 42(S1), 489-526.

Amini, S. M., & Parmeter, C. F. (2012). Comparison of Model Averaging Techniques: Assessing Growth Determinants. *Journal of Applied Econometrics*, 27(5), 870-876.

Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial Reporting Fraud and Other Forms of Misconduct: A Multidisciplinary Review of the Literature. *Review of Accounting Studies*, 23(2), 732-783.

Andrews, I., & Kasy, M. (2019). Identification of and Correction for Publication Bias. *American Economic Review*, 109(8), 2766-2794.

Armour, J., Mayer, C., & Polo, A. (2017). Regulatory Sanctions and Reputational Damage in Financial Markets. *Journal of Financial and Quantitative Analysis*, 52(4), 1429-1448.

Aupperle, K. E., Carroll, A. B., & Hatfield, J. D. (1985). An Empirical Examination of the Relationship Between Corporate Social Responsibility and Profitability. *Academy of Management Journal*, 28(2), 446-463.

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6(2), 159-178.

Bajzík, J., Havránek, T., Irsova, Z., & Schwarz, J. (2020). Estimating the Armington elasticity: The importance of study design and publication bias. *Journal of International Economics*, 127, 103383.

Bauer, R., & Braun, R. (2010). Misdeeds Matter: Long-Term Stock Price Performance After the Filing of Class-Action Lawsuits. *Financial Analysts Journal*, 66(6), 74-92.

Barber, B. M., & Odean, T. (2008). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *The Review of Financial Studies*, 21(2), 785-818.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76, 169-217.

Bhagat, S., Brickley, J. A., & Coles, J. L. (1994). The Costs of Inefficient Bargaining and Financial Distress: Evidence from Corporate Lawsuits. *Journal of Financial Economics*, 35(2), 221-247.

Bhagat, S., & Romano, R. (2002a). Event Studies and the Law: Part I: Technique and Corporate Litigation. *American Law and Economics Review*, 4(1), 141-168.

Bhagat, S., & Romano, R. (2002b). Event Studies and the Law: Part II: Empirical Studies of Corporate Law. *American Law and Economics Review*, 4(2), 380-423.

Black, B. S. (2000). The Legal and Institutional Preconditions for Strong Securities Markets. *UCLA Law Review*, 48, 781-855.

Blanco-Perez, C., & Brodeur, A. (2020). Publication Bias and Editorial Statement on Negative Findings. *The Economic Journal*, 130(629), 1226-1247.

Boehmer, E., Musumeci, J., & Poulsen, A. B. (1991). Event Study Methodology under Conditions of Event-Induced Variance. *Journal of Financial Econometrics*, 30, 252-273.

Bom, P. R., & Rachinger, H. (2019). A Kinked Meta-Regression Model for Publication Bias Correction. *Research Synthesis Methods*, 10(4), 497-514.

Bonini, S., & Boraschi, D. (2012). Corporate Scandals and Capital Structure. In Entrepreneurship, Governance and Ethics (pp. 241-269). Springer, Dordrecht.

Brodeur, A., Lé, M., Sangnier, M., & Zylberberg, Y. (2016). Star Wars: The Empirics Strike Cack. *American Economic Journal: Applied Economics*, 8(1), 1-32.

Brodeur, A., Cook, N., & Heyes, A. (2020). Methods Matter: P-Hacking and Causal Inference in Economics. *American Economic Review*, 110(11), 3634-60.

Bruns, S. B., & Ioannidis, J. P. A. (2016). P-Curve and P-Hacking in Observational Research. *PLoS One*, 11 (2), e0149144.

Chakravarty, S., Choo, L., Fonseca, M. A., & Kaplan, T. R. (2021). Should Regulators Always Be Transparent? A Bank Run Experiment. *European Economic Review*, *136*, 103764.

Choi, S., & Kahan, M. (2007). The Market Penalty for Mutual Fund Scandals. *Boston University Law Review*, 87, 1021-1057.

Choi, S. J., & Pritchard, A. C. (2016). SEC Investigations and Securities Class Actions: An Empirical Comparison. *Journal of Empirical Legal Studies*, 13(1), 27-49.

Christensen, H. B., Hail, L., & Leuz, C. (2016). Capital-Market Effects of Securities Regulation: Prior Conditions, Implementation, and Enforcement. *The Review of Financial Studies*, 29(11), 2885-2924.

Cummins, J. D., Lewis, C. M., & Wei, R. (2006). The Market Value Impact of Operational Loss Events for US Banks and Insurers. *Journal of Banking & Finance*, 30(10), 2605-2634.

de Batz, L. (2020). Financial Impact of Regulatory Sanctions on Listed Companies. *European Journal of Law and Economics*, 49(2), 301-337.
Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC. *Contemporary Accounting Research*, 13(1), 1-36.
Desai, H., Hogan, C. E., & Wilkins, M. S. (2006). The Reputational Penalty for Aggressive Accounting: Earnings Restatements and Management Turnover. *The Accounting Review*, 81(1), 83-112.
Djama, C. (2013). Fraudes à l'Information Financière et Contrôle de l'AMF : Une Etude des Réactions du Marché Financier Français. *Revue Française de Gestion*, 231, 133-157.
Djankov, S., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2008). The Kaw and Economics of Self-Dealing. *Journal of Financial Economics*, 88(3), 430-465.
Doucouliagos, C., & Stanley, T. D. (2013). Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity. *Journal of Economic Surveys*, 27(2), 316-339.
Dolley, J. C. (1933). Characteristics and Procedure of Common Stock Split-Ups. *Harvard Business Review*, 11(3), 316-326.
Dupont, Q., & Karpoff, J. M. (2019). The Trust Triangle: Laws, Reputation, and Culture in Empirical Finance Research. *Journal of Business Ethics*, 163, 1-22.
Dyck, A., Morse, A., & Zingales, L. (2010). Who Blows the Whistle on Corporate Fraud? *The Journal of Finance*, 65(6), 2213-2253.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in Meta-Analysis Detected by a Simple, Graphical Test. *The British Medical Journal*, 315(7109), 629-634.
Eicher, T. S., Papageorgiou, C., & Raftery, A. E. (2011). Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants. *Journal of Applied Econometrics*, 26(1), 30-55.
Endrikat, J. (2016). Market Reactions to Corporate Environmental Performance Related Events: A Meta-Analytic Consolidation of the Empirical Evidence. *Journal of Business Ethics*, 138(3), 535-548.
Engelen, P.J. (2011). Legal Versus Reputational Penalties in Deterring Corporate Misconduct. In M. Ugur & D. Sunderland (Eds.), Does Economic Governance Matter?: Governance Institutions and Outcomes, 71-94.
Engelen, P.J., & Van Essen, M. (2011). Reputational Penalties in Financial Markets: An Ethical Mechanism? In W. Vandekerckhove, J. Leys, K. Alm, B. Scholtens, S. Signori, & H. Schäfer (Eds.), Responsible Investment in Times of Turmoil (Vol. 31, pp. 55-74): Springer Netherlands.
Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
Fama, E. F. (1990). Contract Costs and Financing Decisions, *The Journal of Business*, 63(1), S71-S91.
Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, 10(1), 1-21.
Fang, L. H. & Peress, J. (2009). Media Coverage and the Cross-Section of Stock Returns. *The Journal of Finance*, 64(5), 2023-2052.
Fang, L. H., Peress, J. & Zheng, L. (2014). Does Media Coverage of Stocks affect Mutual Funds' Trading and Performance. *The Review of Financial Studies*, 27(12), 3441-3466.
Feldkircher, M. (2012). Forecast Combination and Bayesian Model Averaging: A Prior Sensitivity Analysis. *Journal of Forecasting*, 31(4), 361-376.
Feldkircher, M., & Zeugner, S. (2012). The Impact of Data Revisions on the Robustness of Growth Determinants/ A Note on 'Determinants of Economic Growth: Will Data Tell?'. *Journal of Applied Econometrics*, 27(4), 686-694.
Fernandez, C., Ley, E., & Steel, M. F. (2001). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, 100(2), 381-427.
Feroz, E. H., Park, K., & Pastena, V. S. (1991). The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases. *Journal of Accounting Research*, 29, 107-142.
Frooman, J. (1997). Socially Irresponsible and Illegal Behavior and Shareholder Wealth: A Meta-Analysis of Event Studies. *Business & Society*, 36(3), 221-249.
Furukawa, C. (2019). Publication Bias Under Aggregation Frictions: From Communication Model to New Correction Method. *Unpublished Paper, Massachusetts Institute of Technology*.
Gande, A., & Lewis, C. M. (2009). Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers. *Journal of Financial and Quantitative Analysis*, 44(4), 823-850.
Garoupa, N. (2001). Optimal Magnitude and Probability of Fines. *European Economic Review*, *45*(9), 1765-1771.
Gechert, S., Havránek, T., Irsova, Z., & Kolcunova, D. (2021). Measuring Capital-Labor Substitution: The Importance of Method Choices and Publication Bias. Forthcoming in *Review of Economic Dynamics*.
George, E. I. (2010). Dilution priors: Compensating for Model Space Redundancy. In Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown (pp. 158-165). *Institute of Mathematical Statistics*.
Geyer-Klingeberg, J., Hang, M., & Rathgeber, A. (2020). Meta-Analysis in Finance Research: Opportunities, Challenges, and Contemporary Applications. *International Review of Financial Analysis*, 101524.

Goldman, E., Peyer, U., & Stefanescu, I. (2012). Financial Misrepresentation and its Impact on Rivals. *Financial Management*, 41(4), 915-945.
Guy, D.M., & Pany, K. (1997). Fraud in a Financial Statement Audit. *Working paper*, AICPA, New York.
Hansen, B. E. (2007). Least Squares Model Averaging. *Econometrica*, 75(4), 1175-1189.
Haslem, B., Hutton, I., & Smith, A. H. (2017). How Much do Corporate Defendants Really Lose? A New Verdict on the Reputation Loss Induced by Corporate Litigation. *Financial Management*, 46(2), 323-358.
Havránek, T., & Irsova, Z. (2017). Do Borders Really Slash Trade? A Meta-Analysis. *IMF Economic Review*, 65(2), 365-396.
Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., ... & van Aert, R. C. M. (2020). Reporting Guidelines for Meta-Analysis in Economics. *Journal of Economic Surveys*, 34(3), 469-475.
Havránek, T., & Sokolova, A. (2020). Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say "Probably Not". *Review of Economic Dynamics*, 35, 97-122.
Hawley, D. D. (1991). Business Ethics and Social Responsibility in Finance Instruction: An Abdication of Responsibility. *Journal of Business Ethics*, 10(9), 711-721.
Hennes, K. M., Leone, A. J., & Miller, B. P. (2008). The Importance of Distinguishing Errors from Irregularities in Restatement Research: The Case of Restatements and CEO/CFO Turnover. *The Accounting Review*, 83(6), 1487-1519.
Hubler, J., Louargant, C., Laroche, P., & Ory, J. N. (2019). How Do Rating Agencies' Decisions Impact Stock Markets? A Meta-Analysis. *Journal of Economic Surveys*, 33(4), 1173-1198.
Ioannidis, J. P., Stanley, T. D., & Doucouliagos, H. (2017). The Power of Bias in Economics Research. *The Economic Journal*, 127(4), 236-265.
Jackson, H. E., & Roe, M. J. (2009). Public and Private Enforcement of Securities Laws: Resource-Based Evidence. *Journal of Financial Economics*, 93(2), 207-238.
Karpoff, J. M. (2012). Does Reputation Work to Discipline Corporate Misconduct? In Barnett, M. L., & Pollock, T. G. (Eds.). The Oxford handbook of corporate reputation. Oxford University Press, Chapter 18.
Karpoff, J. M. (2020). Financial Fraud and Reputational Capital. In Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation, Chapter 6, pp. 153-177.
Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and Databases in Financial Misconduct Research. *The Accounting Review*, 92(6), 129-163.
Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008). The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis*, 43(3), 581-611.
Karpoff, J. M., & Lott Jr, J. R. (1993). The Reputational Penalty Firms Bear from Committing Criminal Fraud. *The Journal of Law and Economics*, 36(2), 757-802.
Karpoff, J. M., Lott, Jr, J. R., & Wehrly, E. W. (2005). The Reputational Penalties for Environmental Violations: Empirical Evidence. *The Journal of Law and Economics*, 48(2), 653-675.
Kočenda, E., Iwasaki, I. (2021). Bank Survival around the World: A Meta-Analytic Review. Forthcoming in *Journal of Economic Surveys*.
Kolari, J. W., & Pynnönen, S. (2010). Event Study Testing with Cross-Sectional Correlation of Abnormal Returns. *The Review of Financial Studies*, 23(11), 3996-4025.
Kothari, S. P., Shu, S., & Wysocki, P. D. (2009). Do Managers Withhold Bad News? *Journal of Accounting Research*, 47(1), 241-276.
Kothari, S. P. & Warner, J. B. (2008). Econometrics of Event Studies. In Handbook of Empirical Corporate Finance 1, Elsevier, Chapter 1, 3-36.
Kothari, S. P., & Warner, J. B. (1997). Measuring Long-Horizon Security Price Performance. *Journal of Financial Economics*, 43(3), 301-339.
Lane, T. (2016). Discrimination in the Laboratory: A Meta-Analysis of Economics Experiments. *European Economic Review*, *90*, 375-402.
La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2006). What Works in Securities Laws? *The Journal of Finance*, 61(1), 1-32.
Leuz, C., Nanda, D., & Wysocki, P. D. (2003). Earnings Management and Investor Protection: An International Comparison. *Journal of Financial Economics*, 69, 505-527.
Lev, B., Ryan, S. G., & Wu, M. (2008). Rewriting Earnings History. *Review of Accounting Studies*, 13(4), 419-451.
Lewis, J. B., & Linzer, D. A. (2005). Estimating Regression Models in Which the Dependent Variable is Based on Estimates. *Political Analysis*, 13(4), 345-364.
Ley, E., & Steel, M.F. (2009). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression. *Applied Economics*, 24, 651-674.
Liang, H., & Renneboog, L. (2017). On the Foundations of Corporate Social Responsibility. *The Journal of Finance*, 72(2), 853-910.

Liu, C., & Yawson, A. (2020). Financial Misconduct and Market-Based Penalties. Corruption and Fraud. In C. Alexander, & D. Cumming (Eds.), Handbook Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation Chapter 4, 65-120.
MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1), 13-39.
Marascuilo, L. A., & Serlin, R. C. (1988). Statistical Methods for the Social and Behavioral Sciences. *WH Freeman/Times Books/Henry Holt & Co*.
Matousek, J., Havránek, T., & Irsova, Z. (2021). Individual Discount Rates: A Meta-Analysis of Experimental Evidence. *Experimental Economics*, 1-41.
Miller, G. S. (2006). The Press as a Watchdog for Accounting Fraud. *Journal of Accounting Research*, 44, 1001-1033.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS med*, 6(7), e1000097.
Morris, B. C., Egginton, J. F., & Fuller, K. P. (2019). Return and Liquidity Response to Fraud and SEC Investigations. *Journal of Economics and Finance*, 43(2), 313-329.
Murphy, D. L., Shrieves, R. E., & Tibbs, S. L. (2009). Understanding the Penalties Associated with Corporate Misconduct: An Empirical Examination of Earnings and Risk. *The Journal of Financial and Quantitative Analysis*, 44(1), 55-83.
Nainar, S. M. K., Rai, A., & Tartaroglu, S. (2014) Market Reactions to Wells Notice: An Empirical Analysis. *International Journal of Disclosure and Governance*, 11(2), 177-193.
Nelson, C., Gilley, S., & Trombley, G. (2009). Disclosures of SEC Investigations Resulting in Wells Notices, *Securities Litigation Journal*, 19(4), 19-21.
Nourayi, M. M. (1994). Stock Price Responses to the SEC's Enforcement Actions. *Journal of Accounting and Public Policy*, 13(4), 333-347.
Ozeki, N. (2019). Determinants of Market Reaction to Disclosure of Accounting Misconduct: Evidence from Japan, *Securities Analysts Journal*, 57(3), 72-84.
Palmrose, Z.V., Richardson, V., & Scholz, S. (2004). Determinants of Market Reactions to Restatement Announcements. *Journal of Accounting and Economics*, 37, 59-89.
Parsons, C. A., Sulaeman, J., & Titman, S. (2018). The Geography of Financial Misconduct. *The Journal of Finance*, 73(5), 2087-2137.
Peltzman, S. (1976). Toward A More General Theory of Regulation. *The Journal of Law and Economics*, 19(2), 211-240.
Peress, J. (2014). The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes. *The Journal of Finance*, 69, 2007-2043.
Pritchard, A. C., & Ferris, S. P. (2001). Stock Price Reactions to Securities Fraud Class Actions under the Private Securities Litigation Reform Act. *Michigan Law and Economics Research Paper*, n°01-009.
Raftery, A.E., Madigan, D., Hoeting, J., 1997. Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, 92, 179–191.
Reurink, A. (2018). Financial Fraud: A Literature Review. *Journal of Economic Surveys*, 32(5), 1292-1325.
Ripken, S. K. (2006). The Dangers and Drawbacks of the Disclosure Antidote: Toward a More Substantive Approach to Securities Regulation. *Baylor Law Review*, 58(1), 139-204.
Rusnák, M., Havránek, T., & Horváth, R. (2013). How to Solve the Price Puzzle? A Meta-Analysis. *Journal of Money, Credit and Banking*, 45(1), 37-70.
Sharpe, W. F. (1970). Portfolio Theory and Capital Markets. McGraw-Hill, New York.
Shleifer, A. (2005). Understanding Regulation. *European Financial Management*, 11(4), 439-451.
Sokolova, A., & Sorensen, T. (2021). Monopsony in Labor Markets: A Meta-Analysis. *ILR Review*, 74(1), 27-55.
Soltani, B. (2014). The Anatomy of Corporate Fraud: A Comparative Analysis of High Profile American and European Corporate Scandals. *Journal of Business Ethics*, 120(2), 251-274.
Solomon, D. H., & Soltes, E. F. (2019). Is 'Not Guilty' the Same as 'Innocent'? Evidence from SEC Financial Fraud Investigations. *Working paper*.
Stanley, T. D. (2005). Beyond Publication Bias. *Journal of Economic Surveys*, 19(3), 309-345.
Stanley, T. D., & Doucouliagos, H. (2010). Picture this: A Simple Graph that Reveals Much Ado About Research. *Journal of Economic Surveys*, 24(1), 170-191.
Stanley, T. D., & Doucouliagos, H. (2012). Meta-Regression Analysis in Economics and Business (Vol. 5). Routledge.
Stanley, T. D., & Doucouliagos, H. (2017). Neither Fixed nor Random: Weighted Least Squares Meta-Regression. *Research Synthesis Methods*, 8(1), 19-42.
Stanley, T. D. & Doucouliagos, C. (2019). Practical Significance, Meta-Analysis and the Credibility of Economics. *IZA Discussion Papers n° 12458, Institute of Labor Economics (IZA)*.
Stigler, G.J., 1964. Public Regulation of the Securities Market. *Journal of Business*, 37(2), 117-142.

Taffler, R. J., Lu, J., & Kausar, A. (2004). In Denial? Stock Market Underreaction to Going-Concern Audit Report Disclosures. *Journal of Accounting and Economics*, 38, 263-296.
Tibbs, S. L., Harrell, D.H., & Shrieves, R.E. (2011). Do Shareholders Benefit from Corporate Misconduct? A Long-Run Analysis. *Journal of Empirical Legal Studies*, 8(3), 449-476.
Tukey, J. W. (1977). Exploratory Data Analysis (Vol. 2, pp. 131-160).
Veld, C., Verwijmeren, P., & Zabolotnyuk, Y. (2018). Wealth Effects of Seasoned Equity Offerings: A Meta-Analysis. *International Review of Finance*, 20(1), 77-131.
Zigraiova, D., Havranek, T., Irsova, Z., & Novak, J. (2021). How Puzzling is the Forward Premium Puzzle? A Meta-Analysis. *European Economic Review*, 134, 103714.

Table 1: Average Abnormal Returns per Day for Different Subsets of Data

Table 1 details, for the whole sample and different subsets, the average abnormal returns per day ($AARDs$), complemented by the standard errors (SE) and a 95% confidence interval. Averages are simple averages or they are weighted by the inverse of the number of estimates reported per study. Some categories are not mutually exclusive. The definitions of the subsets are available in Appendix E (Table A.3).

	Nb. Observations [1]	Unweighted				Weighted			
		Mean	SE	95% conf. int.		Mean	SE	95% conf. int.	
		[2]	[3]	[4]	[5]	[6]	[6]	[7]	[8]
1. Structural characteristics:									
Geographical specificities:									
U.S. only	304	**-2.47%**	0.18%	-2.83%	-2.11%	**-3.07%**	0.20%	-3.47%	-2.67%
Common-law countries	328	**-2.35%**	0.17%	-2.69%	-2.02%	**-2.91%**	0.19%	-3.28%	-2.53%
Code-law countries	151	**-0.66%**	0.11%	-0.87%	-0.44%	**-0.98%**	0.15%	-1.28%	-0.67%
Emerging countries	110	**-0.50%**	0.09%	-0.67%	-0.33%	**-0.47%**	0.08%	-0.63%	-0.32%
China only	89	**-0.54%**	0.10%	-0.74%	-0.34%	**-0.53%**	0.09%	-0.71%	-0.35%
Event under review:									
Exclusively accounting fraud	150	**-2.86%**	2.76%	-3.41%	-2.32%	**-3.71%**	0.29%	-4.28%	-3.13%
Exclusively violations of securities laws	194	**-1.32%**	0.15%	-1.62%	-1.03%	**-1.46%**	0.16%	-1.77%	-1.15%
Violations of securities laws (including accounting fraud)	135	**-1.37%**	0.23%	-1.82%	-0.92%	**-1.74%**	0.29%	-2.32%	-1.16%
Step of the enforcement:									
Alleged crimes (allegation in the press, initiation of regulatory procedures, investigation, class-action or lawsuit filing, etc.)	293	**-2.21%**	0.18%	-2.56%	-1.87%	**-2.74%**	0.20%	-3.15%	-2.34%
Convicted crimes (verdict of regulatory procedures, verdict of lawsuits or class-actions, accounting restatement)	186	**-1.20%**	0.17%	-1.53%	-0.86%	**-2.03%**	0.21%	-2.45%	-1.61%
Type of procedure:									
Public enforcement	255	**-1.83%**	0.18%	-2.19%	-1.46%	**-2.42%**	0.22%	-2.85%	-1.99%
Private enforcement (stock market procedure, class action, lawsuit, settlement)	188	**-1.58%**	0.18%	-1.93%	-1.23%	**-2.00%**	0.21%	-2.41%	-1.60%
Source of the news:									
Newspaper articles	202	**-1.71%**	0.18%	-2.07%	-1.36%	**-1.96%**	0.18%	-2.31%	-1.61%
Regulatory communication	319	**-1.78%**	0.16%	-2.10%	-1.46%	**-2.87%**	0.26%	-3.38%	-2.35%
Corporation communication	122	**-2.08%**	0.25%	-2.58%	-1.58%	**-2.50%**	0.20%	-2.89%	-2.11%
2. Estimation characteristics:									
Estimation model:									
Market model	398	**-1.83%**	0.14%	-2.10%	-1.55%	**-2.18%**	0.14%	-2.45%	-1.90%
Other models	81	**-1.78%**	0.33%	-2.43%	-1.13%	**-3.80%**	0.54%	-4.87%	-2.73%
Event windows:									
Before the event (t < 0)	67	**-0.78%**	0.15%	-1.07%	-0.49%	**-0.72%**	0.16%	-1.03%	-0.40%
On the event day (t = 0)	84	**-3.27%**	0.48%	-4.22%	-2.32%	**-4.57%**	0.57%	-5.70%	-3.45%
Around the event day (including t = 0)	270	**-1.85%**	0.15%	-2.14%	-1.56%	**-2.40%**	0.16%	-2.72%	-2.08%
After the event day (t > 0)	58	**-0.76%**	0.22%	-1.19%	-0.32%	**-0.52%**	0.19%	-0.90%	-0.13%
"Exotic" event windows	123	**-0.92%**	0.14%	-1.20%	-0.63%	**-1.03%**	0.14%	-1.31%	-0.75%
Exclusion of confounding events	140	**-1.22%**	0.15%	-1.51%	-0.93%	**-1.24%**	0.14%	-1.52%	-0.96%
Complementary estimations:									
Reputational penalty estimates	65	**-0.88%**	0.14%	-1.17%	-0.60%	**-1.38%**	0.18%	-1.74%	-1.02%
Cross-sectional regressions of (C)AARs	298	**-2.02%**	0.17%	-2.34%	-1.69%	**-2.50%**	0.17%	-2.85%	-2.16%
3. Publication status:									
Published papers/chapters	388	**-1.86%**	0.14%	-2.14%	-1.58%	**-1.31%**	0.17%	-1.66%	-0.97%
Unpublished papers	91	**-1.64%**	0.28%	-2.21%	-1.08%	**-2.80%**	0.40%	-3.59%	-2.02%
All estimates	**479**	**-1.82%**	**0.13%**	**-2.07%**	**-1.57%**	**-2.44%**	**0.15%**	**-2.74%**	**-2.15%**

Sources: Studies, Authors' calculations

Table 2: Meta-Regression Analysis of Publication Selection Bias

Table 2 details the results of the publication selection bias analysis, based on the FAT-PET test (Eq. (1)) for the full sample (all, column 1) and four sub-samples (the U.S. versus other countries, columns 2 and 3; accounting fraud versus other securities law violations, columns 4 and 5). The standard errors (SE) control for the publication bias (FAT) and the intercepts (PET) control for the means beyond bias. As each study reports on average four estimates, data dependence is corrected for by clustering standard errors by studies. Eq. (1) is estimated with three types of estimator: 1) unweighted estimations in panel A (OLS, study-level fixed effects, study-level between effects, hierarchical Bayes, and using the number of observations reported by the study as an instrument variable; 2) weighted least squares estimations in panel B (by the inverse of the number of estimates reported by the study and by the precision, i.e. the inverse of the standard errors); and 3) three recent non-linear estimations in panel C, with the weighted average of the adequately powered estimates (WAAP) developed by Ioannidis et al. (2017), the selection model of Andrews and Kasy (2019), and the stem-based bias correction method (Furukawa, 2019).

	All [1]		**U.S.** [2]		**Other countries** [3]		**Accounting fraud** [4]		**Violations of securities laws** [5]	
Panel A. Unweighted estimations										
1. OLS										
SE (*publication bias*)	-1.40	***	-1.52	***	-0.67		-1.39	***	-1.30	***
	(0.242)		(0.228)		(0.468)		(0.373)		(0.292)	
Intercept (*effect beyond bias*)	-0.68%	***	-0.94%	***	-0.38%	**	-1.12%	***	-0.54%	***
	(0.002)		(0.002)		(0.002)		(0.004)		(0.001)	
2. Study-level fixed effects										
SE (*publication bias*)	-1.19	***	-1.30	***	-0.59	***	-1.51	***	-1.09	***
	(0.114)		(0.150)		(0.137)		(0.192)		(0.149)	
Intercept (*effect beyond bias*)	-0.85%	***	-1.16%	***	-0.41%	***	-0.97%	***	-0.67%	***
	(0.002)		(0.002)		(0.001)		(0.003)		(0.001)	
3. Study-level between effects										
SE (*publication bias*)	-1.58	***	-1.78	***	-0.81	***	-1.49	***	-1.43	***
	(0.147)		(0.179)		(0.174)		(0.195)		(0.265)	
Intercept (*effect beyond bias*)	-0.83%	***	-0.96%	***	-0.42%		-1.24%	**	-0.63%	**
	(0.002)		(0.003)		(0.002)		(0.004)		(0.003)	
4. Hierarchical Bayes										
SE (*publication bias*)	-1.51	***	-1.78	***	-1.05	***	-1.35	***	-1.48	***
	(0.216)		(0.291)		(0.343)		(0.377)		(0.257)	
Intercept (*effect beyond bias*)	-0.62%	***	-0.66%	**	-0.32%		-0.92%	***	-0.51%	*
	(0.026)		(0.041)		(0.071)		(0.377)		(0.038)	
5. IV number of observations reported by the study										
SE (*publication bias*)	-1.59	***	-1.70	***	-0.86	**	-1.86	***	-1.79	***
	(0.428)		(0.526)		(0.433)		(0.395)		(0.604)	
Intercept (*effect beyond bias*)	-0.53%	**	-0.76%	*	-0.29%	**	-0.55%	*	-0.25%	
	(0.003)		(0.004)		(0.001)		(0.003)		(0.003)	
Panel B. Weighted least square estimations										
1. Weighted by the precision (inverse of the standard error)										
SE (*publication bias*)	-1.90	***	-1.91	***	-1.21	***	-2.01	***	-1.75	***
	(0.199)		(0.211)		(0.361)		(0.285)		(0.253)	
Intercept (*effect beyond bias*)	-0.28%	***	-0.52%	***	-0.12%	*	-0.34%	*	-0.27%	***
	(0.001)		(0.001)		(0.001)		(0.002)		(0.001)	
2. Weighted by the inverse of the number of estimates reported by studies										
SE (*publication bias*)	-1.49	***	-1.66	***	-0.75		-1.48	***	-1.25	***
	(0.262)		(0.225)		(0.525)		(0.363)		(0.309)	
Intercept (*effect beyond bias*)	-0.91%	***	-1.10%	***	-0.45%	**	-1.36%	***	-0.77%	***
	(0.002)		(0.003)		(0.002)		(0.004)		(0.003)	

	All [1]	U.S. [2]	Other countries [3]	Accounting fraud [4]	Violations of securities laws [5]
Panel C. Non-linear estimations					
1. Weighted average of the adequately powered (WAAP, Ioannidis et al., 2017)					
Effect beyond bias	-0.15% ***	-0.26% ***	-0.07% ***	-0.13% ***	-0.15% ***
	(0.0003)	(0.0007)	(0.0002)	(0.0006)	(0.0004)
2. Selection model (Andrews and Kasy, 2019)[2]					
Effect beyond bias	-0.38% ***	-0.73% ***	-0.07% ***	-1.03% ***	-0.43% ***
	(0.080)	(0.091)	(0.011)	(0.828)	(0.066)
3. Stem-based bias correction method (Furukawa, 2019)					
Effect beyond bias	-0.13%	-0.52%	-0.01%	-0.71%	-0.14%
	(0.004)	(0.006)	(0.003)	(0.006)	(0.004)
Number of observations[1]	**479**	**304**	**175**	**150**	**329**

Source: Authors' estimations.
*Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. Stars for the hierarchical Bayes are presented only as an indication of the parameter's statistical importance to keep visual consistency with the rest of the table. All standard errors (with the exception of the hierarchical Bayes) are clustered by studies and are reported in parentheses.*
[1] The available number of observations is reduced for the weighted average of adequately powered and stem-based methods.
[2] Complementary results of the selection model are displayed in Appendix F, Figure A.7, with funnel plots and histograms of Z-Statistics for the whole sample and sub-samples.

Table 3: Potential Sources of Endogeneity

Table 3 investigates potential sources of endogeneity by augmenting Eq. (1) with five variables (not reported in the table) and these variables interacted with the standard errors (SE). Estimations include a study fixed effect. Detailed definitions of the variables are available in Appendix E, Table A.3.

	Only U.S.	**Excl. accounting fraud**	**Crimes disclosed by firms**	**Nb Google quotes**	**Scopus cite score**	**All**
Standard error	-0.587	-1.05 *	-1.038 ***	-0.974 *	-1.511 ***	-0.514
(*publication bias*)	(0.429)	(0.561)	(0.352)	(0.546)	(0.506)	(0.584)
Constant	-0.013 ***	-0.0010 ***	-0.009 ***	0.079 ***	-0.005 *	0.0805 ***
(*mean beyond bias*)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.005)
SE x Only U.S.	-0.711					-0.821
	(0.601)					(0.592)
SE x Exclusively accounting fraud		-0.326				0.24
		(0.646)				(0.629)
SE x Crime disclosed by firms			-0.725			-0.551
			(0.646)			(0.579)
SE x Nb Google quotes				-0.14		-0.492 **
				(0.172)		(0.205)
SE x Scopus cite score					0.154	0.378 *
					(0.186)	(0.213)
Observations	479	479	479	479	479	479

Source: Authors' estimations.
*Notes: All standard errors are clustered by studies and are reported in parenthesis. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.*

Table 4: Why Do the Estimated $AARDs$ Vary?

Table 4 details the results for the Bayesian Model Averaging (BMA) [1], the Frequentist Model Averaging (FMA) [3], and the frequentist check (OLS) [2]. The BMA is estimated using the unit information prior and the uniform prior. For the FMA, Mallow's weights (Hansen, 2007) and the orthogonalization of the covariate space (Amini and Parmenter, 2012) are used. In the frequentist check, we only include the 11 variables with PIPs above 50% from the BMA (deemed not insignificant). The results, which may suffer from an omitted variable bias, hold when raising this threshold to 75% (deemed substantial).

Response variable:	**AARDs**	**Bayesian Model Averaging [1]**			**Frequentist Check (OLS) [2]**			**Frequentist Model Averaging [3]**		
Observations:	**479**	Post. Mean	Post. std. er.	PIP	Coef.	Std. er.	*p*-value	Coef.	Std. er.	*p*-value
Constant		1.42%	NA	1.00	1.81%	0.009	0.03	-0.07%	0.023	0.98
Standard error		-1.312	0.092	1.00	-0.469	0.075	0.00	-1.241	0.079	0.00
Structural characteristics:										
Geographical scope:	Only U.S.	0.00%	0.000	0.01				0.04%	0.005	0.94
	Emerging economies	0.01%	0.001	0.03				-1.39%	0.012	0.24
	Market liquidity (WB)*	0.00%	0.001	0.01				0.42%	0.003	0.20
	Rule of law (WB)	-0.02%	0.001	0.08				-0.56%	0.008	0.48
	Regulation (Fraser)	-0.02%	0.001	0.08				0.11%	0.003	0.74
	Confidence in the government (WVS)*	0.10%	0.003	0.12				1.83%	0.009	0.05
Period under review:	Mid-point year	-0.06%	0.002	0.10				-0.13%	0.006	0.82
	Length of the period under review	0.00%	0.000	0.00				0.27%	0.003	0.30
Types of event:	Exclusively accounting fraud	-0.40%	0.004	0.58	-1.05%	0.002	0.00	-0.65%	0.002	0.01
	Alleged fraud	-1.13%	0.002	1.00	-1.10%	0.002	0.00	-1.33%	0.002	0.00
	Regulatory procedures	0.00%	0.000	0.00				-0.11%	0.002	0.62
	Crimes disclosed by the press	0.00%	0.000	0.00				0.16%	0.002	0.49
	Crimes disclosed by firms	-0.08%	0.002	0.14				-0.35%	0.003	0.16
Estimation characteristics:										
Sample characteristics	Initial sample size published	0.02%	0.001	0.03				0.37%	0.003	0.16
	Confounding events excluded	0.00%	0.000	0.01				0.06%	0.002	0.80
	Number of events in the sample (log)	-0.14%	0.001	0.56	-0.35%	0.001	0.00	-0.28%	0.001	0.01
Model	Market model	0.00%	0.000	0.01				0.45%	0.003	0.12
	Equally weighted market index	-0.18%	0.003	0.28				-0.12%	0.002	0.62
	Source of the data: CRSP	-0.05%	0.002	0.09				-0.07%	0.003	0.82
	Number of estimates reported per study	0.00%	0.000	0.02				0.01%	0.000	0.74
	Estimation window specified	-0.21%	0.003	0.30				-0.92%	0.003	0.00
	Long-term event windows	0.00%	0.000	0.00				-0.10%	0.002	0.68
Event window of the reported AARDs	Length of the event window	0.00%	0.000	0.03				-0.02%	0.000	0.51
	Event window strictly before the event	0.96%	0.003	0.98	0.36%	0.002	0.12	0.71%	0.003	0.01
	Event window = event	-0.24%	0.003	0.36				-0.55%	0.002	0.02
	"Exotic" event windows	0.19%	0.003	0.31				0.59%	0.003	0.06
Statistical significance indicators	Significance level ("stars")	0.75%	0.004	0.86	0.79%	0.003	0.00	0.94%	0.003	0.00
	t-statistics	0.00%	0.000	0.01				0.08%	0.002	0.74
	p-values	-0.17%	0.003	0.23				-0.48%	0.003	0.09
	z-statistics	0.36%	0.004	0.56	0.53%	0.002	0.02	0.62%	0.002	0.01
Complementary results	Cross-sectional regression	-0.53%	0.004	0.68	-0.80%	0.003	0.01	-0.78%	0.002	0.00
	Reputational penalty estimation	0.11%	0.003	0.15				0.43%	0.003	0.15
Publication characteristics:										
Characteristics of the article	Number of authors	-0.18%	0.002	0.52	-0.95%	0.004	0.01	-0.43%	0.001	0.00
	Multiple authorships	-0.05%	0.002	0.11				-0.38%	0.002	0.11
	Publication year (log)	-0.03%	0.001	0.05				-0.07%	0.006	0.91
	Cross-disciplinary journal	0.01%	0.001	0.01				0.33%	0.003	0.32
Quality of the publication	Number of Google quotes (log)	-0.35%	0.002	0.90	-0.39%	0.001	0.00	-0.36%	0.001	0.00
	Scopus cite score	0.15%	0.001	0.70	0.20%	0.001	0.00	0.18%	0.001	0.01

*Source: Authors' estimations. Notes: Std. er. = Standard Error; PIP = Posterior Inclusion Probability. * Normalized*

Table 5: $AARDs$ Elasticity Implied by "Higher-Confidence" Estimation

Table 5 details the results for the estimation of Eq. (2) with a one-variable vector X: lower confidence, depending on the definition of the variable and on subsamples by geographies (Panel B), types of event (Panel C), and characteristics of the methodology (Panel D). Generally speaking, the dummy variable "lower confidence" was hinted at by Bajzík et al. (2020): 1 for estimates in which we have a "lower" confidence and 0 otherwise (i.e. "higher" confidence), based on the hypotheses below. Consequently, the constant of the regression corresponds to the mean reported $AARDs$ conditional on a higher confidence and corrected for publication bias. Four definitions of higher confidence are compared (Panel A). The strictest definition of higher confidence (Confidence 1) covers estimates for which the study was published in a high-quality peer-reviewed journal (better-ranked than the *International Review of Law and Economics*) and the event study methodology was precisely described (including the estimation window and *z*-statistics of the estimates). Confidence levels 2 and 3 are less strict combinations: being published in a high-quality peer-reviewed journal (Confidence 2) and being published in any peer-reviewed journal and precisely describing the event study methodology (Confidence 3). The lowest confidence (Confidence 4) covers estimates published in any peer-reviewed journal. For most of the estimations of Panel B-D, we used the strictest definition of confidence (Confidence 1) but when the mean of the variable "lower confidence" exceeded 99%, we used instead the Confidence 2 definition.

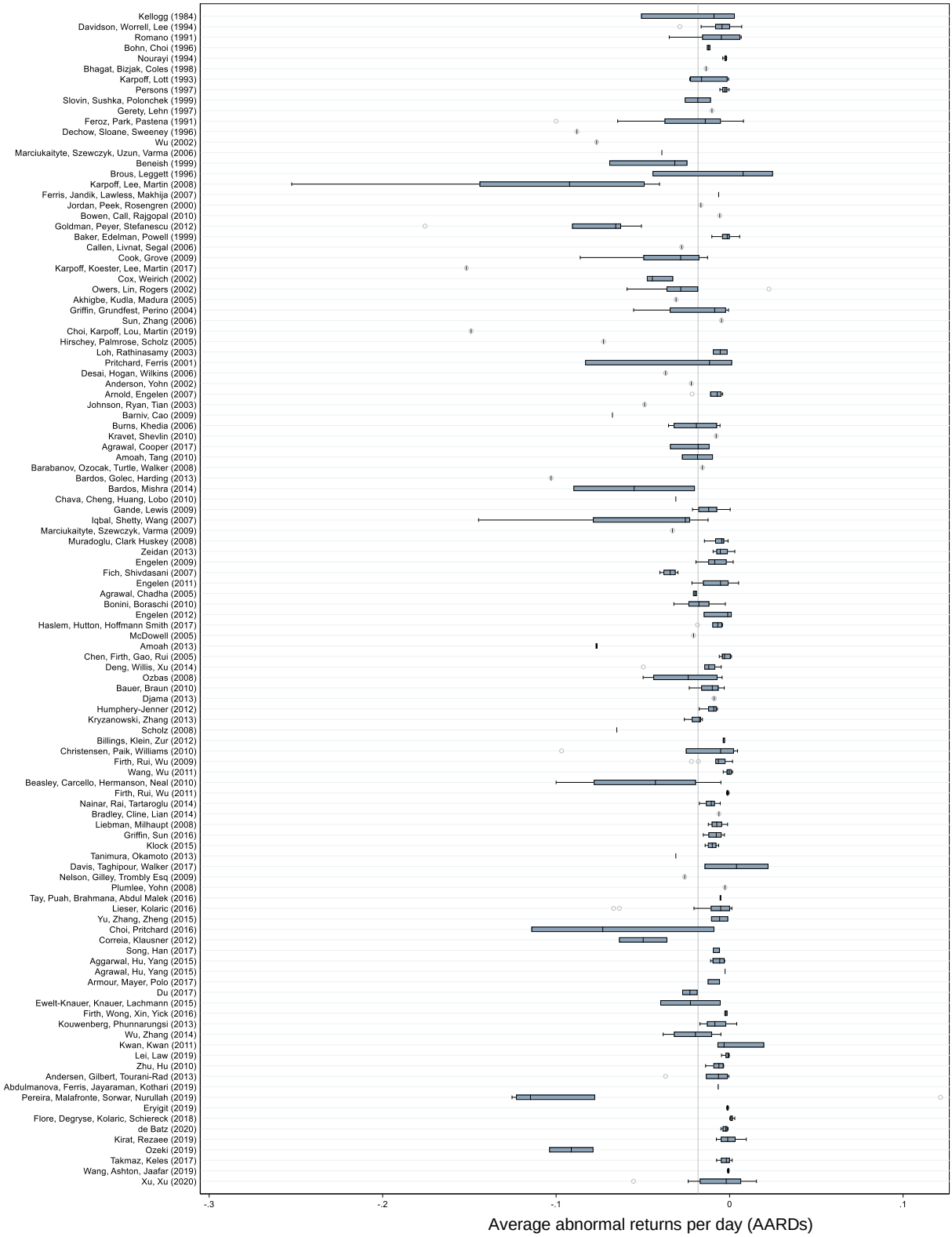
Panel A: Levels of confidence	**Confidence 1**	**Confidence 2**	**Confidence 3**	**Confidence 4**
Constant (*corrected AARDs*)	-1.55% ***	-1.41% ***	-0.22%	-0.58% ***
	(0.005)	(0.005)	(0.003)	(0.002)
Standard error (*publication bias*)	-1.408 ***	-1.405 ***	-1.427 ***	-1.422 ***
	(0.243)	(0.237)	(0.229)	(0.244)
Lower confidence	0.009 **	0.00826 *	-0.00604 *	-0.00445
	(0.005)	(0.005)	(0.003)	(0.005)
Observations	479	479	479	479
Panel B: Geography	**Only U.S.[1]**	**Common-law countries[2]**	**European Countries[2]**	**Emerging countries[2]**
Constant (*corrected AARDs*)	-1.98% ***	-1.57% ***	-0.76% ***	-0.51% ***
	(0.002)	(0.005)	(0.000)	(0.002)
Standard error (*publication bias*)	-1.523 ***	-1.527 ***	-0.113 **	-0.826
	(0.228)	(0.221)	(0.054)	(0.607)
Lower confidence	0.0108 ***	0.00817	0.00325 **	0.00263 **
	(0.002)	(0.005)	(0.001)	(0.001)
Observations	304	328	58	110
Panel C: Types of event	**Exclusively accounting frauds[1]**	**Other crimes[1]**	**Regulatory procedures[1]**	**Other procedures[1]**
Constant (*corrected AARDs*)	-1.83% ***	-1.46% **	-1.37% ***	-2.62% ***
	(0.000)	(0.007)	(0.005)	(0.001)
Standard error (*publication bias*)	-1.403 ***	-1.301 ***	-1.238 ***	-1.884 ***
	(0.376)	(0.292)	(0.273)	(0.270)
Lower confidence	0.74% *	0.94%	0.72%	2.11% ***
	(0.004)	(0.007)	(0.005)	(0.002)
Observations	150	329	255	224
Panel D: Methodology	**Exotic ev. windows[1]**	**Non-exotic ev. windows[1]**	**Only stars[2]**	**Other qualifications of stat. signif.[1]**
Constant (*corrected AARDs*)	-0.62% ***	-1.79% ***	-0.95% ***	-1.66% ***
	(0.000)	(0.006)	(0.003)	(0.005)
Standard error (*publication bias*)	0.0038 ***	0.00981	0.00789 ***	0.00894 *
	(0.001)	(0.006)	(0.002)	(0.005)
Lower confidence	-1.845 ***	-1.336 ***	-1.991 ***	-1.169 ***
	(0.100)	(0.270)	(0.183)	(0.297)
Observations	123	356	76	403

Source: Authors' estimations.

*Notes: All standard errors are clustered by studies and are reported in parenthesis. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. [1] Confidence 1 [2] Confidence 2*

Figure 1: Distribution of Average Abnormal Returns per Day ($AARDs$) Across Studies

This figure shows a box plot of the estimated average abnormal returns per day reported for every one of the 111 studies in the scope of this meta-analysis. Following Tukey (1977), the length of each bow represents the interquartile range (P25-P75), and the dividing line inside the box is the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and the lower quartiles, if such estimates exist. The grey vertical line denotes the naïve average (-1.82%). Studies are sorted by the median year of the sampled data, in ascending order.

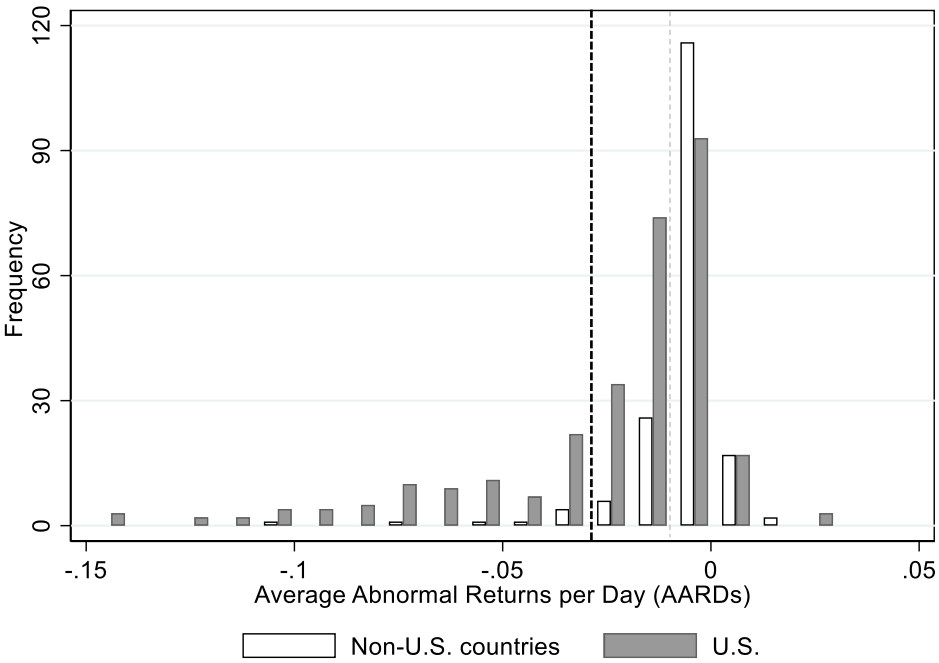


Source: Authors

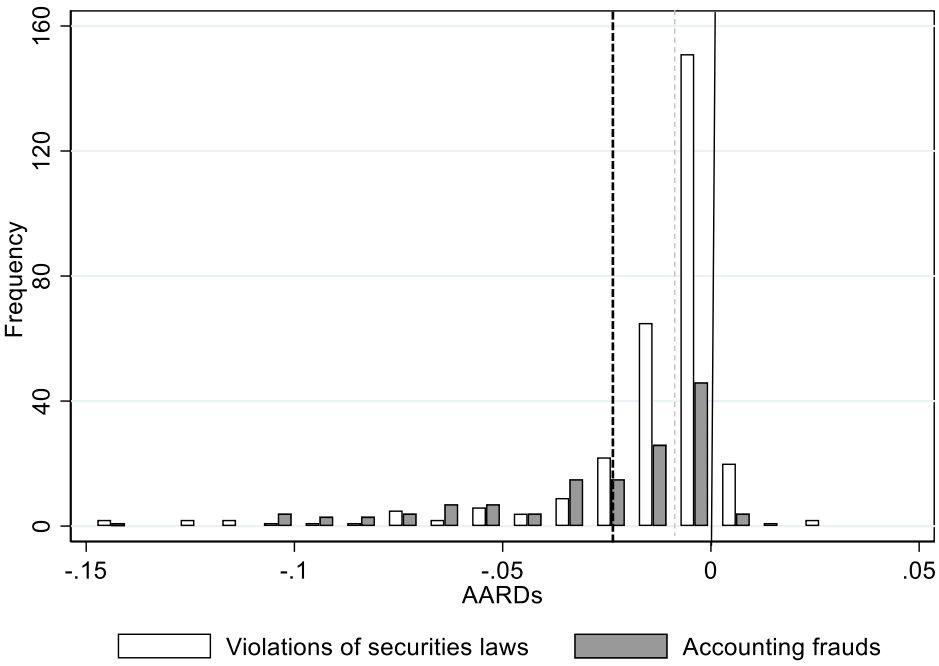
Figure 2: Frequency Distributions of $AARDs$

Figure 2 shows the histograms of the estimates of average abnormal returns per day reported in the individual studies. AARDs are split between the U.S. and non-U.S. countries (panel A) and between exclusively accounting fraud and other violations of securities laws (Panel B). Outliers are excluded from the figures but included in all the tests.

Panel A. Financial Crimes Committed in the U.S. or in Other Countries (Respectively -2.47% and -0.69% on Average, *see lines on the figure*).



Panel B. Exclusively Accounting Frauds versus Other Violations of Securities Laws (Respectively -2.86% and -1.34% on Average, *see lines on the figure*).



Source: Authors

Figure 3: Chronological Ordering of $AARDs$

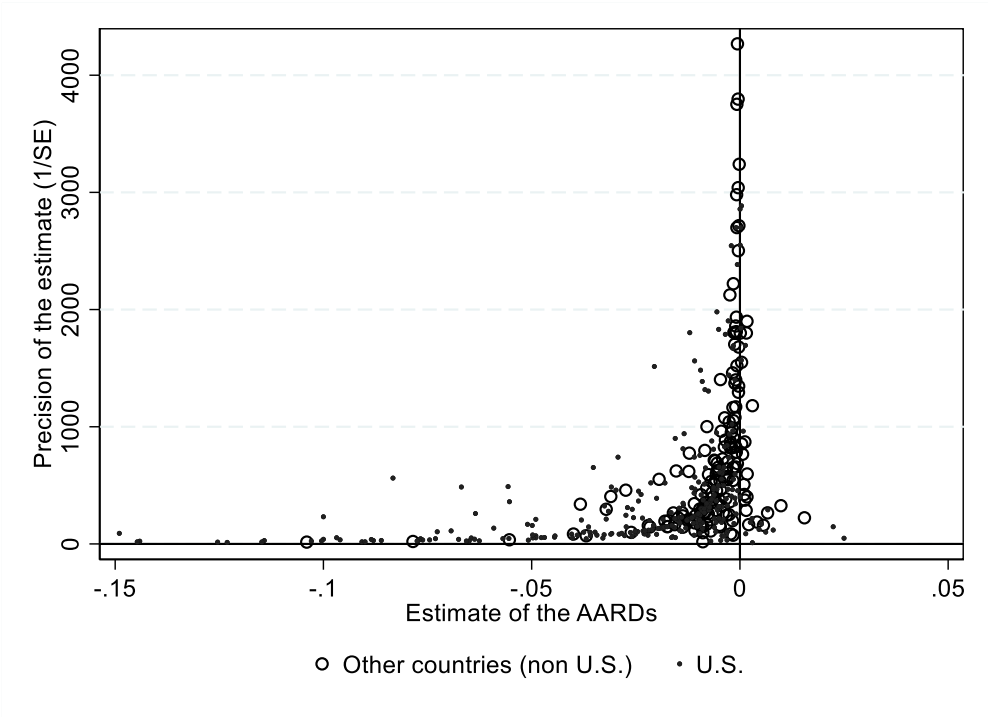
Figure 3 depicts graphically the chronological ordering of median $AARDs$ reported in individual studies, based on the median year of the data used in the corresponding study, ranging from 1973 to 2016. The dashed line is the time trend.

Source: Authors

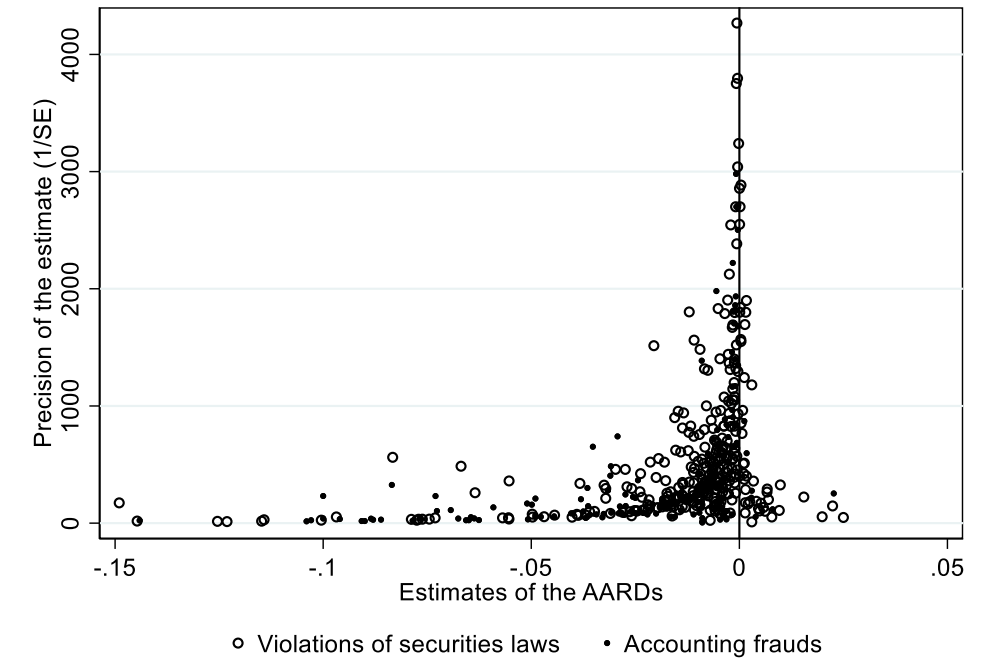
Figure 4: Funnel Graphs of the Impact of Financial Crimes

The following funnel graphs scatter the estimated average abnormal returns per day of the disclosure of financial crimes ($AARDs$) against these estimates' precision (i.e. the inverse of the estimated standard errors). The 479 estimates are split by geographical scope, depending on whether the study investigates the U.S. or other countries (Panel A) and on whether the crimes were exclusively accounting fraud or violations of securities laws (Panel B). The distribution is expected to be symmetrical around the true value of the estimate, in the absence of publication bias.

Panel A. Financial Crimes Committed in the U.S. or in Other Countries.



Panel B. Accounting Frauds versus Violations of Securities Laws.

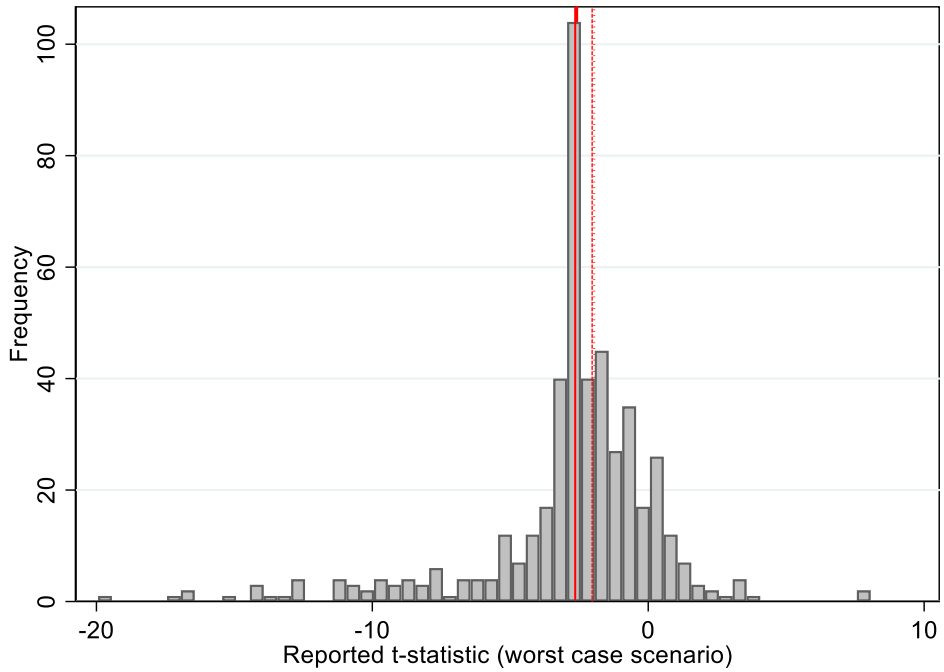


Source: Authors

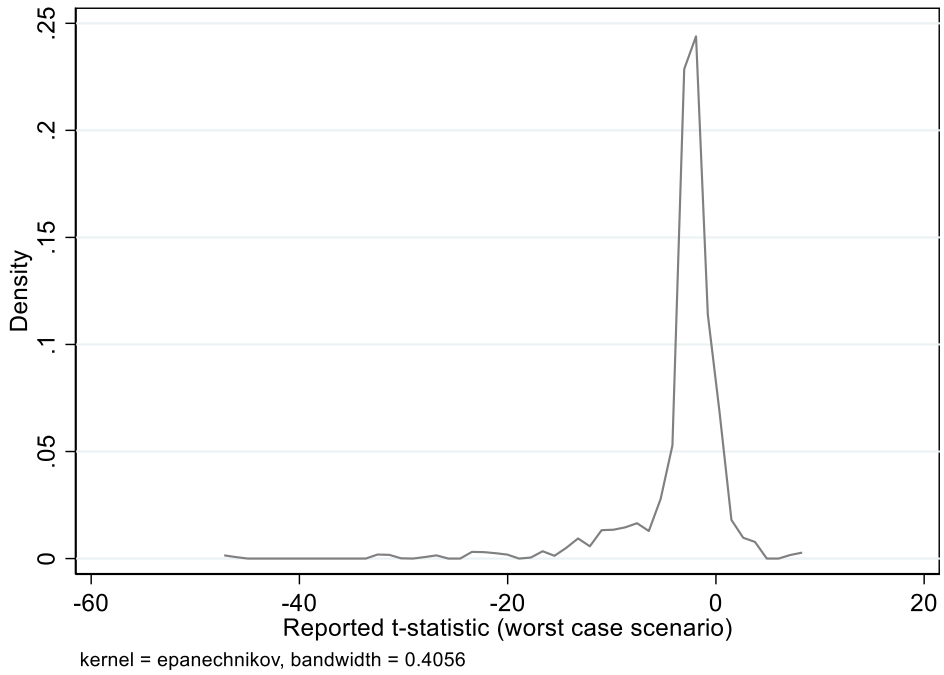
Figure 5: Distributions of *t*-Statistics

Without publication bias, the frequency distributions of Student's *t*-statistics from the 111 individual studies should be approximately normal. Panel A is a histogram of the *t*-statistics reported or estimated. The dotted vertical red line and the solid vertical red line symbolize the critical values of the *t*-statistics (1.96 and 2.58, respectively), associated with a statistical significance of 5% and 1%, respectively. They are both assorted with jumps in the frequency. For the sake of the presentation of these figures, extreme negative values were excluded. Similar results are depicted in Panel B with the Kernel density estimate distribution for the estimated *t*-statistics of the sample of estimates from the 111 individual studies.

Panel A. Frequency Distribution of *t*-Statistics.



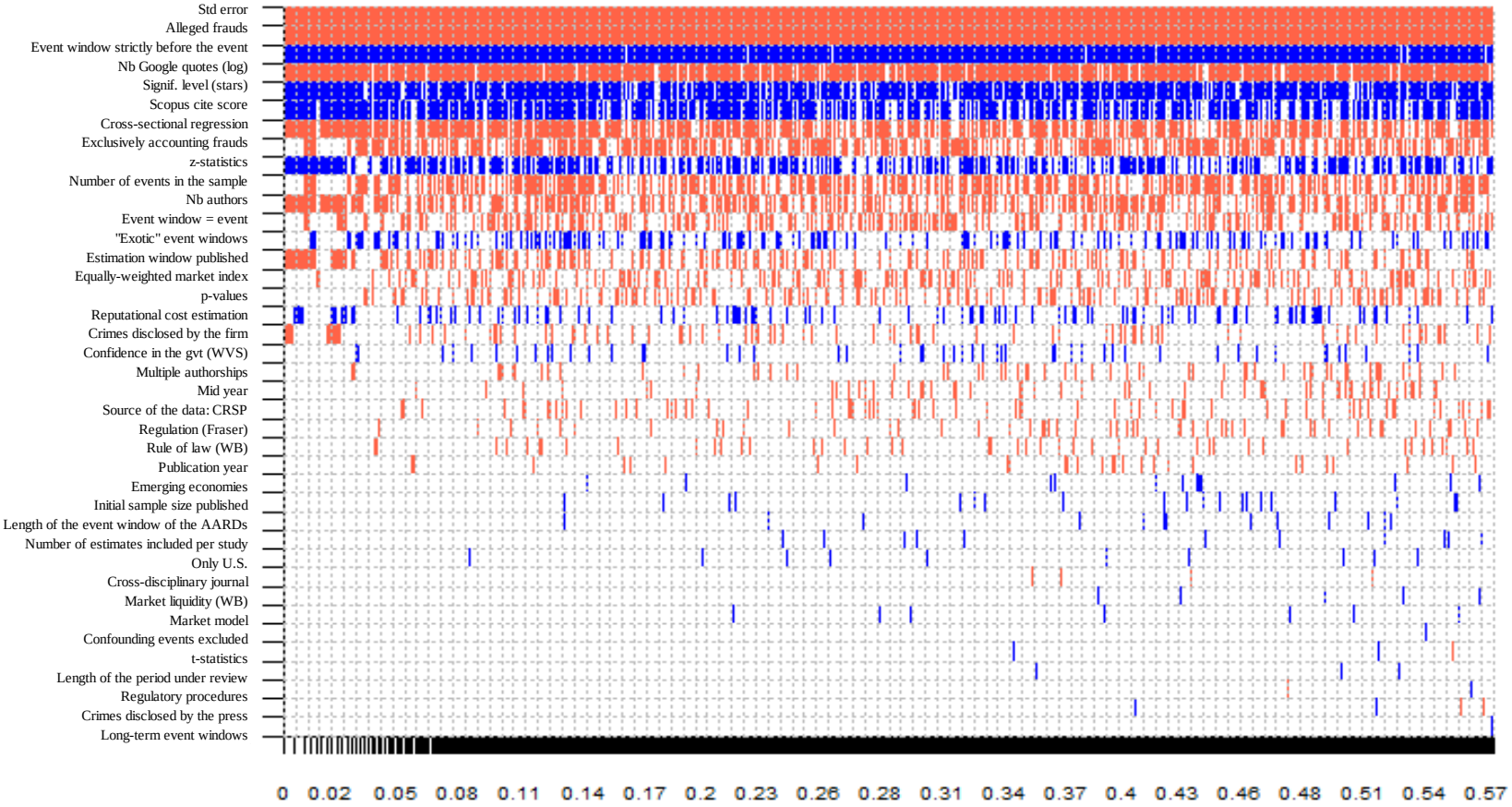
Panel B. Kernel Density Estimate of *t*-Statistics.



Source: Authors

Figure 6: Model Inclusion in Bayesian Model Averaging

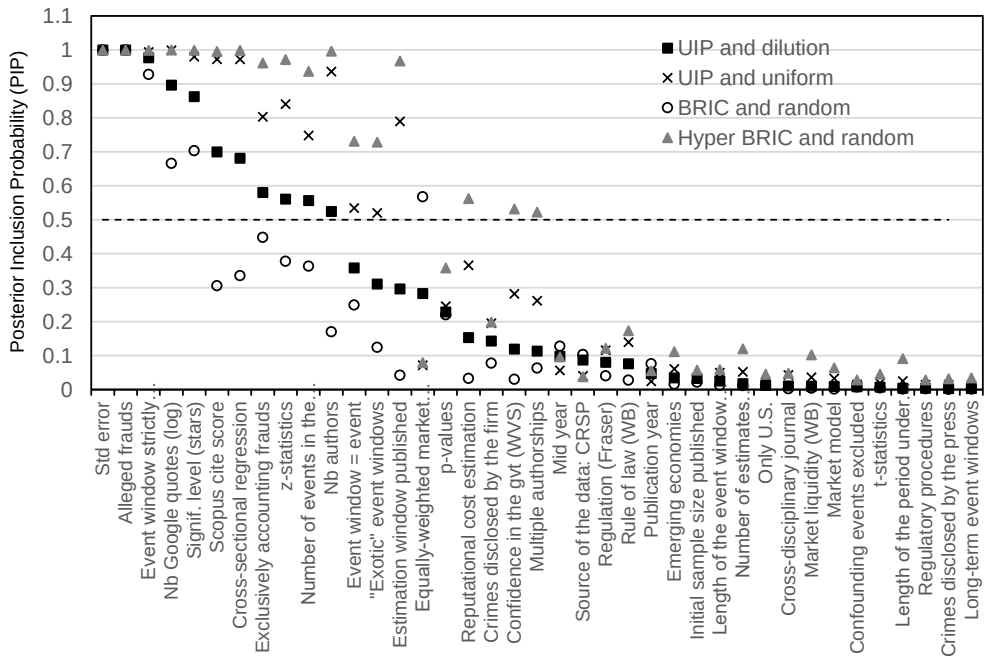
Figure 6 depicts the model inclusion in Bayesian Model Averaging, with the average abnormal returns per day as the response variable. Each column denotes an individual model. Our baseline specification uses the unit information prior recommended by Eicher et al. (2011) and the dilution prior suggested by George (2010), which addresses collinearity. On the vertical axis, the explanatory variables are ranked according to their Posterior Inclusion Probability (PIP) by descending order. A detailed description of all variables is available in Table A.3 of Appendix E. The horizontal axis shows the values of the cumulative posterior model probability for each model, ranked from the highest on the left to the lowest on the right. The blue color (or darker in grayscale) means that the estimated parameter of the explanatory variable is positive. Conversely, the red color (or lighter in grayscale) indicates a negative sign for the estimated parameter. No color denotes that the variable is not included in the model. Numerical results are displayed in Table 4; a sensitivity analysis of the PIP to the prior setting is displayed in Figure 7; and robustness checks are in Appendix G.



Source: Authors

Figure 7: Sensitivity of Posterior Inclusion Probabilities across Prior Settings

Figure 7 compares the Posterior Inclusion Probabilities (PIPs) between different priors. Four combinations are depicted: 1) our baseline specification, as in Bajzík et al. (2020), with a unit information prior for the parameters (UIP) and a dilution model prior for model space (Dilution), adjusting the model probabilities by the determinant of the correlation matrix of the variables included in the model; 2) a unit information prior for the parameters (UIP) and uniform model prior for model space (Uniform), as recommended by Eicher et al. (2011) given the good predictive power of these priors, as done in Havránek and Sokolova (2020); 3) a benchmark g-prior for the parameters (BRIC) and a beta-binomial model prior for the model space (Random), which sets an equal prior probability to each model (Ley and Steel, 2009), as suggested by Fernandez et al. (2001) and Ley and Steel (2009); and 4) a data-dependent hyper-g prior (Hyper BRIC) suggested by Feldkircher (2012) and Feldkircher and Zeugner (2012), which should be less sensitive to the noise in the data, and a beta-binomial model prior for the model space (Random).



Source: Authors

Appendix A: Event Study Methodology (for Online Publication)

Event studies have long been used to challenge the information content of a wide range of corporate news, called "events" (for example Dolley (1933), MacKinlay (1997), and Kothari and Warner (2008)).[32] The goal is to quantify an "abnormal" market reaction following the event by deducing estimated "normal" market parameters from "actual" observed market parameters. A wide range of impact measure variables have been used: returns (the most frequent, on which this work focuses), the bid-ask spread, volatility, turnover, clients, cost of financing (interest rates), financing mix (debt versus equity), top management turnover, analysts' forecasts, etc.

The impact of each event is measured as abnormal returns. For every "event", the abnormality of daily returns is tested over an event window by comparing "actual" ex-post returns with "normal" returns. The latter are the expected returns without conditioning on the event occurring, estimated over an estimation window preceding the event window. The abnormal returns consecutive to a given step of the procedure are taken as unbiased estimates of the total financial consequences of the event.

The finance literature has considered several models of expected returns to describe the behavior of returns and to sort out, to the maximum possible extent, changes in returns caused by the "event" itself from those caused by any other unrelated movement in prices. The event is assumed to be exogenous with respect to the firm. They can be classified as statistical or economic models:

A. Statistical models:

- Constant-mean-return model: $R_{i,t} = \mu_i + \varepsilon_{i,t}$, where $R_{i,t}$ is the returns in t for stock i, μ_i is the mean return of stock i, and $\varepsilon_{i,t}$ is the disturbance term.
- Market model (or single factor market model): $R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$, with $E(\varepsilon_{i,t}) = 0$ and $Var(\varepsilon_{i,t}) = \sigma_\varepsilon^2$, where $R_{i,t}$ and $R_{m,t}$ are the returns in t respectively on the stock i and on the market portfolio. $\varepsilon_{i,t}$ is the zero-mean disturbance term. α_i, β_i, and σ_ε^2 are the firm-specific parameters of the model.
- Factor models: adding other factors than the market trend, for example a sector index (Sharpe, 1970).
- Market-adjusted-return model: restricted market model with $\alpha_i = 0$ and $\beta_i = 1$, when no data is available before the event, for example.

B. Economic models:

- Capital Asset Pricing Model (CAPM): $R_{i,t} = R_f + \beta_i(R_{m,t} - R_f) + \varepsilon_{i,t}$, with $E(\varepsilon_{i,t}) = 0$ and $Var(\varepsilon_{i,t}) = \sigma_\varepsilon^2$, where R_f is the risk-free rate, $R_{i,t}$ and $R_{m,t}$ are the returns in t respectively on the stock i, and on the market portfolio. $\varepsilon_{i,t}$ is the zero-mean disturbance term. β_i, is the beta or systemic risk of stock i.
- Arbitrage Pricing Theory (Fama-French): $R_{i,t} = \delta_0 + \delta_{i,1}F_{1,t} + \delta_{i,2}F_{2,t} + \cdots + \delta_{i,n}F_{n,t} + \varepsilon_{i,t}$, where $F_{i,t}$, $i \in [\![1;n]\!]$, are the n factors that generate returns and $\delta_{i,y}, y \in [\![1;n]\!]$ are the factor loadings.

In the sample of this meta-analysis, by far the most frequently used is the market model. It assumes a stable linear relation between the security return and the market return. It also hypotheses a jointly multivariate normal and temporally independent distribution of returns.

For a firm i, over the period τ, the abnormal returns are:

$$AR_{i,\tau} = R_{i,\tau} - E(R_{i,\tau}/X_\tau)\ . \qquad (I)$$

$AR_{i,\tau}$, $R_{i,\tau}$ and $E(R_{i,\tau}/X_\tau)$ capture abnormal, actual, and expected normal returns, respectively, on the security i over τ, given the conditioning information X_τ for the normal performance model. Equity returns are defined as the daily log difference in the value of the equity.

For every security i of sector s, the market model is in t:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}, \text{ with } E(\varepsilon_{i,t}) = 0 \text{ and } Var(\varepsilon_{i,t}) = \sigma_\varepsilon^2. \qquad (II)$$

$R_{i,t}$ and $R_{m,t}$ are the returns in t on the stock i and on the market portfolio, respectively. $\varepsilon_{i,t}$ is the zero-mean disturbance term. α_i, β_i, and σ_ε^2 are the parameters of the model.

Under general conditions, abnormal returns parameters ($\widehat{\alpha}_i$ and $\widehat{\beta}_i$) are estimated for every event using the selected model over an estimation window preceding the event with Ordinary Least Squares, as recommended

[32] Event studies have been used for decades to assess market reactions to corporate misconduct ranging from product safety and product recalls (airplane crashes, drug recalls, product or automobile recalls, etc.) to any kind of corporate malfeasance (bribery, criminal fraud, tax evasion, illegal political contributions, criminal antitrust violations and price fixing, employee discrimination, environment accidents, environment and wildlife offenses, business ethics, breach of contracts, misleading advertising, etc.) and financial misconduct (insider trading, accounting fraud, option backdating, etc.).

by MacKinlay (1997). On each day t of the event window, the deviation in an individual stock's daily return (typically including reinvested dividends) from what is expected based on Eq. (II) (i.e. the prediction error or "abnormal" returns) is taken as an unbiased estimate of the financial effects of the "event" on stock i in t:

$$AR_{i,t} = R_{i,t} - \widehat{\alpha}_i - \widehat{\beta}_i \, R_{m,t} \; . \qquad (III)$$

$R_{i,t}$ is the actual returns on the security i in t and $AR_{i,t}$ is the estimated abnormal returns for firm i in t. $\widehat{\alpha}_i$, and $\widehat{\beta}_i$ are the estimates of α_i and β_i from Eq. (II) over the estimation window. Abnormal returns over the event window capture the impact of the event on the value of the firm, under the assumption that the event is exogenous with respect to the given security. Abnormal returns are calculated over an event window, including the event day ($t = 0$).

The market-adjusted model merely assumes the following: $AR_{i,t} = R_{i,t} - R_{m,t}$.

The event window can start before the event to investigate for potential anticipation by the market (for example from leaks of information in the days preceding the event). Its length can challenge the persistence over time of the price effect. Under the null hypothesis H_0, the "event" has no impact on the distribution of returns (mean or variance effect). Individual parametric t-statistics are calculated for each firm's abnormal return and for every event day.

Abnormal returns must be aggregated to draw overall inferences for the event of interest, through time and across firms. In fact, on a case-by-case basis, the statistical significance is difficult to detect because of the volatility in firms' stock returns. Hence, abnormal returns are then cumulated over time ($CAR_{i,[t_1;t_2]}$) and averaged across the n victims to get the Cumulative Average Abnormal Returns ($\mathrm{CAAR}_{[t_1;t_2]}$) over the period $[\![t_1; t_2]\!]$, including the event (Eq. (IV)). All events are treated as a group, for which the p-value on the constant of the regression for every period gives the significance of the CAR across all sanctions with robust standard errors.

$$CAAR_{[t_1;t_2]} = \frac{1}{n}\sum_{i=1}^{n} CAR_{i,[t_1;t_2]} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=t_1}^{t_2} AR_{i,t} \; . \qquad (IV)$$

Appendix B: Main Features of Financial Crimes and Enforcement (for Online Publication)

Table A.1: Main Features of Some Securities Enforcers

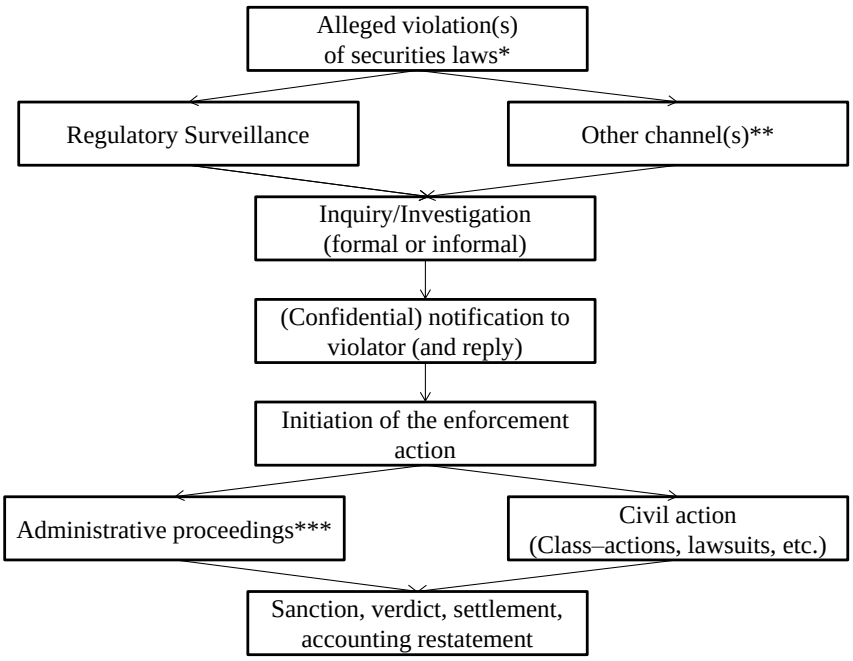
Table A.1 compares the main features of securities law enforcement in the four most frequent countries in the sample: the U.S., China, the UK, and France. Each country has its own enforcement mix, with different weights given to public (higher in code-law countries) or private (higher in common-law countries, typically the U.S.) enforcement and to self-regulation of the market (Djankov et al., 2008). Enforcement can also rely more on informal discussions and administrative guidance (such as in the UK, Japan, and France), or on formal legal actions against wrongdoers (like in the U.S.). Financial regulations can be enforced by either several bodies (at different levels of government such as federal, province, or state levels or depending on the sector with splits between banks, insurance companies, etc.) or one single financial supervisory agency.

	U.S.	**China**	**UK**	**France**
Securities regulator	Securities and Exchange Commission (SEC)	China Securities Regulatory Commission (CSRC)	Financial Conduct Authority (FCA, FSA until 2012)	*Autorité des Marchés Financiers* (AMF since 2003)
Civil actions can be taken by the securities regulator	Yes	No	Yes	Yes
Major types of sanction	Cease and desist orders, suspension or revocation of broker-dealer and investment advisor registrations, censures, bars from association with the securities industry, monetary penalties and disgorgements	Warning, fines, disgorgement of illegal gains, banning of market entry, rectification notice, regulatory concern and letter of warning, public statements and regulatory interview	Variation/cancellation /refusal of authorization/approva l/permissions, financial penalties, public censure, prohibition and suspension	Warning, blame, prohibition and suspension from activity, financial penalties
Most frequent type of sanction	Monetary penalties	Non-monetary penalties	Non-monetary penalties	Monetary penalties
Possibility of class actions	Yes	Yes	No	No
Regulatory communication before sanction	Yes	No	No	No
Settlements	Yes	Yes (mediations)	Yes	Yes (since 2012)
Type of law	Common	Code	Common	Code
Legal origins	English	Socialist	English	French

Source: Authors

Figure A.1: Common features of Financial Crime Prosecution

Figure A.1 presents a simplified view of the consecutive steps of public or private prosecution of financial crimes. Most code-law countries (France, Germany, Italy, Spain, etc.) do not communicate any information before the sanction is pronounced. Conversely, common-law countries, and most frequently in the U.S., enforcers and defendants can communicate through official ways during the procedure. For example, for the U.S., the following steps were investigated by the literature: Accounting and Auditing Enforcement (AAER), SEC formal or informal investigations and sanctions, Wells Notice issuance, sanctions by Department of Justice and Securities Exchange Commission, class action filing, and accounting restatement publications.

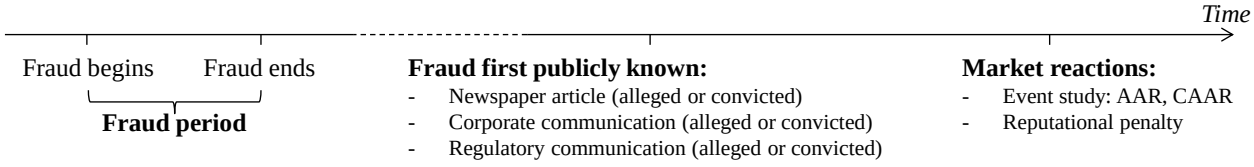


* Securities laws, including enforced accounting standards (U.S. GAAP in the U.S., IFRS, etc.).
** Self-regulatory organizations (stock exchanges, Justice Ministries, etc.), media, external auditors, complaints from shareholders or stakeholders, whistleblowing, etc.
*** Examples of securities law enforcers: Australian ASIC, Canadian OSC, Chinese CRSC, French AMF, German BaFin, U.K. FCA, U.S. SEC, U.S. Department of Justice, U.S. Comptroller of Currency.

Source: Authors

Figure A.2: Chronology of Financial Crimes

This figure shows the typical succession of events that lead to market reactions when learning about a corporate financial crime. The sequence of events is representative for most crimes in the scope of this study but may differ in certain cases.

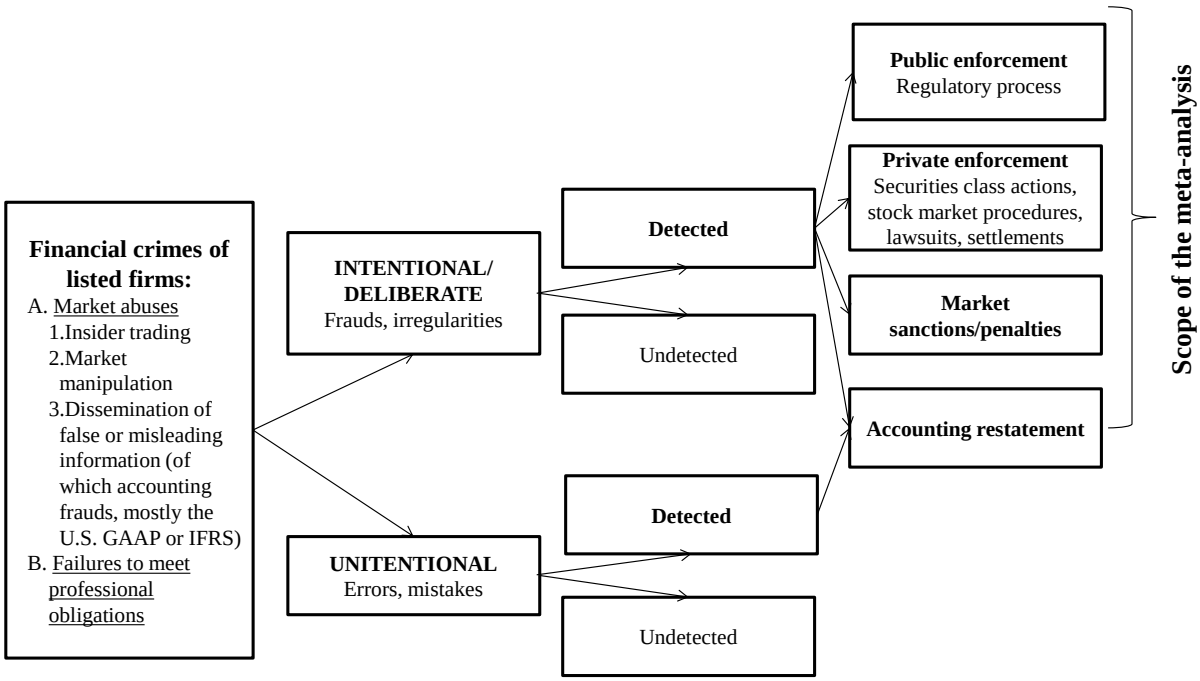


Source: Authors

Appendix C: Scope of the Meta-Analysis (for Online Publication)

Figure A.3: Graphical Presentation of the Scope of the Meta-Analysis

This figure graphically describes the inclusion criteria of the meta-analysis. From a wide range of studies on financial crimes by listed firms, the scope was reduced to the literature investigating detected and intentional crimes and the subsequent market reactions that are based on an event-study methodology. Financial crimes cover the following range of misconduct: market abuses with insider trading (insider dealing, soundings, research), price manipulation (spoofing/layering, new issue/M&A support, ramping, squeeze/corner, bull/bear raids, circular trading,[33] improper order handling,[34] and improper price influence[35]), and the dissemination of false information (collusion and information sharing with pools and information disclosure; misleading customers with guarantees, window dressing, misrepresentation), to which we may add any breach of regulations or professional obligations for listed firms.

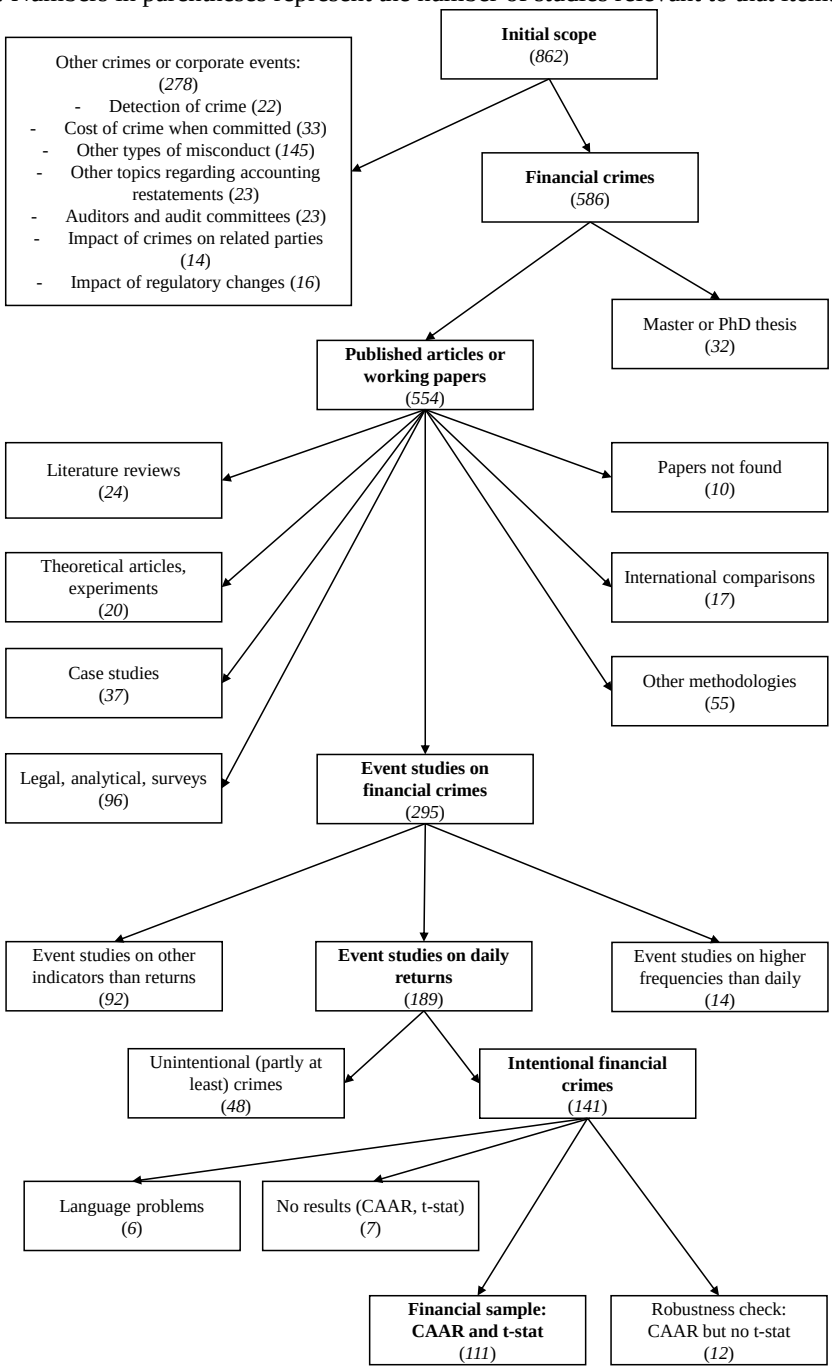


Source: Authors

[33] Wash trades, matched trades, money pass and compensation trades, and parking/warehousing.
[34] Front running, cherry picking and partial fills, and stop losses and limits.
[35] Benchmarks, closing prices, reference prices, portfolio trades, and barriers.

Figure A.4: PRISMA Statement

The following PRISMA flow diagram shows the details of the information flow in each stage of the literature search in our meta-analysis, as recommended by Moher et al. (2009) and Havránek et al. (2020). From an initial sample of 862 studies reviewed, we end up with a sample of 111 articles to which we add 12 more articles for robustness checks for which no details were given on statistical significance. The details of each category are available upon request. Bold titles illustrate how we ended with the final sample. This graphical illustration has its limit as many studies cumulated reasons for being excluded but, for the sake of presentation, they were allocated into one category. Numbers in parentheses represent the number of studies relevant to that item.



Source: Authors

Appendix D: Studies Subject to Meta-Analysis – Key Facts (for Online Publication)

Table A.2: The Meta Dataset

Table A.2 describes the main features of the studies included in the meta-analysis. Financial crimes are sorted into three categories: exclusively accounting fraud, regulatory securities fraud (excluding accounting fraud), and all regulatory securities fraud (including accounting fraud). The country codes are the following, by alphabetical order: AU-Australia, BE-Belgium, CA-Canada, CN-China, DE-Germany, ES-Spain, FR-France, JP-Japan, KR-South Korea, LU-Luxembourg, MY-Malaysia, NL-Netherlands, SW-Sweden, TH-Thailand, TR-Turkey, UK-United Kingdom, US-United States of America. The "sample size" variable is the number of financial crimes that were included in the event study to assess the size effect on returns. The variable "AAR per day" is the average of all abnormal returns published, whatever the event window (between -10 to +10 days around the event day), divided by the number of days in the event window. The average AAR per day is weighted by the number of estimates per study. The variable "Stat. signif." is a dummy variable for statistically significant abnormal returns following the financial crimes. Finally, the variable "Nb. est." stands for the number of estimates included in the dataset per study.

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/day	Std. dev.	Stat. signif.	Nb. est.
Abdulmanova, Ferris, Jayaraman, Kothari	2019	WP	Regulatory securities fraud	US	2004	2013	462	-0.7%	0.0%	yes	2
Aggarwal, Hu, Yang	2015	Journal of Portfolio Management	Regulatory securities fraud (incl. accounting fraud)	CN	2001	2011	750	-0.6%	0.4%	yes/no	5
Agrawal, Chadha	2005	Journal of Law and Economics	Accounting fraud	US	2000	2001	119	-2.0%	0.2%	yes/no	2
Agrawal, Cooper	2017	Quarterly Journal of Finance	Accounting fraud	US	1997	2002	419	-2.1%	1.2%	yes	3
Akhigbe, Kudla, Madura	2005	Applied Financial Economics	Accounting fraud	US	1991	2001	77	-3.1%	.	yes	1
Amoah	2013	Advances in Public Interest Accounting	Regulatory securities fraud	US	1996	2006	301	-7.7%	0.1%	yes	2
Amoah, Tang	2010	Advances in Accounting	Accounting fraud	US	1997	2002	143	-1.8%	1.3%	yes/no	2
Andersen, Gilbert, Tourani-Rad	2013	JASSA	Regulatory securities fraud	AU	2004	2012	18	-1.1%	1.3%	yes	7
Anderson, Yohn	2002	WP	Accounting fraud	US	1997	1999	4	-2.2%	.	yes	1
Armour, Mayer, Polo	2017	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	UK	2001	2011	40	-0.8%	0.4%	yes	3
Arnold, Engelen	2007	Management & Marketing	Regulatory securities fraud (incl. accounting fraud)	BE, NL	1994	2003	57	-0.9%	0.7%	yes/no	6
Baker, Edelman, Powell	1999	Business and Professional Ethics Journal	Regulatory securities fraud	US	1991	1996	14	-0.2%	0.5%	yes/no	8
Barabanov, Ozocak, Turtle, Walker	2008	Financial Management	Regulatory securities fraud	US	1996	2003	623	-1.6%	.	yes	1
Bardos, Golec, Harding	2013	Journal of Financial Research	Accounting fraud	US	1997	2002	166	-10.3%	.	yes	1
Bardos, Mishra	2014	Applied Financial Economics	Accounting fraud	US	1997	2002	24	-5.5%	4.9%	yes	2
Barniv, Cao	2009	Journal of Accounting and Public Policy	Accounting fraud	US	1995	2003	61	-6.8%	.	yes	1
Bauer, Braun	2010	Financial Analytical Journal	Regulatory securities fraud (incl. accounting fraud)	US	1996	2007	648	-1.1%	0.6%	yes	15
Beasley, Carcello, Hermanson, Neal	2010	COSO	Accounting fraud	US	1998	2007	213	-4.8%	3.4%	yes/no	8

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/day		Stat. signif.	Nb. est.
Beneish	1999	The Accounting Review	Accounting fraud	US	1987	1993	50	-4.2%	2.4%	yes	3
Bhagat, Bizjak, Coles	1998	Financial Management	Regulatory securities fraud	US	1981	1983	46	-1.4%	.	yes	1
Billings, Klein, Zur	2012	WP	Regulatory securities fraud (incl. accounting fraud)	US	1996	2008	408	-0.3%	0.1%	yes	3
Bohn, Choi	1996	University of Pennsylvania Law Review	Regulatory securities fraud	US	1975	1986	103	-1.2%	0.1%	yes	2
Bonini, Boraschi	2010	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	US	1996	2005	686	-1.8%	0.9%	yes	11
Bowen, Call, Rajgopal	2010	The Accounting Review	Regulatory securities fraud (incl. accounting fraud)	US	1989	1996	78	-0.6%	.	yes	1
Bradley, Cline, Lian	2014	Journal of Corporate Finance	Regulatory securities fraud (incl. accounting fraud)	US	1996	2011	1530	-0.6%	.	yes	1
Brous, Leggett	1996	Journal of Financial Research	Regulatory securities fraud (incl. accounting fraud)	US	1989	1991	62	-0.6%	3.4%	yes/no	3
Burns, Khedia	2006	Journal of Financial Economics	Accounting fraud	US	1997	2001	215	-2.0%	1.5%	yes	4
Callen, Livnat, Segal	2006	Journal of Investing	Accounting fraud	US	1986	2001	385	-2.8%	.	yes	1
Chava, Cheng, Huang, Lobo	2010	International Journal of Law and Management	Regulatory securities fraud	US	1995	2004	85	-3.1%	.	yes	1
Chen, Firth, Gao, Rui	2005	Journal of Accounting and Public Policy	Regulatory securities fraud	CN	1999	2003	169	-0.2%	0.3%	yes/no	10
Choi, Karpoff, Lou, Martin	2019	WP	Regulatory securities fraud (incl. accounting fraud)	US	1978	2015	942	-14.5%	.	yes	1
Choi, Pritchard	2016	Journal of Legal Studies	Regulatory securities fraud	US	2004	2007	231	-6.5%	5.3%	yes	3
Christensen, Paik, Williams	2010	Journal of Forensic & Investigative Accounting	Regulatory securities fraud (incl. accounting fraud)	US	2001	2003	151	-2.1%	3.9%	yes/no	6
Cook, Grove	2009	Journal of Forensic & Investigative Accounting	Regulatory securities fraud (incl. accounting fraud)	US	1984	2005	88	-3.6%	2.3%	yes	14
Correia, Klausner	2012	WP	Accounting fraud	US	2000	2011	683	-5.0%	2.0%	yes	2
Cox, Weirich	2002	Managerial Auditing Journal	Accounting fraud	US	1992	1999	27	-4.2%	0.8%	yes	3
Davidson, Worrell, Lee	1994	Journal of Business Ethics	Accounting fraud	US	1965	1990	34	-0.6%	0.9%	yes/no	16
Davis, Taghipour, Walker	2017	Managerial Finance	Regulatory securities fraud	US	1996	2013	2153	0.3%	2.4%	yes	2
de Batz	2020	European Journal of Law and Economics	Regulatory securities fraud (incl. accounting fraud)	FR	2004	2016	52	-0.3%	0.1%	yes/no	10
Dechow, Sloane, Sweeney	1996	Contemporary Accounting Research	Accounting fraud	US	1982	1992	78	-8.8%	.	yes	1
Deng, Willis, Xu	2014	Journal of Financial and Quantitative Analysis	Regulatory securities fraud (incl. accounting fraud)	US	1996	2006	156	-1.7%	1.7%	yes	6
Desai, Hogan, Wilkins	2006	The Accounting Review	Accounting fraud	US	1997	1998	146	-3.7%	.	yes	1
Djama	2013	Revue Française de Gestion	Accounting fraud	FR	1995	2008	36	-0.9%	0.0%	yes/no	3
Du	2017	Journal of Business Finance & Accounting	Accounting fraud	US	2001	2011	17	-2.3%	0.6%	yes	2

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/day		Stat. signif.	Nb. est.
Engelen	2009	WP	Regulatory securities fraud	BE, DE, FR, LU, NL, UK	1995	2005	83	-0.8%	0.7%	yes/no	12
Engelen	2011	Book chapter	Regulatory securities fraud	BE, DE, FR, LU, NL, UK	1995	2005	101	-0.7%	1.0%	yes/no	6
Engelen	2012	CESifo Economic Studies	Regulatory securities fraud	US	1993	2008	122	-0.5%	0.9%	yes/no	3
Eryiğit	2019	Journal of Financial Crime	Accounting fraud	TR	2005	2015	160	-0.1%	0.0%	yes/no	4
Ewelt-Knauer, Knauer, Lachmann	2015	Journal of Business Economics	Regulatory securities fraud	DE	1998	2014	126	-2.3%	2.5%	yes	2
Feroz, Park, Pastena	1991	Journal of accounting research	Accounting fraud	US	1982	1989	58	-2.6%	3.1%	yes/no	11
Ferris, Jandik, Lawless, Makhija	2007	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	US	1982	1999	194	-0.6%	.	yes	1
Fich, Shivdasani	2007	Journal of Financial Economics	Regulatory securities fraud	US	1998	2002	200	-3.5%	0.5%	yes	4
Firth, Rui, Wu	2009	Journal of Accounting and Public Policy	Regulatory securities fraud	CN	1999	2005	61	-0.7%	0.7%	yes/no	10
Firth, Rui, Wu	2011	Journal of Corporate Finance	Accounting fraud	CN	2000	2005	267	-0.1%	0.1%	yes/no	8
Firth, Wong, Xin, Yick	2016	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	CN	2003	2010	75	-0.2%	0.1%	yes	2
Flore, Degryse, Kolaric, Schiereck	2018	WP	Regulatory securities fraud (incl. accounting fraud)	DE, ES, FR, NL, SW, UK, US	2005	2015	251	0.1%	0.1%	yes/no	5
Gande, Lewis	2009	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	US	1996	2003	605	-1.2%	0.7%	yes/no	7
Gerety, Lehn	1997	Managerial and Decision Economics	Accounting fraud	US	1981	1987	37	-1.0%	.	yes	1
Goldman, Peyer, Stefanescu	2012	Financial Management	Accounting fraud	US	1976	2010	444	-8.3%	3.7%	yes	5
Griffin, Grundfest, Perino	2004	Abacus	Regulatory securities fraud	US	1990	2002	2133	-1.8%	2.5%	yes/no	4
Griffin, Sun	2016	Accounting and Finance Research	Regulatory securities fraud	US	2001	2007	80	-0.8%	0.5%	yes/no	4
Haslem, Hutton, Hoffmann Smith	2017	Financial Management	Regulatory securities fraud	US	1995	2006	594	-0.8%	0.5%	yes	6
Hirschey, Palmrose, Scholz	2005	WP	Accounting fraud	US	1995	1999	405	-7.3%	.	yes	1
Humphery-Jenner	2012	Journal of Financial Intermediation	Regulatory securities fraud	US	1996	2007	416	-1.1%	0.4%	yes	5
Iqbal, Shetty, Wang	2007	Journal of Financial Research	Regulatory securities fraud	US	1996	2003	298	-5.0%	4.4%	yes	10
Johnson, Ryan, Tian	2003	WP	Accounting fraud	US	1992	2005	87	-4.9%	.	yes	1
Jordan, Peek, Rosengren	2000	Journal of Financial Intermediation	Regulatory securities fraud (incl. accounting fraud)	US	1989	1994	35	-1.7%	.	yes	1
Karpoff, Koester, Lee, Martin	2017	The Accounting Review	Accounting fraud	US	1978	2011	1052	-14.5%	.	yes	1
Karpoff, Lee, Martin	2008	Journal of financial and quantitative analysis	Accounting fraud	US	1978	2002	371	-9.4%	4.5%	yes	6
Karpoff, Lott	1993	Journal of Law and Economics	Accounting fraud	US	1978	1987	4	-1.3%	1.1%	yes/no	5

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/day		Stat. signif.	Nb. est.
Kellogg	1984	Journal of Accounting and Economics	Accounting fraud	US	1967	1979	26	-1.9%	2.8%	yes/no	3
Kirat, Rezaee	2019	Applied Economics	Regulatory securities fraud (incl. accounting fraud)	FR	2004	2017	54	0.0%	0.6%	yes	7
Klock	2015	Journal of Business & Securities Law	Regulatory securities fraud (incl. accounting fraud)	US	1996	2012	714	-1.0%	0.3%	yes	4
Kouwenberg, Phunnarungsi	2013	Pacific-Basin Finance Journal	Regulatory securities fraud (incl. accounting fraud)	TH	2003	2010	111	-0.7%	0.9%	yes/no	4
Kravet, Shevlin	2010	Review of Accounting Studies	Accounting fraud	US	1997	2001	299	-0.8%	.	yes	1
Kryzanowski, Zhang	2013	Journal of Multinational Financial Management	Accounting fraud	CA	1997	2006	210	-1.9%	0.5%	yes	4
Kwan, Kwan	2011	International Review of Business Research Papers	Regulatory securities fraud	MY	2005	2009	41	0.3%	1.5%	yes/no	3
Lei, Law	2019	WP	Regulatory securities fraud (incl. accounting fraud)	CN	1999	2015	1188	-0.1%	0.2%	yes/no	8
Liebman, Milhaupt	2008	Columbia Law Review	Regulatory securities fraud	CN	2001	2006	68	-0.7%	0.4%	yes/no	8
Lieser, Kolaric	2016	WP	Regulatory securities fraud (incl. accounting fraud)	US	1996	2014	1377	-1.3%	2.2%	yes/no	15
Loh, Rathinasamy	2003	Review of Pacific Basin Financial Markets and Policies	Regulatory securities fraud (incl. accounting fraud)	US	1996	1998	290	-0.5%	0.6%	yes	2
Marciukaityte, Szewczyk, Uzun, Varma	2006	Financial Analysts Journal	Regulatory securities fraud (incl. accounting fraud)	US	1978	2001	28	-3.9%	.	yes	1
Marciukaityte, Szewczyk, Varma	2009	Financial Analysts Journal	Accounting fraud	US	1997	2002	187	-3.3%	.	yes	1
McDowell	2005	WP	Accounting fraud	US	1998	2003	174	-2.1%	.	yes	1
Muradoglu, Clark Huskey	2008	WP	Regulatory securities fraud (incl. accounting fraud)	US	1995	2004	296	-0.6%	0.4%	yes/no	12
Nainar, Rai, Tartaroglu	2014	International Journal of Disclosure and Governance	Regulatory securities fraud	US	1999	2007	77	-1.1%	0.4%	yes/no	6
Nelson, Gilley, Trombley	2009	Securities Litigation Journal	Regulatory securities fraud	US	2002	2007	58	-2.6%	.	yes	1
Nourayi	1994	Journal of Accounting and Public Policy	Regulatory securities fraud (incl. accounting fraud)	US	1977	1984	82	-0.2%	0.1%	yes	4
Owers, Lin, Rogers	2002	International Business and Economics Research Journal	Accounting fraud	US	1994	1997	13	-2.5%	2.6%	yes	6
Ozbas	2008	WP	Accounting fraud	US	1999	2003	75	-2.5%	2.2%	yes/no	4
Ozeki	2019	Securities Analysts Journal	Accounting fraud	JP	2005	2016	218	-9.1%	1.8%	yes/no	2
Pereira, Malafronte, Sorwar, Nurullah	2019	Journal of Financial Services Research	Regulatory securities fraud (incl. accounting fraud)	US	2004	2015	1387	-8.4%	6.1%	yes/no	5
Persons	1997	Journal of Business Research	Regulatory securities fraud	US	1972	1993	95	-0.3%	0.2%	yes	4
Plumlee, Yohn	2008	WP	Accounting fraud	US	2003	2006	1303	-0.3%	.	yes	1
Pritchard, Ferris	2001	WP	Regulatory securities fraud	US	1995	1999	89	-3.1%	4.6%	yes/no	3
Romano	1991	Journal of Law, Economics, and Organization	Regulatory securities fraud	US	1970	1987	66	-0.8%	1.6%	yes/no	6

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/day		Stat. signif.	Nb. est.
Scholz	2008	US Department of Treasury	Accounting fraud	US	1997	2006	264	-6.5%	.	yes	1
Slovin, Sushka, Polonchek	1999	Journal of Financial Economics	Regulatory securities fraud (incl. accounting fraud)	US	1975	1992	61	-1.8%	1.1%	yes	2
Song, Han	2017	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	KR	2001	2010	220	-0.7%	0.2%	yes	3
Sun, Zhang	2006	WP	Regulatory securities fraud	CN	1990	2002	144	-0.5%	.	yes	1
Takmaz, Keles	2017	Journal of Business Research Turk	Regulatory securities fraud	TR	2007	2016	72	-0.2%	0.4%	yes/no	4
Tanimura, Okamoto	2013	Asian Economic Journal	Accounting fraud	JP	2000	2008	39	-3.1%	.	yes	1
Tay, Puah, Brahmana, Abdul Malek	2016	Journal of Financial Crime	Regulatory securities fraud (incl. accounting fraud)	MY	1996	2013	17	-0.5%	0.0%	no	3
Wang, Ashton, Jaafar	2019	The British Accounting Review	Accounting fraud	CN	2007	2016	433	-0.1%	0.0%	yes/no	7
Wang, Wu	2011	China Journal of Accounting Research	Accounting fraud	CN	1999	2005	67	-0.1%	0.2%	yes/no	5
Wu	2002	WP	Accounting fraud	US	1977	2000	932	-7.7%	.	yes	1
Wu, Zhang	2014	China Journal of Accounting Studies	Regulatory securities fraud	CN	2002	2011	157	-2.1%	1.4%	yes	6
Xu, Xu	2020	International Review of Law and Economics	Regulatory securities fraud (incl. accounting fraud)	CN	2014	2018	107	-0.7%	2.1%	yes/no	10
Yu, Zhang, Zheng	2015	Financial Management	Accounting fraud	CN	1999	2011	195	-0.6%	0.7%	yes	2
Zeidan	2013	Journal of Business Ethics	Regulatory securities fraud	US	1990	2009	163	-0.4%	0.5%	yes/no	4
Zhu, Hu	2010	WP	Accounting fraud	CN	2006	2008	88	-0.7%	0.4%	yes/no	7
Overall	**2009**				**1994**	**2004**	**293**	**-1.8%***			**4.3**

Source: Authors

Appendix E. Variable Definitions and Descriptive Statistics (for Online Publication)

Table A.3: Variable Definitions and Descriptive Statistics

Table A.3 describes most of the variables for the full sample of financial crimes (479 estimates from 111 studies). Simple means are compared with weighted means, using the inverse of the number of estimates per study. Simple means are calculated for the sample of articles. In fact, on average, four estimates are reported by each study. Some categories are not mutually exclusive.

Variables	Description	Mean	Std dev.	Weighted mean
Effect: AAR per day (AARD)	Average abnormal returns per day (of the event window), equal to the reported average abnormal returns (for one-day event windows) or to the reported cumulative average abnormal returns, estimated using a daily event study methodology, divided by the number of days in the event window.	-1.82%	2.79%	-2.44%
Standard Error	Reported standard error of the estimated abnormal returns, or conservative standard errors, estimated with the conservative (worst case scenario) *t*-statistics.	0.81%	1.30%	1.03%
1. Data characteristics				
Geographical scope:	1 if only one country in the scope.	0.95	0.23	0.97
	1 if the estimate's sample is the U.S. (i.e. most frequent country in the sample).	0.63	0.48	0.70
	1 if the estimate's sample is Asia.	0.23	0.42	0.19
	1 if the estimate's sample is Europe.	0.12	0.33	0.10
	1 if the estimate's sample is China (i.e. the 2nd most frequent country in the sample).	0.19	0.39	0.13
	1 if the estimate's sample is emerging economies (China, Malaysia, South Korea, Thailand, Turkey).	0.23	0.42	0.18
	1 if the legal origin of the commercial law of a country is English common law (Australia, Canada, Malaysia, Thailand, U.K., U.S.), and zero otherwise, considering the geographic distribution of the sample, as in Djankov et al. (2008).	0.68	0.47	0.76
Period under review:	Beginning of period under review.	1994	10.24	1994
	End of period under review.	2005	8.11	2004
	Average year of the period under review.	2000	8.67	1999
	Mid-point, as the logarithm of the mean year of the data used, minus the earliest mean year in the data.	3.23	0.51	3.22
	Length of the period under review.	11.39	6.26	10.77
Event types:				
Types of regulatory breach:	1 if the scope of crimes is limited to exclusively accounting fraud.	0.31	0.46	0.41
	1 if the scope of crimes is limited to any violation of securities laws.	0.41	0.49	0.35
	1 if the scope of crimes covers all violations of securities laws (incl. accounting fraud).	0.28	0.45	0.24
Source of the news/origin of the data under review:	1 if the crimes were disclosed in the press (typically WSJ in the U.S.).	0.42	0.49	0.37
	1 if the crimes were disclosed by regulatory communication.	0.67	0.47	0.67
	1 if the crimes were disclosed by corporate communication.	0.25	0.44	0.30
Steps of enforcement procedure:	1 if the fraud is alleged (not convicted).	0.61	0.49	0.58
	1 if the crimes were being investigated.	0.11	0.31	0.08
	1 if the crimes went through settlement.	0.04	0.19	0.03
	1 if the crimes led to an accounting restatement.	0.13	0.34	0.22
	1 if the crimes were convicted by an authority/court (verdict of regulatory procedures, verdict of lawsuits or class-actions, accounting restatement).	0.41	0.49	0.45
Types of enforcement procedure:	1 if the crimes led to a regulatory procedure.	0.53	0.50	0.53
	1 if the crimes led to a stock exchange procedure.	0.09	0.28	0.08
	1 if the crimes led to a class-action.	0.24	0.43	0.22
	1 if the crimes led to a private lawsuit.	0.10	0.30	0.10
2. Estimation characteristics				
Model:	1 if market model used to estimate abnormal returns (not Fama-French models, CAPM or market-adjusted model).	0.83	0.38	0.84
	1 if equally weighted market index.	0.50	0.50	0.54
	1 if CRSP dataset used for returns (Center for Research in Securities Prices).	0.44	0.50	0.55
	Number of estimates reported per study, to avoid unintentional weighting of articles reporting multiple estimates as recommended by Havránek and Irsova (2017). We used the raw number of estimates, as most of the articles in the sample did not include the estimate's variances.	7.18	4.19	4.33

Variables	Description	Mean	Std dev.	Weighted mean
Sample characteristics:	1 if the initial sample size is published.	0.78	0.42	0.74
	1 if confounding events are explicitly excluded from the final sample.	0.29	0.46	0.24
	Final number of events (financial crimes) in the sample (in particular excluding confounding events and events with data problems).	264	378	261
	Number of observations, as the logarithm of the final number of events in the sample.	4.82	1.25	4.80
Estimation window:	1 if estimation window specified in the study.	0.72	0.45	0.64
	Beginning of the estimation window (in days, relative to the event in t = 0).	-154	129	-153
	End of the estimation window (in days, relative to the event in t = 0).	-20	30	-21
Event window:	Beginning of the event window (in days, relative to the event in t = 0).	-15	35	-17
	End of the estimation window (in days, relative to the event in t = 0).	18	39	19
	Length of the event window (in days).	32.80	60.48	36.50
	1 if event windows beyond $[\![-10;+10]\!]$.	0.28	0.45	0.27
Event window of the estimates:	Beginning of the event window of the estimate (in days, relative to the event in t = 0).	-1.6	3.1	-1.2
	End of the estimation window of the estimate (in days, relative to the event in t = 0).	1.4	2.8	1.3
	Length of the event window of the estimate (in days).	4.0	4.5	3.6
	1 if the event window is strictly before the event date (t = 0).	0.14	0.35	0.09
	1 if the event window is limited to the event date (t = 0).	0.18	0.38	0.17
	1 if the event window is around the event date (t = 0).	0.56	0.50	0.65
	1 if the event window is strictly after the event date (t = 0).	0.12	0.33	0.09
	1 if "exotic" event window (standing for less than 5% of the compounded event windows, i.e. less than 24 estimates)	0.26	0.44	0.22
Statistical significance:	Conservative (worst case scenario) *t*-statistics, invert of the absolute value: *t*-stat when available; when the *t*s were not published, they were obtained as follows (Frooman, 1997): 1) the statistical significance levels were converted into conservative levels of significance; 2) the *z*s were directly changed into *t*s on the assumption that as sample size increases, Student's t distribution approaches the normal distribution (Marascuilo and Serlin, 1988); and 3) the *p* values were converted into ts by using a *t* table and the appropriate degrees of freedom. Finally, three studies (Desai et al. (2006), Nelson et al. (2009), and Goldman et al. (2012)), including seven estimates, stated that the abnormal returns are significant but without including *t*-statistics nor the statistical significance. We made the conservative hypothesis that the statistical significance level was 10% for each.	-3.27	4.85	-3.42
	1 if abnormal returns are significant.	0.72	0.45	0.80
	1 if statistical significance level ("stars").	0.80	0.40	0.81
	1 if *t*-statistics.	0.55	0.50	0.47
	1 if *p*-value.	0.41	0.49	0.33
	1 if *z*-statistics (Patel, Corrado, rank test).	0.17	0.37	0.18
	1 if other statistical significance indicators (Boehmer et al. (1991), Kolari and Pynnonen (2010)).	0.03	0.17	0.04
Complementary results:	1 if complementary cross-sectional regression for the determinants of the stock market reaction to the event (i.e. between the estimated abnormal returns and the characteristics specific to the event, sample, etc.)	0.62	0.49	0.66
	1 if additional estimates of reputational penalties.	0.14	0.34	0.12
3. <u>Publication characteristics</u>				
Characteristics of the article:	Number of authors of the paper.	2.32	0.86	2.36
	1 if multiple authorships in the sample.	0.29	0.46	0.33
	Year of publication.	2009	7.77	2009
	Publication year, as the logarithm of the year of publication.	3.20	0.45	3.20
	Month of publication (1 to 12).	5.61	3.88	5.60
Journal fields:	1 if financial journal.	0.36	0.48	0.37
	1 if accounting journal.	0.31	0.46	0.41
	1 if law journal.	0.24	0.43	0.23
	1 if business journal.	0.23	0.43	0.23
Quality of the publication:	1 if published in a cross-disciplinary journal. As stated in Amiram et al. (2018), studies on financial misconduct belongs to three perspectives: law, accounting, and finance.	0.90	0.30	0.87
	1 if published in a refereed journal or chapter in a book. We expect published studies to exhibit higher quality on average and to contain fewer mistakes in reporting their results. Still, the inclusion of unpublished papers is unlikely to alleviate publication bias (Rusnák et al., 2013): researchers write their papers with the intention to publish. Otherwise, the article is a working paper.	0.81	0.39	0.81
	1 if published in a Scopus journal.	0.63	0.48	0.67

Variables	Description	Mean	Std dev.	Weighted mean
	Number of citations in Google Scholar (as number).	135	352	180
	Citations, as the logarithm of the number of per-year citations of the study since its first appearance on Google Scholar.	1.39	1.19	1.60
	Scopus Cite Score in 2018.	1.65	1.89	1.64
	Scopus Cite Score of the year of publication (2011 to 2018, otherwise 2011).	1.11	1.32	1.17
	IDEAS/RePec Recursive Discounted impact factor.	0.40	1.11	0.39
4. <u>Control variables (on the mid-period under review)</u>				
Economic development index:	Log of nominal current USD GDP (source: World Bank, data available from 1960 to 2018), as in Jackson and Roe (2009).	8.46	0.99	8.55
	Log of GNI per capita in USD (source: World Bank, data available from 1960 to 2018), as in Hubler et al. (2019) and Jackson and Roe (2009).	9.75	1.16	9.90
	Log of GDP per capita (source: World Bank, data available from 1960 to 2018), as in Djankov et al. (2008).	2.79	1.54	2.95
Financial market indicators:	Domestic market capitalization, as % of GDP (source: World Bank, data available from 1975 to 2018), as in Djankov et al. (2008) and Hubler et al. (2019).	93.80	41.39	97.85
	Market liquidity indicator (stocks traded, turnover ratio of domestic shares (%) (source: World Bank, data available from 1975 to 2018), as in La Porta et al. (2006).*	115.37	48.25	114.01
	Log of the average number of domestic listed firms to its population in millions, (source: World Bank, data available from 1975 to 2018), as in Djankov et al. (2008).	2.44	1.62	2.64
	Total value of stock traded, as % of GDP (current USD, source: World Bank, data available from 1975 to 2018).	0.81	0.86	0.90
	Domestic credit provided by financial sector (% of GDP) (source: World Bank, data available from 1960 to 2018).	102.48	33.05	97.93
Legal environment:	Rule of law (mid-period under review, or previous year if not published, or 1996 if before) (source World Bank, data available from 1996 to 2018), supported by the conclusion of La Porta et al. (2006) that financial markets do not prosper when left to market forces alone.	1.11	0.82	1.21
	Regulation sub-index of the economic freedom indicator from the Fraser Institute (mid-period under review, or the closest-available, or average year when not available), with data available from 1970 to 2017, as in Hubler et al. (2019).*	7.70	1.20	7.87
	Credit market regulation sub-index of the economic freedom indicator from the Fraser Institute (mid-period under review, or the closest-available or average year when not available), with data available from 1970 to 2017, as in Hubler et al. (2019).*	8.95	1.17	9.10
	Most people can be trusted (source: World Value Survey).*	39.79	10.06	38.98
	Quite a lot of confidence in the government (source: World Value Survey).*	32.84	10.20	31.32
	Quite a lot of confidence in the judicial system/courts (source: World Value Survey).*	41.45	8.87	40.64
Main sectors:	1 if specified that the most frequent sector involved in financial misconduct is industry.	0.33	0.47	0.33
	1 if specified that the most frequent sector involved in financial misconduct is finance. In some articles (Bonini and Boraschi (2010); Ozeki (2019)), financial firms were excluded from the sample.	0.17	0.38	0.18

*Sources: Studies, World Bank, Fraser Institute, Authors' calculations * Normalized in the BMA and FMA.*

[1] In some studies, no split was made between alleged and convicted financial crimes. All crimes were treated jointly. Consequently, the sum of the two variables exceeds one.

[2] Private enforcement is defined as the combination of the following types of procedure: private lawsuits, stock exchange procedures, and class actions.

[3] The Fraser economic freedom index measures the degree of economic freedom present in five major areas (with 26 components): size of government, legal system and security of property rights, sound money, freedom to trade internationally, and regulation. Each component and sub-component is placed on a scale from 0 to 10 that reflects the distribution of the underlying data.

[4] In two articles (Bauer and Braun (2010) and Ozeki (2019)), financial firms were excluded from the sample.

Appendix F. Additional Funnel Plots (for Online Publication)

Figure A.5: Funnel Plot for the U.S. versus other Countries

Figure A.5 is a funnel plot specifically for the U.S. or other countries depending on whether the financial crime is alleged or convicted.

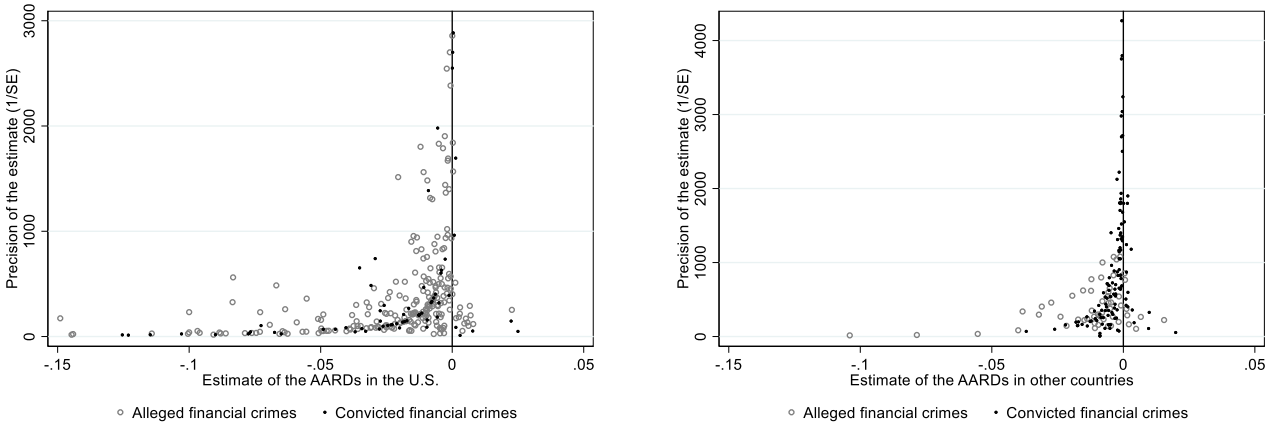
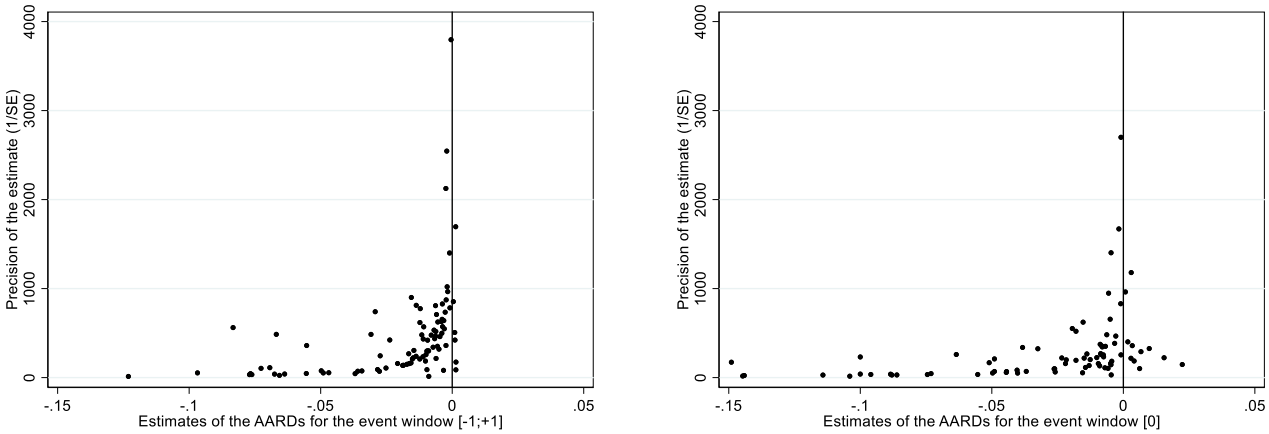
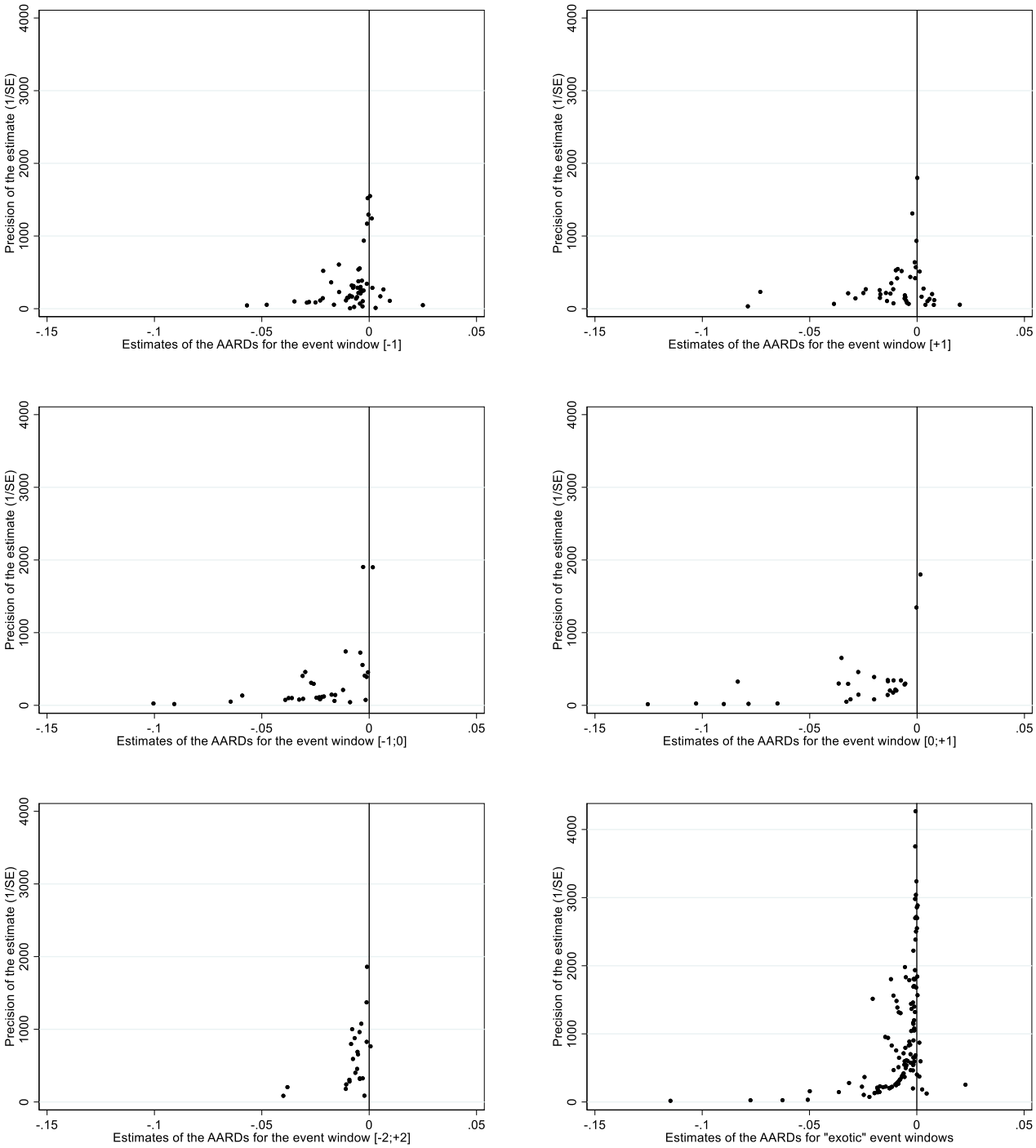


Figure A.6: Funnel Plots by Event Windows

Figure A.6 details the funnel plots depending on the event windows of the estimates, for the most frequent event windows (i.e. standing for more than 5% of the sample of 479 estimates). The following event windows stood for more than 24 estimates, by declining order of frequency. [-1;+1], [0], [-1], [+1], [-1;0], [0;+1], and [-2;+2]. The other event windows (123) are called "exotic" in the sense that the authors may have been tempted to publish these event windows to publish statistically significant abnormal returns and hence maximize their probability of publication.



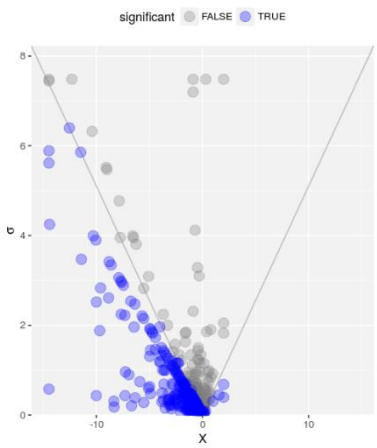


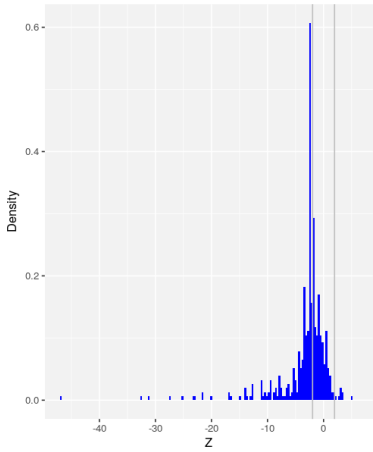
Source: Authors

Figure A.7: Funnel Plots and Histograms of Z-Statistics (Non-Linear Approach by Andrews and Kasy, 2019)

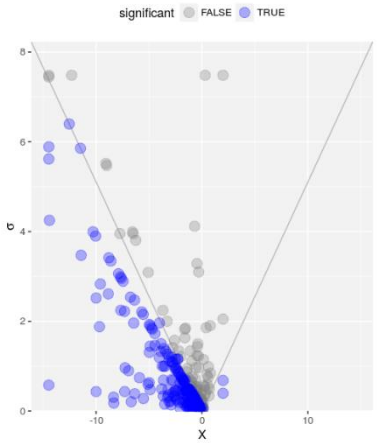
The following figures depict the funnel plots and the histograms of z-statistics for the full sample (panel A) and sub-samples (limited to the U.S. (Panel B.1.) or any other countries (Panel B.2.) and limited to exclusively accounting frauds (Panel C.1.) or any violation of securities laws (Panel C.2.)). The calculations are done using Maximilian Kasy's online application, which allows the estimation of models of selection publication using meta-studies.

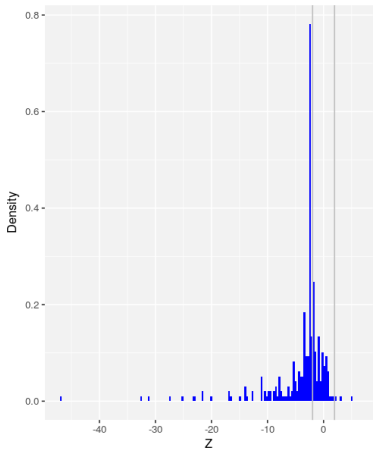
Panel A. Full Sample



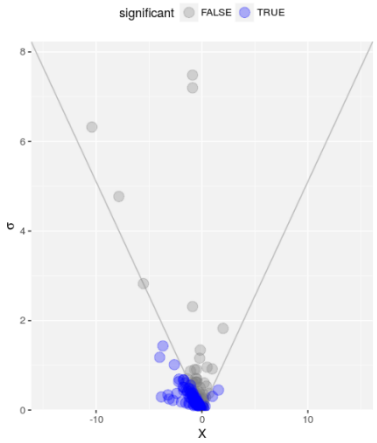


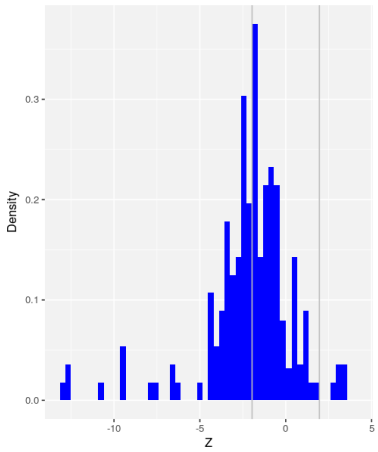
Panel B.1. Financial Crimes Committed in the U.S.



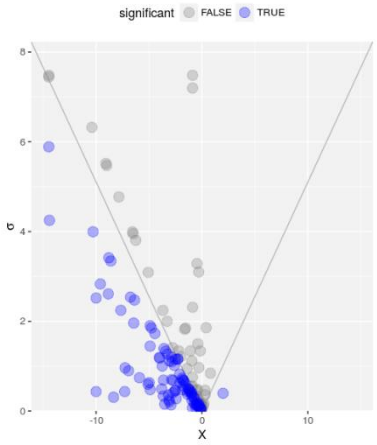


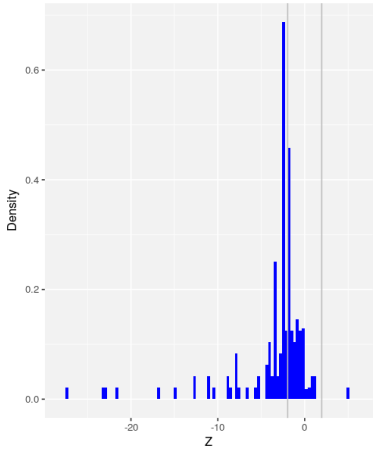
Panel B.2. Financial Crimes Committed in Other Countries



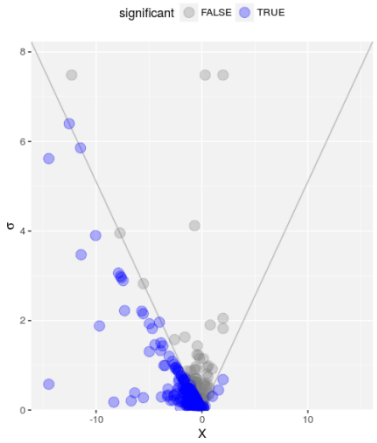


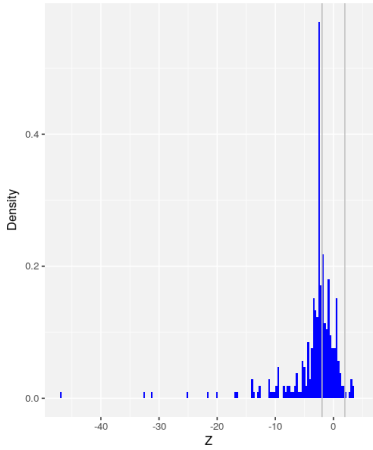
Panel C.1. Exclusively Accounting Frauds





Panel C.2. Any Violation of Securities Laws





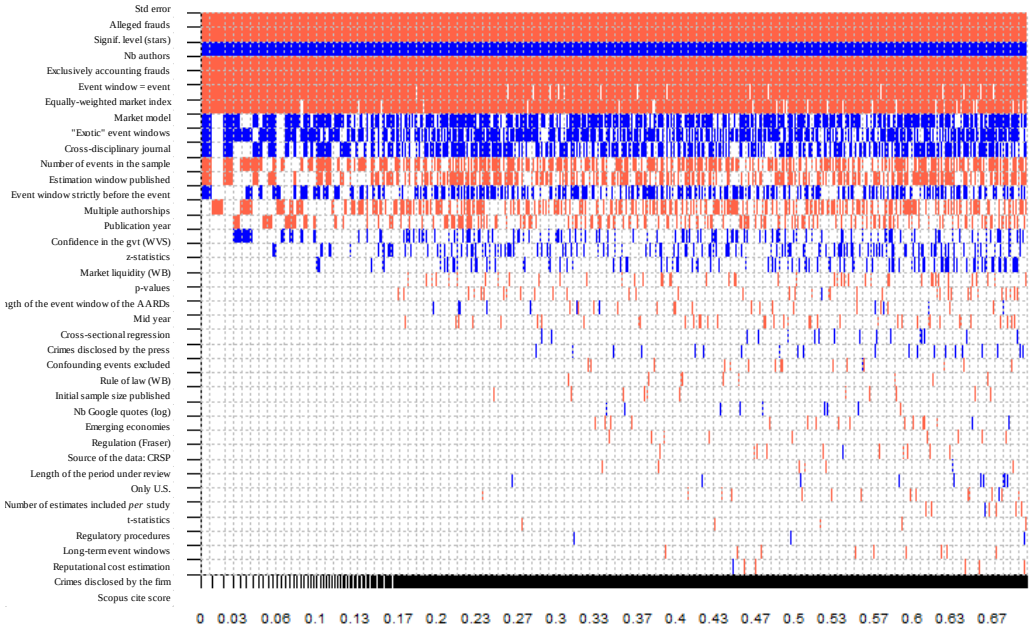
Source: Authors, calculations https://maxkasy.github.io/home/metastudy/

Appendix G. Robustness Checks of BMA (for Online Publication)

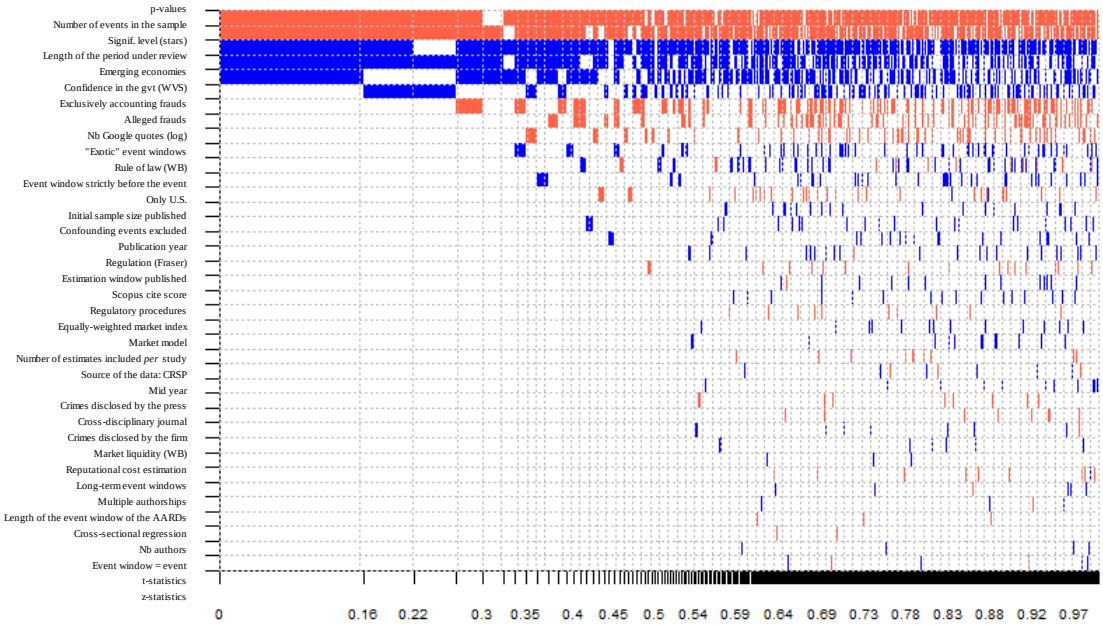
Figure A.8: Model Inclusion in Bayesian Model Averaging

Figure A.8 is a robustness check of the BMA results of Figure 6. This figure depicts the model inclusion in Bayesian Model Averaging, with the average abnormal returns per day as the response variable. All variables are weighted by the inverse of the number of estimates reported per study (Panel A) or by the inverse of the standard errors (Panel B). Each column denotes an individual model. The variables are sorted by Posterior Inclusion Probability (PIP) in descending order. The horizontal axis denotes the cumulative posterior model probabilities for the 10,000 best models. The blue color (or darker in grayscale) means that the estimated parameter of the explanatory variable is positive. Conversely, the red color (or lighter in grayscale) indicates a negative sign for the estimated parameter. No color denotes that the variable is not included in the model. A detailed description of all variables is available in Table A.3 of Appendix E. We use our baseline specification with the unit information prior recommended by Eicher et al. (2011) and the dilution prior suggested by George (2010), which addresses collinearity.

Panel A. Variables Weighted by the Number of Estimates Reported per Study



Panel B. Variables Weighted by the Inverse of the Standard Errors



Source: Authors

Appendix H. Hedges' Test for Publication Bias (for Online Publication)

As a robustness check, the results on the publication bias of the literature on financial crimes are complemented by Hedges' model (1992)[36] and the augmented model by Ashenfelter et al. (1999).[37] Hedges' model assumes that the probability of publication of estimates is determined by their statistical significance, with jumps for the psychologically important *p*-value. These thresholds are typically 0.01, 0.05, and 0.1 in economics. All estimates significant or insignificant at the conventional levels should have the same probability of being published in the absence of publication bias. Ashenfelter et al. (1999) allowed for heterogeneity related to publication bias in the estimates of the underlying effect.

As in Havránek and Sokolova (2020),[38] we assume four intervals of *p*-values reflecting different levels of conventional statistical significance of the estimates: below 0.01, between 0.01 and 0.05, between 0.05 and 0.1, and above 0.1. For the first step, *p*-value < 0.01, we normalize ω to 1 and evaluate whether the remaining three weights differ from this value. Regarding the characteristics of the estimates, we control for the following publication characteristics, which might be related to publication bias: the publication year, the number of citations in Google Scholar, publication in Scopus journal, and the RePec impact factor of the journal.

Table A.4 shows the estimation results for two models: 1) an unrestricted model, assuming a publication bias and 2) a restricted model, with $\omega_2 = \omega_3 = \omega_4 = 1$, assuming no publication bias (in other words, all coefficients have the same probability of being published, different statistical significance notwithstanding). Part A details the results of Hedges' model without heterogeneity in the estimates of excess sensitivity (simple model). The restriction is rejected, which suggests publication bias: estimates significant at the 1% level are much more likely to get published than all other estimates (the differences among the three remaining groups are not statistically significant). Part B displays similar results when allowing for heterogeneity in the estimates of excess sensitivity that might potentially be related to publication bias.

Table A.4: Hedges' Test of Publication Bias

	A. Simple model				**B. Model controlling for publication characteristics**			
	Unrestricted model		**Restricted model (ω_j=1)**		**Unrestricted model**		**Restricted model (ω_j=1)**	
	Coeff.	Standard error	Coeff.	Standard error	Coeff.	Standard error	Coeff.	Standard error
ω_2	-8.626	6.604			-4.788	3.928		
ω_3	-10.741	9.399			-5.539	5.183		
ω_4	-82.172	9.399			-40.417	15.734		
Publication year					-0.001	0.000	-0.001	0.000
Citations in Google Scholar					-0.009	0.002	-0.009	0.002
Scopus journal					0.015	0.006	0.026	0.005
RePec impact factor					0.000	0.002	-0.002	0.002
Constant	0.040	0.017	-0.035	0.002	0.035	0.013	0.007	0.011
σ	-0.108	0.007	-0.042	0.002	-0.037	0.002	-0.039	0.002
Log likelihood	1089.262		1174.252		1453.452		1199.135	
Observations	479		479		479		479	
	χ^2 (H_0: all estimates have the same probability of publication): 170.0, p-value < 0.001.				χ^2 (H_0: all estimates have the same probability of publication): 508.6, p-value < 0.001.			

Notes: Without publication bias, all estimates, whatever their statistical significance, should have the same probability of being reported. ω_1, the weight associated with the probability of publication for estimates significant at the 1% level, is set to 1. ω_2, ω_3, and ω_4 show the relative probabilities of publication for estimates significant at the 5% level, significant at the 10% level, and insignificant, respectively. σ is the estimated measure of heterogeneity (standard deviation) of the estimates of excess sensitivity.

[36] Hedges, L. V. (1992). Meta-Analysis. *Journal of Educational Statistics*, 17(4), 279-296.
[37] Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A Review of Estimates of the Schooling/Earnings Relationship, With Tests for Publication Bias. *Labour Economics*, 6(4), 453-470.
[38] Havránek, T., & Sokolova, A. (2020). Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say "Probably Not". *Review of Economic Dynamics*, 35, 97-122.

Appendix I: Studies Included in the Meta-analysis Dataset (for Online Publication)

Abdulmanova, A., Ferris, S. P., Jayaraman, N., & Kothari, P. (2019). The Effect of Investor Attention on Fraud Discovery and Value Loss in Securities Class Action Litigation. *Georgia Tech Scheller College of Business Research Paper*, (17-34).

Aggarwal, R., Hu, M., & Yang, J. (2015). Fraud, Market Reaction, and the Role of Institutional Investors in Chinese Listed Firms. *The Journal of Portfolio Management*, 41(5), 92-109.

Agrawal, A., & Chadha, S. (2005). Corporate Governance and Accounting Scandals. *The Journal of Law and Economics*, 48(2), 371-406.

Agrawal, A., & Cooper, T. (2017). Corporate Governance Consequences of Accounting Scandals: Evidence from Top Management, CFO and Auditor Turnover. *Quarterly Journal of Finance*, 7(1), 1650014.

Akhigbe, A., Kudla, R. J., & Madura, J. (2005). Why are Some Corporate Earnings Restatements More Damaging? *Applied Financial Economics*, 15(5), 327-336.

Amoah, N. Y., & Tang, A. P. (2010). Board, Audit Committee and Restatement-Induced Class Action Lawsuits. *Advances in Accounting*, 26(2), 155-169.

Amoah, N.Y. (2013). What is Fraud in Private Securities Lawsuit? *Advances in Public Interest Accounting*, 16, 39-63.

Andersen, A., Gilbert, A., & Tourani-Rad, A. (2013). Breach of Continuous Disclosure in Australia. *JASSA*, 4, 21-26.

Anderson, K.L., & Yohn, T. (2002). The Effect of 10-K Restatements on Firm Value, Information Asymmetries, and Investors' Reliance on Earnings. *Working paper, Georgetown University*.

Armour, J., Mayer, C., & Polo, A. (2017). Regulatory Sanctions and Reputational Damage in Financial Markets. *Journal of Financial and Quantitative Analysis*, 52(4), 1429-1448.

Arnold, M., & Engelen, P. J. (2007). Do Financial Markets Discipline Firms for Illegal Corporate Behaviour. *Management & Marketing*, 2(4), 103-110.

Baker, H. K., Edelman, R. B., & Powell, G. E. (1999). The Effect of Announcements of Corporate Misconduct and Insider Trading on Shareholder Returns. *Business and Professional Ethics Journal*, 18(1), 47-64.

Barabanov, S. S., Ozocak, O., Turtle, H. J., & Walker, T. J. (2008). Institutional Investors and Shareholder Litigation. *Financial Management*, 37(2), 227-250.

Bardos, K. S., Golec, J., & Harding, J. P. (2013). Litigation Risk and Market Reaction to Restatements. *Journal of Financial Research*, 36(1), 19-42.

Bardos, K. S., & Mishra, D. (2014). Financial Restatements, Litigation and Implied Cost of Equity. *Applied Financial Economics*, 24(1), 51-71.

Barniv, R. R., & Cao, J. (2009). Does Information Uncertainty Affect Investors' Responses to Analysts' Forecast Revisions? An Investigation of Accounting Restatements. *Journal of Accounting and Public Policy*, 28(4), 328-348.

Bauer, R., & Braun, R. (2010). Misdeeds Matter: Long-Term Stock Price Performance After the Filing of Class-Action Lawsuits. *Financial Analysts Journal*, 66(6), 74-92.

Beasley, M. S., Carcello, J. V., Hermanson, D. R., & Neal, T. L. (2010). Fraudulent Financial Reporting: 1998-2007: An Analysis of US Public Companies. *COSO, Committee of Sponsoring Organizations of the Treadway Commission*.

Beneish, M. D. (1999). Incentives and Penalties Related to Earnings Overstatements that Violate GAAP. *The Accounting Review*, 74(4), 425-457.

Bhagat, S., Bizjak, J., & Coles, J. L. (1998). The Shareholder Wealth Implications of Corporate Lawsuits. *Financial Management*, 27, 5-27.

Billings, M. B., Klein, A., & Zur, E. (2012). Shareholder Class Action Suits and the Bond Market. *Working paper* available at SSRN 1984666.

Bohn, J., & Choi, S. (1996). Fraud in the New-Issues Market: Empirical Evidence on Securities Class Actions. *University of Pennsylvania Law Review*, 144(3), 903-982.

Bonini, S., & Boraschi, D. (2010). Corporate Scandals and Capital Structure. *Journal of Business Ethics*, 95, 241-269.

Bowen, R. M., Call, A. C., & Rajgopal, S. (2010). Whistle-Blowing: Target Firm Characteristics and Economic Consequences. *The Accounting Review*, 85(4), 1239-1271.

Bradley, D., Cline, B. N., & Lian, Q. (2014). Class Action Lawsuits and Executive Stock Option Exercise. *Journal of Corporate Finance*, 27, 157-172.

Brous, P. A., & Leggett, K. (1996). Wealth Effects of Enforcement Actions Against Financially Distressed Banks. *Journal of Financial Research*, 19(4), 561-577.

Burns, N., & Kedia, S. (2006). The Impact of Performance-Based Compensation on Misreporting. *Journal of Financial Economics*, 79(1), 35-67.

Callen, J. L., Livnat, J., & Segal, D. (2006). Accounting Restatements: Are They Always Bad News for Investors? *The Journal of Investing*, 15(3), 57-68.

Chava, S., Cheng, C. A., Huang, H., & Lobo, G. J. (2010). Implications of Securities Class Actions for Cost of Equity Capital. *International Journal of Law and Management*, 52 (2), 144-161.

Chen, G., Firth, M., Gao, D. N., & Rui, O. M. (2005). Is China's Securities Regulatory Agency a Toothless Tiger? Evidence from Enforcement Actions. *Journal of Accounting and Public Policy*, 24(6), 451-488.

Choi, H. M., Karpoff, J. M., Lou, X., & Martin, G. S. (2019). Enforcement Waves and Spillovers. *Working paper*, available at SSRN 3526555.

Choi, S. J., & Pritchard, A. C. (2016). SEC Investigations and Securities Class Actions: An Empirical Comparison. *Journal of Empirical Legal Studies*, 13(1), 27-49.

Christensen, T.E., D.G. Paik, & C.D. Williams, (2010). Market Efficiency and Investor Reactions to SEC Fraud Investigations. *Journal of Forensic & Investigative Accounting*, 2(3), 1-30.

Cook, T., & Grove, H. (2009). The Stock Market Reaction to Allegations of Fraud and Earnings Manipulation. *Journal of Forensic & Investigative Accounting*, 1(2), 1-29.

Correia, M., & Klausner, M. (2012). Are Securities Class Actions "Supplemental" to SEC Enforcement? An Empirical Analysis. *Working Paper Stanford Law School Stanford.*
Cox, R. A., & Weirich, T. R. (2002). The Stock Market Reaction to Fraudulent Financial Reporting. *Managerial Auditing Journal*, 17(7), 374-382.
Davidson, W. N., Worrell, D. L., & Lee, C. I. (1994). Stock Market Reactions to Announced Corporate Illegalities. *Journal of Business Ethics*, 13(12), 979-987.
Davis, F., Taghipour, B., & Walker, T. J. (2017). Insider Trading Surrounding Securities Class Action Litigation and Settlement Announcements. *Managerial Finance*, 43(1) 124-140.
de Batz, L. (2020). Financial Impact of Regulatory Sanctions on Listed Companies. *European Journal of Law and Economics*, 49(2), 301-337.
Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC. *Contemporary Accounting Research*, 13(1), 1-36.
Deng, S., Willis, R. H., & Xu, L. (2014). Shareholder Litigation, Reputational Loss, and Bank Loan Contracting. *Journal of Financial and Quantitative Analysis*, 49(4), 1101-1132.
Desai, H., Hogan, C. E., & Wilkins, M. S. (2006). The Reputational Penalty for Aggressive Accounting: Earnings Restatements and Management Turnover. *The Accounting Review*, 81(1), 83-112.
Djama, C. (2013). Fraudes à l'Information Financière et Contrôle de l'AMF : Une Etude des Réactions du Marché Financier Français. *Revue Française de Gestion*, 231, 133-157.
Du, L. (2017). The CDS Market Reaction to Restatement Announcements. *Journal of Business Finance & Accounting*, 44(7-8), 1015-1035.
Engelen, P.J. (2009). The Reputational Penalty for Illegal Insider Trading by Managers, *Working paper presented at the Academy of Management Conference, Chicago U.S.*
Engelen, P. J. (2011). Legal versus Reputational Penalties in Deterring Corporate Misconduct. In D. Sunderland & M. Ugur (Eds.), Does Governance Matter? Governance Institutions and Outcomes, 71-95.
Engelen, P.J. (2012). What is the Reputational Cost of a Dishonest CEO? Evidence from US Illegal Insider Trading. *CESifo Economic Studies*, 58, 140-163.
Eryiğit, M. (2019). Short-Term Performance of Stocks after Fraudulent Financial Reporting Announcement. *Journal of Financial Crime*, 26(2), 464-476.
Ewelt-Knauer, C., Knauer, T., & Lachmann, M. (2015). Fraud Characteristics and Their Effects on Shareholder Wealth. *Journal of Business Economics*, 85(9), 1011-1047.
Feroz, E. H., Park, K., & Pastena, V. S. (1991). The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases. *Journal of Accounting Research*, 29, 107-142.
Ferris, S. P., Jandik, T., Lawless, R. M., & Makhija, A. (2007). Derivative Lawsuits as a Corporate Governance Mechanism: Empirical Evidence on Board Changes Surrounding Filings. *Journal of Financial and Quantitative Analysis*, 42(1), 143-165.
Fich, E. M., & Shivdasani, A. (2007). Financial Fraud, Director Reputation, and Shareholder Wealth. *Journal of Financial Economics*, 86(2), 306-336.
Firth, M., Rui, O. M., & Wu, W. (2011). Cooking the Books: Recipes and Costs of Falsified Financial Statements in China. *Journal of Corporate Finance*, 17(2), 371-390.
Firth, M., Rui, O. M., & Wu, X. (2009). The Tand Consequences of Disseminating Public Information by Regulators. *Journal of Accounting and Public Policy*, 28(2), 118-132.
Firth, M., Wong, S., Xin, Q., & Yick, H. Y. (2016). Regulatory Sanctions on Independent Directors and Their Consequences to the Director Labor Market: Evidence from China. *Journal of Business Ethics*, 134(4), 693-708.
Flore, C., Degryse, H., Kolaric, S., & Schiereck, D. (2018). Forgive Me All My Sins: How Penalties Imposed on Banks Travel Through Markets. *Working Paper, available at SSRN 3178589*.
Gande, A., & Lewis, C. M. (2009). Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers. *Journal of Financial and Quantitative Analysis*, 44(4), 823-850.
Gerety, M., & Lehn, K. (1997). The Causes and Consequences of Accounting Fraud. *Managerial and Decision Economics*, 18(7-8), 587-599.
Goldman, E., Peyer, U., & Stefanescu, I. (2012). Financial Misrepresentation and its Impact on Rivals. *Financial Management*, 41(4), 915-945.
Griffin, P. A., Grundfest, J. A., & Perino, M. A. (2004). Stock Price Response to News of Securities Fraud Litigation: An Analysis of Sequential and Conditional Information. *Abacus*, 40(1), 21-48.
Griffin, P. A., & Sun, E. (2011). Troublesome Tidings? Investors' Response to a Wells Notice. *Accounting and Finance Research*, 5(1), 99-120.
Haslem, B., Hutton, I., & Smith, A. H. (2017). How Much do Corporate Defendants Really Lose? A New Verdict on the Reputation Loss Induced by Corporate Litigation. *Financial Management*, 46(2), 323-358.
Hirschey, M., Palmrose, Z. V., & Scholz, S. (2005). Long-Term Market Underreaction to Accounting Restatements. *Working Paper, University of Kansas, School of Business*.
Humphery-Jenner, M. L. (2012). Internal and External Discipline Following Securities Class Actions. *Journal of Financial Intermediation*, 21(1), 151-179.
Iqbal, Z., Shetty, S., & Wang, K. (2007). Further Evidence on Insider Trading and the Merits of Securities Class Actions. *Journal of Financial Research*, 30(4), 533-545.
Johnson, S. A., Ryan, H. E., & Tian, Y. S. (2003). Executive Compensation and Corporate Fraud. *Working Paper, Lousiana State University, Baton Rouge, LA*.
Jordan, J. S., Peek, J., & Rosengren, E. S. (2000). The Market Reaction to the Disclosure of Supervisory Actions: Implications for Bank Transparency. *Journal of Financial Intermediation*, 9(3), 298-319.

Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and Databases in Financial Misconduct Research. *The Accounting Review*, 92(6), 129-163.
Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008a). The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis*, 43(3), 581-611.
Karpoff, J. M., & Lott Jr, J. R. (1993). The Reputational Penalty Firms Bear from Committing Criminal Fraud. *The Journal of Law and Economics*, 36(2), 757-802.
Kellogg, R. L. (1984). Accounting Activities, Security Prices, and Class Action Lawsuits. *Journal of Accounting and Economics*, 6(3), 185-204.
Kirat, T., & Rezaee, A. (2019). How Stock Markets React to Regulatory Sanctions? Evidence from France. *Applied Economics*, 51(60), 6558-6566.
Klock, M. (2015). Do Class Action Filings Affect Stock Prices? The Stock Market Reaction to Securities Class Actions Post PSLRA. *Journal of Business & Securities Law*, 15(2), 109-156.
Kouwenberg, R., & Phunnarungsi, V. (2013). Corporate Governance, Violations and Market Reactions. *Pacific-Basin Finance Journal*, 21(1), 881-898.
Kravet, T., and Shevlin, T. (2010). Accounting Restatements and Information Risk. *Review of Accounting Studies*, 15(2), 264-294.
Kryzanowski, L., & Zhang, Y. (2013). Financial Restatements by Canadian Firms Cross-Listed and Not Cross-Listed in the US. *Journal of Multinational Financial Management*, 23(1-2), 74-96.
Kwan, J. H., & Kwan, S. S. (2011). Violation of Listing Requirements and Company Value: Evidence from Bursa Malaysia. *International Review of Business Research Papers*, 7(2), 257-268.
Lei, A. C. H., & Law, P. K. F. (2019). Financial Fraud, CEO Turnover and Regulatory Effectiveness: Evidence from China. *Working Paper presented in the 27th Conference on the Theories and Practices of Securities and Financial Markets.*
Liebman, B. L., & Milhaupt, C. J. (2008). Reputational Sanctions in China's Securities Market. *Business Finance & Accounting, Columbia Law Review*, 108(4), 929-983.
Lieser, P., & Kolaric, S. (2016). Securities Class Action Litigation, Defendant Stock Price Revaluation, and Industry Spillover Effects. *Working paper, European Financial Management Association Conference paper n° 0388*.
Loh, C., & Rathinasamy, R. S. (2003). Do All Securities Class Actions Have the Same Merit? A Stock Market Perspective. *Review of Pacific Basin Financial Markets and Policies*, 6(02), 167-178.
Marciukaityte, D., Szewczyk, S. H., Uzun, H., & Varma, R. (2006). Governance and Performance Changes after Accusations of Corporate Fraud. *Financial Analysts Journal*, 62(3), 32-41.
Marciukaityte, D., Szewczyk, S. H., & Varma, R. (2009). Voluntary vs. Forced Financial Restatements: The Role of Board Independence. *Financial Analysts Journal*, 65(5), 51-65.
McDowell, J. (2005). A Look at the Market's Reaction to the Announcements of SEC Investigations. *Working paper, The Leonard N. Stern School of Business.*
Muradoglu, Y. G., & Clark Huskey, J. (2008). The Impact of SEC Litigation on Firm Value. *Working Paper*.
Nainar, S. M. K., Rai, A., & Tartaroglu, S. (2014). Market Reactions to Wells Notice: An Empirical Analysis. *International Journal of Disclosure and Governance*, 11(2), 177-193.
Nelson, C., Gilley, S., & Trombley, G. (2009). Disclosures of SEC Investigations Resulting in Wells Notices. *Securities Litigation Journal*, 19(4), 19-21.
Nourayi, M. M. (1994). Stock Price Responses to the SEC's Enforcement Actions. *Journal of Accounting and Public Policy*, 13(4), 333-347.
Owers, J., Lin, C., & Rogers, R. (2002). The Information Content and Valuation Ramifications of Earnings Restatements. *International Business and Economics Research Journal*, 1(5), 71-83.
Ozbas, O. (2008). Corporate Fraud and Real Investment. *Working paper*.
Ozeki, N. (2019). Determinants of Market Reaction to Disclosure of Accounting Misconduct: Evidence from Japan, *Securities Analysts Journal*, 57(3), 72-84.
Pereira, J., Malafronte, I., Sorwar, G., & Nurullah, M. (2019). Enforcement Actions, Market Movement and Depositors' Reaction: Evidence from the US Banking System. *Journal of Financial Services Research*, 55(2-3), 143-165.
Persons, O.S. (1997). SEC's Insider Trading Enforcements and Target Firms' Stock Values. *Journal of Business Research*, 39(3), 187-194.
Plumlee, M., & Yohn, T. (2008). Restatements: Investors Response and Firm Reporting Choices. *Working Paper*.
Pritchard, A. C., & Ferris, S. P. (2001). Stock Price Reactions to Securities Fraud Class Actions Under the Private Securities Litigation Reform Act. *Michigan Law and Economics Research Paper, n°01-009*.
Romano, R. (1991). The Shareholder Suit: Litigation Without Foundation? *Journal of Law, Economics, and Organization*, 7, 55-87.
Scholz, S. (2008). The Changing Nature and Consequences of Public Company Financial Restatements. The United States Department of the Treasury, 1-54.
Slovin, M. B., Sushka, M. E., & Polonchek, J. A. (1999). An Analysis of Contagion and Competitive Effects at Commercial Banks. *Journal of Financial Economics*, 54, 197-225.
Song, C., & Han, S. H. (2017). Stock Market Reaction to Corporate Crime: Evidence from South Korea. *Journal of Business Ethics*, 143(2), 323-351.
Sun, P., & Zhang, Y. (2006). Is There Penalty for Crime: Corporate Scandal and Management Turnover in China? *EFA 2006 Zurich Meetings Paper*.
Takmaz, S., & Keles, E. M. (2017). Do Stock Prices React to Illegal Corporate Behaviors? The Turkish Case. *Journal of Business Research Turk*, 9(4), 245-258.
Tanimura, J. K., & Okamoto, M. G. (2013). Reputational Penalties in Japan: Evidence from Corporate Scandals", *Asian Economic Journal*, 27(1), 39-57.

Tay, L. M., Puah, C. H., Brahmana, R. K., & Abdul Malek, N. I. (2016). The Effect of White Collar Crime Announcement on Stock Price Performance: Evidence from Malaysian Stock Market. *Journal of Financial Crime*, 23(4), 1126-1139.
Wang, Y., Ashton, J. K., & Jaafar, A. (2019). Money Shouts! How Effective are Punishments for Accounting Fraud? *The British Accounting Review*, 51(5), 100824.
Wang, X., & Wu, M. (2011). The Quality of Financial Reporting in China: An Examination from an Accounting Restatement Perspective. *China Journal of Accounting Research*, 4(4), 167-196.
Wu, M. (2002). Earnings Restatements: A Capital Market Perspective. *Working paper, New York University, New York, NY.*
Wu, X., & Zhang, J. (2014). Stock Market Reaction to Regulatory Investigation Announcements. *China Journal of Accounting Studies*, 2(1), 37-52.
Xu, W., & Xu, G. (2020). Understanding Public Enforcement of Securities Law in China: An Empirical Analysis of the Enforcement Actions of the CSRC and its Regional Offices Against Informational Misconduct. *International Review of Law and Economics*, 61, 105877.
Yu, X., Zhang, P., & Zheng, Y. (2015). Corporate Governance, Political Connections, and Intra-Industry Effects: Evidence from Corporate Scandals in China. *Financial Management*, 44(1), 49-80.
Zeidan, M.J. (2013). Effects of Illegal Behavior on the Financial Performance of US Banking Institutions. *Journal of Business Ethics*, 112, 313-324.
Zhu, Z., & Hu, C. (2010). Market Reactions to Financial Restatements-Evidence from Chinese Stock Market. *2010 IEEE International Conference on Industrial Engineering and Engineering Management*, 2527-2530.