CESifo WORKING PAPERS

10528
2023

June 2023

Financial Crime and Punishment: A Meta-Analysis

Laure de Batz, Evžen Kočenda

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10528*

Financial Crime and Punishment: A Meta-Analysis

Abstract

We provide the first quantitative synthesis of the literature on how financial markets react to the disclosure of financial crimes committed by listed firms. While consensus expects negative stock price returns, the exact size of the effect is far from clear. We survey 111 studies published over three decades, from which we collect 480 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. Hence, the "punishment" effect, including a reputational penalty, suffers from a serious publication bias. After controlling for this bias, negative abnormal returns suggest the existence of an informational effect. We also document that accounting frauds, crimes committed in common-law countries such as the United States, and allegations are particularly severely sanctioned by financial markets, while the information channels and types of procedures do not influence market reactions.

JEL-Codes: C830, G140, G180, K420, N240.

Keywords: meta-analysis, event study, financial misconduct, trust, information and market efficiency, listed companies, crime.

*Laure de Batz**
Centre d'Économie de la Sorbonne
Université Paris 1 Panthéon-Sorbonne
France – 75642 Paris Cedex 13
laure.debatz@gmail.com
Orcid: 0000-0002-6124-0353

Evžen Kočenda
Institute of Economic Studies
Charles University
Czech Republic – 11000 Prague
evzen.kocenda@fsv.cuni.cz
Orcid: 0000-0003-4110-1169

*corresponding author

Declarations of interest: none

Data Availability Statement: The data and the program files (Stata and R) can be downloaded in the "Research" section (https://kocenda.fsv.cuni.cz/research.htm).

We are thankful to Tomáš Havránek for his valuable advice on the project development. We also benefited from helpful comments from Josef Bajzík, Hristos Doucouliagos, Svatopluk Kapounek, Jerome Geyer-Klingeberg, Tom Stanley, two anonymous referees, and participants at conferences held in Kyoto, London, Madrid, Marseille, Prague, and Strasbourg. Kočenda acknowledges financial support from the GACR grant No. 23-07983S. The usual disclaimer applies.

1. Introduction

Major financial crimes involving listed firms have been hitting the headlines around the world for decades, from accounting frauds, to rog trading, Ponzi schemes, and insider trading. The most notorious scandals include Enron, WorldCom, Bernard L. Madoff Investment Securities LLC, and Theranos in the United States (U.S.)., Parmalat, Wirecard, Jérôme Kerviel (Société Générale), Bruno Iksil (JPMorgan Chase) in Europe, and Satyam, 1MDB, and Nick Leeson (Barings Bank) in Asia, to name just a few. Subsequent damages are substantial and affect a wide range of stakeholders, from shareholders, managers, and employees of those firms to clients, suppliers and bankers. They also impact economies and industries, and challenge the abilities of regulators the in face of the scandals yet to come (Bhaskar et al., 2019). Still, such financial scandals are only the tip of the iceberg of financial crimes. Only a limited share of financial crimes is detected, so-called partial observability (Ashton et al., 2021; Ormosi; 2014), and most of them do not reach the magnitude of a scandal. Still, when disclosed, financial crimes negatively impact firms and their shareholders (Amiram et al., 2018). Financial crimes threaten the existence and efficiency of capital markets, which are based on trust from market participants.[1] When detected, such financial crimes are punished by the relevant authorities. But how do financial markets react to the disclosure of financial crimes committed by listed firms and how large is the reaction in quantitative terms? How extensive is the market punishment? What drives the results found in the related literature? In our meta-analysis, we aim to answer just that.

In line with the academic, practitioner, and policy literature, we define financial crimes committed by listed firms as: insider dealing, price manipulation, breach of public disclosure requirements (*i.e.* the three market abuses), and more generally breaches of financial regulations (for detailed descriptions see Appendix C, Figure C.2). The disclosure of such financial crimes can follow allegations (by enforcers, journalists, analysts, etc.), sanctions imposed by relevant authorities, settlements, trials, etc. In efficient markets (Fama, 1970), this disclosure is expected to trigger immediate market reactions: stock prices should decline to reflect the expected or proven direct costs of crime (fines, legal fees, potential subsequent accounting restatement, etc.), possibly complemented with indirect costs, due to the firm's deteriorated reputation. This so-called reputational penalty (Peltzman, 1981; Karpoff and Lott; 1993) arises from the expected loss in the present value of future cash flows (Kapoff et al., 2008), due to deteriorated relationships with clients and suppliers and to higher cost of doing business (contracting, compliance, and financing costs).

Most of the literature we analyze is in some agreement that financial crimes are punished on the financial markets by a drop of stock prices, *i.e.* negative returns. But, 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 a financial crime. The evidence is often mixed or less than fully observed (Karpoff et al., 2017). Most studies investigate a limited number of financial crimes (264 on average in our sample, with a standard deviation of 378), challenging the meaningfulness of some results due to small sample biases,

[1] Trust is a pillar of investment decisions on capital markets (Guiso et al., 2008), especially during periods of distress (Sapienza and Zingales, 2012) and regarding corporate social responsibility (CSR) labels; Lins et al. (2017) show that during the 2008–2009 financial crisis, firms with high CSR intensity exhibited higher returns than low-CSR firms.

and primarily focus on the U.S. We illustrate the large heterogeneity in studies dealing with market reactions to the disclosure of financial crimes in Figure 1. First, Panel A depicts the heterogeneity of the distribution of the sample sizes of financial crimes investigated in surveyed literature over time, reflecting the partial observability of crimes. Second, in Panel B, we illustrate the heterogeneity in estimated abnormal market reactions to the disclosure of financial crimes over time with a mild positive trend, indicating lower market sanctions imposed as time passes by.

Because the literature is to an extent fragmented, we analyze the extant literature both to provide the most accurate estimation of the extent of the market punishment of financial crimes and to uncover sources of displayed heterogeneity (Stanley and Doucouliagos, 2019). We believe that a meta-analysis represents a particularly relevant tool to generalize results and to assess their robustness by aggregating conclusions of individual studies (Geyer-Klingeberg et al., 2020). To the best of our knowledge and confirmed by a recent review of meta-analyses in finance (Geyer-Klingeberg et al., 2020), 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. Our goal is to systematically and quantitatively synthesize previous empirical results regarding market reactions subsequent to the disclosure of financial crimes and to dig into the drivers of the market reactions found in the related literature.

We surveyed all available literature until May 1, 2020, and identified 862 articles published from 1978 to 2020. We selected 111 articles that use the event-study methodology and collected a large sample of 480 estimates of stock price reactions quantified as abnormal returns that followed disclosures of financial crimes (see Table 1 and Appendix A for descriptive statistics of the meta dataset). This way, we are able to synthesize market reactions after the disclosure of 32,500 crimes in total, committed in 17 countries between 1965 and 2018 (see Figures 2).[2] Meta-analyzing this literature is particularly relevant as it tries to circumvent the partial observability of crimes (Ashton et al., 2021; Ormosi; 2014) by enlarging to the maximum possible extent the sample of financial crimes, while using the event study methodology which offers a directly-available and comparable size effect and is widely recognized in policy and financial analyses (Fama, 1990; Bhagat and Romano, 2002a, b; Geyer-Klingeberg et al., 2020).

In our analysis, we (i) employ recent methodological innovations in the literature on meta-analysis in economics and finance, and (ii) follow the standards on its conduct specified in Havránek et al. (2020). We first analyze the entire sample of estimates, graphically and statistically, to investigate for publication bias and for the true effect beyond bias (Stanley and Doucouliagos, 2012; Bajzík et al., 2020), using linear and non-linear (Ioannidis et al., 2017; Andrews and Kasy, 2019; Furukawa, 2019) models. Second, 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). Third, capitalizing on the key differentiation factors in the primary studies and on the

[2] 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. We acknowledge that some financial crimes may potentially overlap. However, based on the information in primary studies, we are not able to disentangle possible overlaps in financial crimes between studies.

rich meta-analysis literature, we circumvent the inherent model uncertainty by using Bayesian and frequentist methods of model averaging to choose the most important factors.[3]

Our contributions 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 an intentional financial crime substantially dampens the wealth of shareholders due to existence of negative abnormal returns over the few days around the event (-1.8% *per* day of the event window, or -4.8% over the event window). However, we bring evidence that the negative abnormal returns published in the meta-analyzed empirical literature seem to be exaggerated by three folds. As such, the effect of financial crime on stock price returns suffers from a serious publication bias (Brodeur et al., 2016), an issue that was voiced in number of studies.[4] 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 -0.5% *per* day over the event window following the disclosure of financial crimes (or -2.1% in terms of cumulative returns), more in line with abnormal market reactions to other types of white-collar crimes (Karpoff, 2012). Over-estimating negative abnormal returns as documented in the literature is disturbing because understanding the true underlying dynamics of the punishment of financial crimes is key to better enforcing the financial regulation of securities markets and protecting investors (La Porta et al., 2006). 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), or non-compliance with legal and regulatory framework (Jo, 2021). Accurate quantification of those consequences in terms of drops in returns is imperative since the violation of securities laws is one of the major causes of corporate failure (Soltani, 2014).

We further show that, beyond standard errors, the sampled countries, and the types of financial crimes contribute to the heterogeneity of reported abnormal returns. In particular, crimes committed in the U.S. (and more generally in common-law countries, where enforcement is more transparent), accounting frauds, and allegations drive down market corrections.

In terms of policy implications, our findings also contribute to a regulatory debate on how to come closer to an optimal level of regulation to deter future crimes: disclosed intentional financial crimes are priced-in by market participants. Hence, if an enforcer's goal is that markets react to their decisions and communications, then enforcement actions (from the issuance of warnings to sanction decisions) serve as a regulatory tool *per se*. Specifically, the evidence on

[3] Amongst others: Bajzík et al. (2020), Havránek and Sokolova (2020), Gechert et al. (2022), Kočenda and Iwasaki (2021), Matousek et al. (2021), Sokolova and Sorensen (2021), and Zigraiova et al. (2021).

[4] The existence of publication bias was documented in number of recent meta-analyses (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., 2022; 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).

negative abnormal returns accentuates how the "name and shame" mechanism[5] could efficiently contribute to enforcement: it penalizes the firms as a (partial) substitute for financial fines, and avoids long and costly enforcement procedures. Such a mechanism implicitly assumes that investors would react negatively to disclosed financial crimes while peers would be encouraged not to break similarly financial laws.

The rest of the article is structured as follows. In Section 2, we bring a short account of the relevant literature on financial crime. Then, we detail in section 3 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 4, followed by a heterogeneity analysis of why market reactions vary between studies (section 5). Finally, section 6 concludes and proposes policy-related interpretations.

2. Literature review

The literature on financial crimes is gaining traction, but our knowledge is constrained by the partial observability of crimes (Ashton et al., 2021; Ormosi; 2014), subsequent to the low share of detected crimes in the first place (Becker, 1968).[6]. Alarmingly, Alawadhi et al. (2020) assess that only 3.5% of financial misrepresentations are eventually caught and sanctioned in the U.S. 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). 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. Among all corporate misconducts, financial crimes, and in particular accounting frauds, trigger the strongest market reactions when disclosed (Karpoff, 2012), adding to the fact that shareholders' wealth can be harmed during the fraud period.[7]

In fact, when financial misconducts of listed firms become public information, the semi-strong efficient market hypothesis implies that their spillovers on the firms should be reflected immediately, fully, and in an unbiased manner in the stock prices (Fama, 1970). Put it differently, should disclosed financial crime be informational to market participants, the market should sanction the firm for not abiding the law and deceiving investors by the compounded forecasted costs, translating into a contraction in its market capitalization. Such costs are comprised of direct costs – reflecting fines, legal fees, compensations, and possibly the cash impact of restating financial accounts for accounting frauds – and of indirect longer-term costs (Dechow et al., 1996; Palmrose et al., 2004). In fact, being involved in financial crimes can result in an increase in the costs of doing business and severely damage corporate reputation (Engelen, 2011; Engelen and van Essen, 2011; Haslem et al., 2017; Karpoff, 2012 and 2020). All in all, direct and indirect costs should reduce firms' values. For that, it is reasonable to

[5] "Name and shame" mechanism has been extensively debated (Kahan and Posner, 1999) and is increasingly being adopted for accounting standards enforcement – for example in the U.S., Germany, and the U.K.

[6] 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.

[7] For example, as summarized in Karpoff (2012), environmental violations would trigger statistically insignificant to low abnormal market reactions (Jones and Rubin, 2001; Karpoff et al., 2005), while returns would contract more modestly subsequently to the disclosure of product recalls (Barber and Darrough, 1996), air safety disasters (Mitchell and Maloney, 1989), and antitrust charges (van den Broek et al., 2010) to name a few examples.

expect negative market reactions (measured as contracting returns) subsequent to disclosed financial crimes, as investors react to extant or expected reduced value of firms. In this sense, financial markets can function as an enforcement channel inducing companies to behave responsibly (Engelen, 2011). Is that a true reflection of reality?

One segment of the extant literature often concludes that the disclosure of financial crimes negatively impacts returns, contrary to environmental violations or foreign bribery (Karpoff, 2012, 2020). An event study methodology (see Appendix B for details) is typically used to estimate the "abnormal" market reaction subsequent to unanticipated news such as the disclosure of financial crimes (Karpoff and Lott, 1993; Alexander, 1999; Karpoff et al., 2008; Murphy et al., 2009; Engelen, 2011; Haslem et al., 2017; Karpoff et al., 2017; Armour et al., 2017; de Batz, 2020). 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). Such abnormal contraction in returns can be accounted for by the direct costs of crimes (imposed or forecasted legal penalties, legal fees, and compensations) as well as for dampened investors' expectations due to foreseen corporate shortcomings or degraded future prospects. 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." This opposite view might be potentially rooted in the fact that undergoing an investigation for alleged financial crimes can be an opportunity for the firm to correct internal problems and improper behaviors. Market participants may then respond positively during the investigation thereby revising forecasts to the upside.

Both stances are further reflected empirically in the literature. Christensen et al. (2016) empirically validate the "no-effect" hypothesis of the U.S. Securities and Exchange Commission's (SEC) enforcement actions on market quality, presented by Stigler (1964) and Peltzman (1976). Amiram et al. (2018) even challenge the rationale for levying fines. 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). The deterrent value of enforcement might rise with offenders' wealth (Garoupa, 2001).

Finally, part of the literature adds granularity in the analysis by isolating the indirect costs of financial crimes from the direct costs, within the compounded market reactions (*i.e.* the estimated abnormal contraction in market capitalizations). Direct costs are measurable: they encompass regulatory fines, legal fees, and remedial measures. Conversely, indirect costs – frequently called a "reputational penalty" after Karpoff and Lott (1993) and Engelen and van Essen (2011) among others – are not measurable and much harder to estimate. They reflect downgraded investors' expectations about the firms, for example due to 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, by publishing false information, or by manipulating others' shares. The spillovers of the disclosed financial crimes include degraded future business prospects, higher costs of doing business, or human resources costs (Karpoff et al., 2008; Armour et al., 2017). Such reputational penalty reflects the changes in the behavior of investors and related parties. Reputational penalties are typically proxied by deducting the direct costs of crime from the estimated abnormal market reactions following the disclosure of this financial crime (Karpoff and Lott, 1993; Cummins et al., 2006; Karpoff et al., 2008; Armour et al., 2017). Research on Anglo-Saxon financial crimes typically finds large and significant abnormal returns, mostly due to reputational penalties (Karpoff and Lott, 1993; Karpoff et al., 2008; Armour et al., 2017), contrary to (for example) foreign bribery or environmental violations or to financial crimes committed in civil laws countries (Karpoff, 2012, 2020). In that sense, financial markets are a complementary enforcement channel inducing companies to behave responsibly by punishing financial crimes (Engelen, 2011).

Negative market reactions to disclosed financial crimes and existing reputational penalties demonstrate that financial markets can efficiently complement (or even substitute for) enforcement of securities laws by punishing firms through negative returns. They can also act as a tool to deter financial crime, as illustrated by enforcers using the "name and shame" mechanism, instead of long, expensive, and uncertain sanction procedures.[8] Finding a good balance between enforcement and the market to encourage compliance with financial regulations is critical to efficiently protect investors and anchor trust.

3 Data selection and stylized facts

3.1 Selection of the data

Following the recent guidelines for meta-analytic research (Havránek et al., 2020), we reviewed and analyzed 862 studies identified from keyword searches performed in Google Scholar and in the major economic databases for specific topics related to financial crimes and punishment, complemented with references in these studies, and their Google Scholar citations. We terminated the search on May 1, 2020.[9]

We formed 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 financial crimes (see graphical illustrations of the scope of the sample in Appendix C); 3) specify the first public disclosure of the crime, whatever the source of information (newspaper articles, enforcers, or firms communication); 4) report (Cumulative) Average Abnormal Return(s) ((C)AARs) and an explicit indication of statistical significance (*t*-statistics, *p*-values, *z*-statistic, non-parametric tests, 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 Ph.D. 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. The complete reference

[8] A defendant can be cleared from charges for example on the ground of prescription limits or of procedural irregularities, which do not acquit the investigated person.

[9] Additional details on the data collection process and on the iterative process of selecting articles is disclosed in Appendix C and graphically illustrated by a PRISMA statement in Figure C.3, as recommended by Havránek et al. (2020).

[10] We made the choice to only include studies containing both (C)AARs and information on statistical significance, which are the natural output of event studies; as a result, 7 studies were excluded. 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. Further, we excluded 14 studies using lower-frequency data (monthly and yearly).

of each study can be found in Appendix K, and descriptive statistics of the meta dataset in Table 1. 81% of these studies were published in academic journals, the rest being 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] Including event windows preceding the events controls for possible 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).[12]

We obtain 480 effect estimates from the disclosure of 32,500 intentional financial crimes committed by listed firms (*i.e.* the events). By collecting all short-term estimates, we 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.

3.2 Directly available and comparable effect size: average abnormal returns

The goal of an event study is to quantify "abnormal" reactions to unanticipated events, here the disclosure of financial crimes (MacKinlay, 1997; Kothari and Warner, 2008; see Appendix B for methodological details). Abnormal returns are estimated for each event over the days of the "event window", possibly cumulated over a specific time interval including the event (so-called the "event window of the reported estimate") and then averaged across events. To do so, "normal" estimated returns – *i.e.* expected without conditioning on the event occurring and estimated with a help of a suitable model (typically a market model for 83% of our sample),[13]

---

[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.

[12] Various reasons can contribute to market inefficiencies, leading to no or postponed reactions: the time to access information (initially unaware, herd behaviors), the light financial education (misunderstanding of financial crimes), the avoidance of financial consequences (fees due to portfolio rebalancing, deterring fiscal consequences, etc.), or no investment alternative. It is also likely that some channels of news scaled in time after the publication, from part of the sanctioned company itself or newspaper articles, will contribute to postponed (or lagged) market abnormal reactions.

Part of the literature using event studies also justifies the length of the event windows by the authors' uncertainty on the event day. This is not relevant to our sampled articles as the disclosure of financial crimes can be precisely dated based on official communication channels. Financial crimes are typically disclosed in newspaper articles, in websites of enforcers (sanction decisions, settlements, legal repositories), or in firm's public statements.

[13] The statistical market model, possibly augmented to control for the sector for example, assumes a stable linear relation between the security return and the market return. Three other models of expected returns are used in this financial literature: market-adjusted model (13%), the Fama & French factor model (3%) and the capital asset pricing model (1%).

Additional methodological specifications, supporting the variables controlling for estimation characteristics variables, are detailed in the Appendix B.

based on recent history (the "estimation window" preceding the event) – are subtracted from "actual" observed returns over the "event window".

Limiting the scope of the meta-analysis to event studies provides a directly available and comparable effect size, with a straightforward economic interpretation: the estimated (cumulative) average abnormal returns ($(\mathrm{C})\mathrm{AARs}$) around the disclosure of financial crimes. This methodology 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 to financial crimes as the event dates are precisely known and unanticipated: they are communicated *via* official channels (typically the press, enforcers, or firms). This also facilitates the search for confounding events and their avoidance. A logical consequence is that we observe, after investigating exhaustively the literature on financial crimes, that this methodology is by far the most frequently used to assess the spillovers of disclosed financial crimes, as illustrated by the PRISMA statement (in Appendix C).

Event studies typically use hypothesis tests for the statistical significance of abnormal returns around the event day and, conventionally, the null hypothesis is that $(\mathrm{C})\mathrm{AARs}$ equal zero. The great majority of studies in the sample (84% of the sample) reports a statistical significance levels ("stars"), usually complemented with some statistics (Student's t-test statistics, z-statistics, p-values, and non-parametric test statistics).[14] 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 (264 financial crimes on average). 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 significance level was at least 10% for each. "Conservative" standard errors are then calculated from the published or estimated "conservative" t-statistics and the $(\mathrm{C})\mathrm{AARs}$ when not included in the study.[16] $(\mathrm{C})\mathrm{AARs}$ and standard errors are winsorized at the 1% level, to ensure that the presence of a few outliers does not result from mistakes in the primary studies.[17]

Still, there is no standard event window, depending on authors' *ad hoc* decisions. The event day ($t = 0$) is at least included in the reported event windows. At the level of primary studies, the average event window covers 33 trading days, ranging from 15 days before the event to 18 days after, with great heterogeneity (see Table 1). The heterogeneity remains for estimates collected in our meta-dataset for "short term" event windows $[\![-10; +10]\!]$: on average, the event window for reported estimates lasts for 4 days, with a standard deviation of 4.5 days, starting -1.6 days before the event and ending 1.4 days after. In order to control for

[14] 20% of the sample only report statistical significance levels ("stars").
[15] Desai et al. (2006), Nelson et al. (2009), and Goldman et al. (2012), standing for seven estimates.
[16] Only two studies published standard errors of $(\mathrm{C})\mathrm{AARs}$, standing for 10 estimates or 2% of the estimates. "Conservative" t-statistics are estimated as in Frooman (1997) as follows, when the t-statistics are not published: 1) the statistical significance levels are converted into conservative levels of significance;[16] 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.
[17] The means of CAARs and AARs, as well as the means AARDs, are slightly impacted by the winsorization from -4.889% to -4.769% and from -1.817% to -1.810%, respectively. The results hold with different levels of winsorization (2.5% and 5%).

this heterogeneity, and in spirit of Veld et al. (2018), we propose a two-step dual approach in the conduct of our meta-analysis: in the first step, we use the original sample of AARs and CAARs and standard errors, as collected from primary studies; in the second step, we employ abnormal returns and standard errors "normalized" by the length of their respective event windows. For that purpose, we create two complementary variables (Average Abnormal Return *per* Day, AARD, and Standard Errors *per* Day, SED), that equal the AAR or SE for one-day event windows or the CAAR or SE divided by the length of the event window (in days) otherwise.[18] AARDs and SED are also winsorized at the 1% level.[20]

By using this dual approach, we enlarge the sample of financial crimes to the maximum possible extent and maximize the number of coefficients that are statistically significant at conventional levels. Motivation for normalization is also supported by the observation of similar patterns of distributions between the original and normalized samples, as described in the following sections, and across event windows (see Figures and Tables in the Appendix D). Restricting the sample to one specific event window would pose small sample bias risks. 38% of the sample used one-day event windows, with high heterogeneity, and the most frequent event window $[\![-1;+1]\!]$ only accounts for 20% of the sample. Normalization also presents a greater consistency of meta-analyzed effect sizes, plus it enables including atypical event windows (so-called "exotic" event windows, illustrated by the frequency distribution of event windows in Figure D.1 of Appendix D),[19] bearing in mind that a set of variables controls for the specific features of each event window in the heterogeneity analysis (see Table 1).

Most Tables and Figures are displayed in the article or in Appendix for the two approaches labeled Panel A for the original sample and Panel B for the normalized sample.

3.3 Other features contributing to the interest of the dataset

First, meta-analyzing the literature investigating for market reactions to the disclosure of intentional financial crimes committed by listed firms enlarges to the maximum possible extent the sample of financial crimes – covering alleged to sanctioned financial crimes, through different procedures and disclosure channels, as illustrated in Figures E.1 and E.2 in Appendix E. This is the best way to circumvent partial observability of crimes (Ashton et al., 2021; Ormosi; 2014)., due the low share of detected crimes (Becker, 1968).

Second, this article contributes to the knowledge about crimes by overcoming the lack of international datasets, despite increasingly internationalized and interconnected financial markets.[20] Two thirds of sampled studies investigate crimes committed in the U.S., echoing the market size, a long history of enforcement, and the high regulatory transparency. The large geographical scope is an important dimension in our analysis to put U.S. results into perspective with 16 other Asian and European countries. Less than 4% of the studies meta-analyzed in this article conduct cross-country analyzes, for a few countries.

Third, we limit the scope of the surveyed studies to short-term event windows,

[18] 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.

[19] Event windows are qualified as "exotic" when they stand for less than 5% of the compounded event windows (*i.e.* less than 24 estimates), relevant to a fourth of the sample of abnormal returns (123).

[20] 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.

$[\![-10;+10]\!]$ around the event, because Kothari and Warner (1997) and Bhagat and Romano (2002a), among others, raised serious concerns about the specification and explanatory power of event studies with long-term event windows. The further from the event, the higher the noise-to-signal ratio and the higher the occurrence of other confounding events interfering with the investigated event.

Finally, the nature of the events supports the meaningfulness of primary studies. 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).[21] Our sample contains only "bad" news, which were proven to be 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, which can be investigated for by meta-analyzing this literature. This point is further accentuated due to the existing biases in the finance literature explaining the cross-section of expected stock returns (Harvey et al., 2016; Harvey, 2017).

3.4 Potential factors explaining heterogeneity among studies

In addition to the estimated (C)AARs and their statistical significances, we build a set of variables to account for the heterogeneity among studies, a heterogeneity introduced by the choices of authors of the primary studies, and for 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, such as Bajzík et al. (2020), Geyer-Klingeberg et al. (2020), and Zigraiova et al. (2021) to name a few. 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 Table 1.

Amid those coded variables, we select 22 characteristics of study design as potential sources of variability in the abnormal returns, and their correlation matrix is displayed in the Appendix F. This choice is grounded in the existing literature on the enforcement of financial regulations, on financial crimes, and on event studies. For ease of exposition, we sort the variables into the following three categories: the structural characteristics of each sample of financial crimes, the estimation characteristics of each event study and of the reported estimates, and the publication characteristics of each article, potentially related to quality, which are not captured by data and estimation characteristics.

3.4.1 Structural characteristics of financial crimes

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

Reactions to financial crimes can differ between countries, regions, and even legal origins (Djankov et al., 2008). Commercial laws 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 16 other countries in our sample. This proportion correlates with

[21] 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.

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 geographies, 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), levels of democracy (Shleifer, 2005), and sources of data (Karpoff et al., 2017). Therefore, we create a dummy variable *only U.S.* set to one when the event study focuses on financial crimes committed in the U.S., and zero otherwise (64% of the sample). As a robustness check, the estimations give similar results when enlarging the sample to *common law countries* (69% of the sample).

Since market reactions could have tamed across time (Panel B, Figure 1), 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 sampled financial crimes (*mid-point year*, in 2000 on average), which is also positively and significantly correlated to the year of publication of the article.

The heterogeneity between financial crimes also needs to be controlled for (see graphical illustrations in Appendix C and E). First, part of the literature surveys all violations of securities laws, for example by exhaustively investigating all the sanctions made by a given authority over a period of time (48% of the sample). Another strand of the literature focuses only on market abuses (50% of the sample), most often exclusively accounting frauds (in the U.S., 33% of the sample), or insider trading (10% of the sample). An accounting fraud typically leads to an accounting restatement, directly impacting shareholder's wealth, contrary to the disclosure of other violations of securities laws, except for fines. 20% of the sample investigates non-accounting violations of securities laws. Based on the above background, we kept the dummy variable for studies investigating *exclusively accounting frauds*. Second, the literature investigates market reactions to alleged, investigated, or condemned financial crimes, along the consecutive steps of enforcement (see Appendix E, Figure E.2). Frauds can be disclosed in newspaper articles, or in firm or regulatory statements (see Appendix E, Figure E.3). 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. In line with the semi-strong efficient market hypothesis (Fama, 1970), the very first hint of financial crimes (including alleged crimes) was shown to trigger 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) underline the persistent-in-time stigma that arises after a fraud allegation by stressing the difference between "not guilty" and "innocent": "even when no charges are ultimately brought [after SEC financial crime investigations], firms that voluntarily disclose an investigation have significant negative returns, underperforming non-sanctioned firms that stayed silent". We codify the *alleged frauds* as a dummy variable set to one when the crimes are not yet sanctioned (*i.e.* alleged and possibly investigated), and zero otherwise (61% of the sample). Third, each country has its own enforcement mix (see Appendix

E, Table E.1), 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 (possibly during the procedure) and liability standards differ as well. Complementarily, the channel of disclosure of financial crimes (by a newspaper article, the enforcer, or the firm itself) may impact the subsequent spillovers of the fraud. The media coverage of financial crimes can be positively 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 financial and business media can even be perceived by investors as a watchdog (Miller, 2006), as its 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). All in all, we use a dummy variable to control the origin of the source: *crimes disclosed in newspaper articles* (42% of the sample), as opposed to the other alternative being crimes disclosed by enforcers or the firms.[22]

3.4.2 Estimation characteristics of event studies and reported estimates

The estimation characteristics control for the main possible divergences in the implementation of the event study methodology in primary studies (see Appendix B for details), in particular regarding transparency, rigor, and depth of the analyses. First, the heterogeneity between samples is controlled for with the *number of sampled events* (as the log of 264 financial crimes sampled *per* article on average). Second, the rigor in the implementation of the event study methodology is captured by the following three dummy variables: 1) whether *the initial sample size* (of financial crimes) *is specified*, before cleaning the data from confounding events, not-daily-listed firms, duplicates, etc. (78% of the sample); 2) whether the article explicitly *excludes confounding events*, *i.e.* events concomitant with the financial crime disclosures that interfere with the abnormal return estimation (29% of the sample); and 3) whether the *estimation window* (over which the parameters of "normal" returns are estimated) *is specified* (72% of the sample). Third, four variables control for the event windows of the reported estimates: 1) the *length of the event window* of the estimated abnormal returns (4 days on average, 1 day for AAR and 2 up to 21 days for CAARs, given the limit put on the reported short-term estimates), as longer event windows might curb AARDs, to control for the normalization of effect sizes; whether the *event window* is 2) *strictly before* (14% of the sample) or 3) *on the event day* when the financial crime is revealed (*i.e.* $AAR(0)$, standing for 17% of the sample), 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"* (26% of the sample). The semi-strong efficient market hypothesis (Fama et

[22] 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).

al., 1969) implies that that the news should be fully priced-in on the day of its publication. Still, 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), which supports extending the event windows before the event itself. Additionally, it can take some time for the market to fully adjust to the news – this is reflected in frequent use of longer windows in event studies (longer than AAR[0] on the day of the event). Some authors might even be tempted to disclose results over atypical "exotic" event windows for which (cumulative) average abnormal returns are statistically significant, echoing the literature on *p*-hacking. Figure D.2 of Appendix D compares the funnels plots of the seven most frequently used event windows complemented with those "exotic" event windows, for which the publication bias appears to be particularly strong. Fourth, another key estimation characteristic is the statistical significance characterization, with a mere *statistical significance level ("stars")*, and/or *non-parametric tests*, most frequently rank tests (respectively 84% and 18% of the sample). Finally, some event studies are complemented with *cross-sectional regressions* (62% of the sample), to investigate for the drivers of abnormal market reactions, and/or *reputational penalty estimations* (9% of the sample), due to the previously described downgraded investors' expectations about the firms, after Karpoff and Lott (1993).[23]

3.4.3 Publication characteristics of the article

Since the sample is comprised of articles published in a wide range of peer-reviewed journals and of 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. As highlighted by Geyer-Klingeberg et al. (2020), the following three publication characteristics[24] are relevant for a meta-analysis: 1) the *number of authors* of the article (2.3 on average); 2) if the author contributed to more than one article in the sample (*multiple authorships*), as a way to assess his level of expertise (29% of the sample); and 3) whether the article was published in a *business journal* (25% of the sample). The field of research on the spillovers of financial crimes is at the intersection between economics, law, finance, accounting, and business. Being published in a more generalist and less-technical business, management, and organization journal could increase the visibility of the findings, but could also be synonym for a less stringent assessment criteria regarding the implementation of the event study methodology. Finally, as is typical for meta-analyses (e.g. Bajzík et al. (2021) or Gechert et al. (2022)), the following two variables can be expected to be correlated with the unobserved quality of the article, which might not be fully captured by the variables previously described: 1) the number of citations of the article recorded in Google Scholar per year since

[23] 14% of the sample mentioned the average cash fines imposed on the regulated persons in five countries (plus a group of European countries) and 9% of the sample conducted an estimation of the reputational penalty. There is a great heterogeneity in fines: 50mn USD for the U.S.; 61.4mn GBP for the U.K (USD 75.82mn late March 2023); 0.9mn EUR for France (USD 1.0mn late March 2023); 6.7bn JPY for Japan (USD 50.5mn late March 2023); and 38.7mn RMB for China (USD 5.6mn late March 2023).

[24] The year of publication is already controlled for by the mid-point year of the data.

publication (in log, *number of Google quotes per year since publication*) and 2) the *Scopus cite score* of the journal.

3.5 Descriptive statistics

The 480 abnormal returns estimated for 264 financial crimes on average (standing for a cumulated 32,500 financial crimes), compiled from the sample of 111 primary studies, provide a diverse set. Figures 2 depict the distributions of average abnormal returns by study. Most studies report as expected negative and statistically significant abnormal returns, with *naïve* averages for (C)AARs of -4.8% and for AARDs of -1.8% *per* day of the event window (-5.9% and -2.4% when weighting by the number of estimates reported *per* a study, as in Bajzík et al. (2020) and Balima and Sokolova (2021)).[25]

The following initial observations can be made. First, Tables 2 and Figures 3 give indications of the potential causes of heterogeneity in abnormal returns by comparing sub-samples. Markets react more to financial crimes (*i.e.* more negative (C)AARs than average and higher variance) with the following characteristics: exclusively accounting fraud (-6.9%), crimes committed in the U.S. (-6.5%, this also holds more generally in common-law countries), alleged crimes (-5.6%), and crimes directly disclosed by the firms (-5.5%) or in newspaper articles (-5%), whereas the type of procedures (public or private enforcement) does not matter.[26] Second, financial crimes were disclosed over a long time span, between 1965 and 2018, as illustrated in Figure 1. The upward trend of average abnormal returns 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 (Ripken, 2006). It could also result from quality improvements in the data and techniques along time. Third, 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,[27] 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 initial observations call for deeper analyses to confirm any potential systematic difference between subsamples of reported abnormal returns, to potentially correct for publication bias, and to account for potential correlations between explanatory variables.

4. Testing for publication bias

4.1. Publication bias and funnel plots

[25] 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, but is not proportional to the inverse variance of the standard errors. The results (available upon request) are consistent though lower when weighting by the interaction between the inverse number of estimates and the inverse variance.

[26] Similarly, AARDs exceed the average (-1.8%) for exclusively accounting fraud (-2.9%), crimes committed in the U.S. (-2.5%), alleged crimes (-2.9%), and crimes directly disclosed by the firms (-2.1%).

[27] This confirms Amiram et al.'s (2018) observation that studies on financial misconduct belong to three perspectives: law, accounting, and finance. For our sample, by declining order of importance, journals can be sorted as follows: finance, accounting, business, and law.

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 subsequent policy recommendations (Bom and Rachinger, 2019). Event studies can be easily subjected to p-hacking, a theme that is receiving increased attention (Brodeur et al., 2020 and Bruns and Ioannidis, 2016, among others). In fact, 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 of authors to search for negative abnormal returns in response to disclosed financial crimes, in line with the hypotheses of efficient markets and rational investors.

We construct funnel plots to graphically analyze the distribution of the reported effect size, which could illustrate a potential publication selection bias (Egger et al., 1997; Stanley and Doucouliagos, 2010). We plot the estimated average abnormal impacts of financial crimes on returns ((C)AARs for Panel A and AARDs for Panel B) on the horizontal axis against a measure of the estimate's precision (the inverse of the standard errors of the abnormal returns) 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 should form an inverted funnel, with the most precise estimates being closer to the true mean, and inversely less precise estimates being more dispersed. Additionally, the dispersion of effect sizes should be negatively correlated with the precision of the estimate. Figures 4.1 compare the distributions of the abnormal returns against their precisions for the whole sample, complemented with sub-samples by countries (Figures 4.2 compare the U.S. to other countries),[28] and by types of financial crime (Figures 4.3 compare exclusively accounting frauds to more generally violations of securities laws, possibly including accounting frauds when investigating all sanctions made by a given authority or all settlements made in a given country), similarly to Griffin et al. (2004), Pritchard and Ferris (2001), and Lane (2016) among others.

Funnel plots are systematically skewed to the left, confirming average negative abnormal returns: disclosed financial crimes are informational and priced in by market participants. What is more puzzling is that the distribution of abnormal returns is clearly asymmetrical to the left (towards more negative abnormal returns). 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 graphical illustrations in Appendix G) and, to a lesser extent, on accounting frauds. Complementarily, and echoing the literature on p-hacking, the histograms of the distribution of the t-statistics (frequency and Kernel densities displayed in Figures 5) indicate jumps in the

[28] 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.

distributions at the critical significance levels (5% and 1%). They also suggest that the main source of publication bias is the underreporting of positive abnormal returns in the literature, even if the true effect of the disclosure of financial crimes is negative.

4.2. Quantification of the publication bias and the true effect of disclosed 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 effects and their standard errors:

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

where ARs are the reported effects, the average abnormal returns estimated over an event window ((C)AARs) or *per* day of the event window (AARDs), SE are the standard errors of the ARs, β_0 and β_1 are the parameters to be estimated, i and j denote the i^{th} estimate from the j^{th} study ($i \in [\![1; 16]\!]$, $j \in [\![1; 111]\!]$), 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 Figures 4). The estimated intercept between the ARs 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 3 for the original sample of (C)AARs (column [1]) and for the normalized AARDs (column [2]). To support the robustness of the results, we compare three types of estimation technique following recent testing innovations. First, Panel 1 is based on unweighted data using the following approaches: 1) a baseline OLS regression; 2) 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); 3) a regression using between-study variance; 4) a hierarchical Bayes (Bajzík et al., 2020); and 5) instrumenting for the standard error with the number of observations reported by study (Havránek and Sokolova, 2020).[29] Second, Panel 2 uses a weighted-least-squares model of Panel 1 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)[30] and 2) by the inverse of the number of estimates reported by the study, to give an equal weight to every study whatever the number of estimates. Third, Panel 3 uses recent non-linear estimation techniques. These techniques relax the implicit assumption made in Panels 1 and 2 that the publication bias is a linear function of

[29] Estimated abnormal returns 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.

[30] 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.

standard errors. This third 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 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;[31] 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., 2022) to minimize the tradeoff between variance and bias.[32] 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 3 confirms the significant publication bias in the analyzed literature towards negative estimates of abnormal returns hinted at by the funnel plots. Whatever the sample of abnormal returns (original or normalized) or the estimation method (Panels 1, 2, and 3), standard errors clustered by studies impact highly statistically significantly and negatively abnormal returns. This publication bias is particularly high when instrumenting for the standard error with the number of observations reported by the study (Panel 1.5), when weighting by precision (panel 2.1), and when adding study-level fixed effects (Panel 1.2). Additionally, the genuine underlying empirical effect beyond the distortion due to publication bias is negative and statistically significant, but much more limited than the *naïve* averaged estimates shown in Table 1 (specially for non-linear techniques). Most of the abnormal returns reported in the primary 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 abnormal returns is more than three times lower than the *naïve* simple means of the reported estimates. Specifically, after the disclosure of financial crimes, on average and across panels, abnormal returns would represent a loss of -1.3% over the event window (column [1]) and -0.5% *per* day of the event window, implying a cumulative loss of -2.1% over the average 4-day event window (column [2]). These abnormal returns are three times smaller than the initially observed average values of -4.8% and -1.8% that do not account for publication bias.[33] It is also worth stressing that the estimated effect beyond bias is

[31] Complementary results for the selection model of Andrews and Kasy (2019) are displayed in Appendix H, with funnel plots and histograms of Z-statistics for the full sample of financial crimes and sub-samples.
[32] The results for the Hedges test are detailed in Appendix I, with similar conclusions.
[33] 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

higher for linear estimators than for non-linear estimators (respectively for (C)AARs: -1.7% and -0.4% and for AARDs: -0.6% and -0.2%). Finally, the publication bias towards negative abnormal returns is consistent across event windows, and in particular for "exotic" event windows (see funnel plots and MRA results by event windows in Appendix D).[34]

Further, the sample of average abnormal returns *per* day (AARDs) is split into the four sub-samples, echoing the previous observations of heterogeneity in the sample: crimes committed in the U.S. or in other countries (columns [3] and [4] of Table 3) and exclusively accounting frauds or violations of securities laws (columns [5] and [6]). When exploring more detailed results grounded in these specific sub-samples, the following three conclusions can be drawn. First, the publication bias for crimes committed in the U.S. is more than twice as high as for other countries (for which there is even no publication bias estimated from the simple OLS, from instrumenting for the standard error with the number of observations reported by the study, and from weighting with the inverse of the number of estimates). The result holds when enlarging the sample to common-law countries (as opposed to code-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 more than two times on average).[35] 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, though accounting frauds and violations of securities laws suffer from a similar significant publication bias, returns corrected for the bias following the publication of accounting frauds would contract three times more than following violations of securities laws (by -0.95% and -0.35% *per* day, respectively). Stronger reactions to (intentional) accounting frauds can be explained by the direct impacts on the Profit & Loss statements subsequent to accounting restatements.

All in all, we find robust evidence of publication bias in the literature towards reporting negative abnormal returns, particularly marked in the U.S., and of genuine empirical evidence in the collected estimates: markets penalize listed firms for engaging in intentional financial crimes (particularly accounting frauds), 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 the next section. The rest of the article will focus on the normalized sample by the length of the event windows (AARDs) given the consistency of the results along the dual approach. The normalization also presents the great advantages of greater consistency between event windows and of enlarging the sample of financial crimes included in the meta-analysis to the maximum possible extent, compared to

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.

[34] On an individual basis, the MRA results by event windows must be analyzed cautiously given the limited sample size (43 observations from 28 studies for AAR[+1] and 52 observations from 43 studies for AAR[-1]).

[35] AARDs contract on average by -0.70% *per* day in the U.S. compared with -0.28% in other countries.

using subsamples by specific event windows. The results based on the original sample of (C)AARs will be systematically referenced to and included in the Appendix section.

5. Why do market reactions vary among studies?

This section is the first attempt to statistically explain sources of heterogeneity among studies. 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.

5.1. 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 22 control variables vector X detailed in section 3.4. Hence, the specification quantifying the sources of heterogeneity between every observation i ($i \in [\![1;16]\!]$) surveyed from study j ($j \in [\![1;111]\!]$) is defined as:

$$AARD_{i,j} = \beta_0 + \beta_1 SE_{i,j} + \sum_{k=1}^{22} \gamma_k X_{k,i,j} + \varepsilon_{i,j}. \qquad (2)$$

However, the number of factors described in Section 3.4 that potentially affect heterogeneity among studies is large and their inclusion in a regression might be problematic. Some variables are supported by strong rationale, theory, or enforcers' practice, 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 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 22 variables to account for structural variations, the event study estimation, and the publication of the study their correlation matrix is displayed in the Appendix F. Three methods outlined below are used to 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), the model uncertainty is circumvented by using Bayesian model averaging (BMA), displayed in Figure 6. BMA runs numerous regression models with different subsets of the 2^{22} possible combinations of explanatory variables to detect the most likely models.[36] 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, 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

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

as one observation in the data. Additionally, as in Bajzík et al. (2020), we use the dilution prior (George, 2010) adjusting the model probabilities by the determinant of the correlation matrix of the variables included in the model, to alleviate potential collinearity between the 22 explanatory variables.

Second, as in Gechert et al (2022), we run a hybrid frequentist-Bayesian model as a robustness check. The ten variables deemed unimportant based on their posterior probabilities of inclusion (PIPs) from a BMA are excluded (Eicher et al, 2011).[37] The resulting model with twelve deemed non-negligible variables is estimated using a standard OLS technique, clustering standard errors at the study level.

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^{22} potential models.

Complementarily to these baseline estimates, robustness checks include weighted alternatives by the number of estimates reported *per* study and by the precision, displayed in Figures J.1 of Appendix J. A sensitivity analysis to the sets of priors used in the BMA developed in the next section is depicted in Figure 7. Our baseline set of priors is compared to three other sets used recently in the literature to check the robustness of the results: 1) 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); 2) 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 3) 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).

5.2.Results

Figure 6 illustrates graphically the results of the Bayesian Model Averaging.[38] The vertical axis depicts the standard error and the 22 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.

The estimation results are reported in Table 4, with the Bayesian Model Averaging in column [1], the OLS frequentist check in column [2], and the Frequentist Model Averaging in column [3]. All models we run confirm the prevalence of a publication bias in the literature on the spillovers of financial crimes, with across-the-board significant coefficients on the standard error. This demonstrates that the reported abnormal returns following the disclosure of financial

[37] 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.

[38] The BMA results for the original sample of non-normalized (C)AARs are displayed in Appendix J, Figure J.1.

crimes 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 -1.6 to -1.4. Additionally, Figure 7 depicts a limited sensitivity of our results to the priors. Our baseline priors are relatively conservative though all the priors point to similar results, as the PIPs rankings of the variables are broadly preserved, possibly with higher PIPs. In total, the four sets of priors tend to converge in the twelve control variables with significant PIPs.

Structural characteristics. The evidence on the systematic importance of structural characteristics is 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 fact that crimes are disclosed in the U.S. is almost statistically significant for the FMA with a negative coefficient, pointing to higher response in the market. 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, which also reflects the year of publication of the article) would curb market reactions is not confirmed by the BMA nor the FMA, when controlling for study design. Finally, the channel of the disclosure of the crime, through *newspaper articles*, is not either significant.[39] As a robustness check, we added the dummy variable of public enforcement procedures, which did not change the results and turned also insignificant, consistently with initial observations.

Estimation characteristics. Studies with a greater *number of sampled financial crimes* produce more negative estimates of AARDs, which might reflect a small-sample bias. Conversely, regarding the event windows of the reported estimates, our results suggest that the event windows *strictly before the event* and *"exotic" event windows* are associated with lower reported AARDs. This echoes the observations of some anticipation of news stressed in the literature, though the greatest part of the abnormal reaction occurs on the event day and after. Such anticipation might be to some extent country- and event-specific. The significance of *"exotic" event windows* assorted with a positive coefficient underlines the relevance of the *p*-hacking for this field of the literature, when authors investigate numerous "exotic" event windows to conclude with significant – though less meaningful or conventional in economic sense – abnormal returns. Finally, as expected, AARDs estimated on the *event day* are close to significant to exhibit as expected economic effect with more negative results, contrary to the

[39] As a robustness check, the other two channels (enforcers or firms) were also envisioned, with similarly insignificant results.

length of the event windows, both economically and statistically insignificant. The longer the event window, the lower the impact *per* day of the event window, under the efficient market hypothesis.

Four of the seven estimation characteristics controlling for the quality and the rigor of the event study methodology application in every study[40] are economically and statistically significant: the *specification of the estimation window*, the qualification of the *statistical significance with stars* and/or with *non-parametric tests*, and when the event study is enriched with a *cross-sectional regression*. Conversely to the negative impact of *specification of the estimation window*, the variables controlling for how the sample is built (*initial sample size specification* and *exclusion of confounding events*) do not economically and significantly explain heterogeneity. This might reflect a search for brevity and the limited attention brought to the description of the data in the primary studies, despite the fact that the data was duly collected and cleaned, a prerequisite to any quality event study. The baseline specification suggests that higher (or less negative) estimated AARDs are characterized with statistical significance levels ("*stars*") and/or *non-parametric tests*. Additionally, the fact that an event study is complemented by further analyses with *cross-sectional regressions* is statistically significant and negatively correlated with the magnitude of AARDs, echoing the fact that the most prominent articles encourage such estimations (such as Karpoff and Lott, 1993, 2008; Armour et al, 2017) – contrary to *reputational penalty estimations*, bearing in mind the small sample size.

Publication characteristics. Our results suggest that co-authored articles (*nb authors*) exhibit higher *AARDs* (more negative), even though we do not find any significant effect of expertise with *multiple authorships.* Conversely, articles published in possibly less technical *business journals* tend to conclude with smaller estimated AARDs. Our results also stress that the *Google Scholar quotes per year since publication* are robustly and negatively correlated with the reported AARDs. Hence, 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 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.

5.3.Implied AARDs

The main takeaways of the previous investigations are that 1) the reported abnormal returns following the disclosure of intentional financial crimes are exaggerated by a publication bias and that 2) these abnormal returns vary systematically depending on the types of financial crime, on the estimation characteristics, and on the publication characteristics. Even though the scope of the studies covers 17 countries, the great majority investigates the U.S. and these

[40] 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 inclusion of z-statistics; 6) the use of so-called "exotic" event windows, as a strategy to publish statistically significant AARDs; 7) complementary analyses are undergone with a cross-sectional regression; and 8) a reputational penalty estimation.

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 echoing the BMA results: (i) the quality of the outlet (high-quality journals with a RePEc impact factor above the average of the sample, *i.e.* above the *International Review of Law and Economics,* which stands for a reference field journal,[41] or medium-quality journals for articles published in a peer-reviewed journal), and (ii) the rigor in the application of the event study methodology (with the disclosure of the estimation window and the inclusions of non-parametric statistical tests). We then regress the reported AARD on a confidence dummy and on the standard error of the estimate. Results are shown in Table 5 for the full sample comparing different levels of "confidence" (Panel 1) and for subsamples by geography (Panel 2), by types of events (Panel 3), and by specifications of the event study methodology (Panel 4). 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 1 depending on the definition of confidence. The resulting significant mean AARDs range from -1.4% up to -0.5% for the lowest confidence, which exceed the simple mean AARDs corrected for the publication bias (-0.5%, in Table 3) but remain lower than the initial *naïve* estimated impact (-1.8%). Given the substantial effect of the variable "lower confidence", Panel 1 stresses that better estimates – according to our definitions – tend to be significantly more negative than the less reliable ones. Such results corrected for the publication bias appear in line with abnormal market reactions following other types of white-collar crimes (Karpoff, 2012).[42] Digging into the details of the subsamples, the following conclusions can be drawn. Differences between countries and financial crimes are confirmed, with much more responsive financial markets in the U.S. (-1.6%) than in Europe (-0.8%) and, to a lesser extent, in emerging economies (-0.5%), for accounting frauds (-2.3%) and alleged frauds (-1.8%). Interestingly, and in line with previous results, studies publishing estimates for "exotic" event windows and using statistical significance levels to conclude with less responsive financial markets (-0.5% and -0.2%, respectively), conversely to studies disclosing non-parametric tests (-1.8%).

6. Conclusions

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

[41] We proxy peer-review quality according to the recursive discounted impact factor from the Web of Science (article influence score). As the threshold for high-quality peer-review, we follow Bajzík et al. (2020) and choose the International Review of Law and Economics, which published several influential contributions in this field of literature but still leaves enough better-ranked journals to allow for sufficient variation in the confidence dummy variable.

[42] To name a few examples: three-day cumulative abnormal returns of -2.3% subsequent to antitrust charges (van den Broek et al., 2010); two-day cumulative abnormal returns of -1% subsequent to environmental violations (Karpoff et al., 2005), two-day cumulative abnormal returns of -0.32% subsequent to U.S. car recalls (Barber and Darrough, 1996); one-day abnormal returns of -1.7% subsequent to air safety disasters due to pilot errors (Mitchell and Maloney, 1989).

are estimated as average abnormal returns around the disclosure of such crimes. Subsequent market reactions are key in terms of enforcement as financials markets can complement or even substitute for enforcers by imposing reputational penalties and consequently by encouraging best practices, possibly more quickly and at lower cost. This reflects a regulatory shift towards the "naming and shaming" of financial criminals, for the less serious but still weighty crimes.

We examine a total of 480 estimates of abnormal returns following the publication of 32,500 intentional financial crimes committed by listed firms that were reported in 111 research studies. We perform a meta-analysis to examine the relationship between these abnormal returns and the characteristics of the sample of misconducts under review, of the estimations, and of the publication. We propose an original dual approach, based on the original sample of estimates collected from the primary studies, complemented by a normalized sample by the length of the respective event windows. This dual approach enables using the widest possible sample with concurrent conclusions.

The results of the meta-analysis reveal a strongly negative publication selection bias in the literature, which is in line with the *a priori* hypothesis of efficient markets and rational investors: markets are expected to react negatively to bad news like the disclosure of financial crimes (Karpoff et al., 2008). Our results (funnel asymmetry tests, meta-regression analyses, 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 more than 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. Such crimes are bad news regarding the firms, 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. 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. Conversely, the channel through which the crimes is disclosed, and the types of procedures do not significantly impact market reactions.

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, impact factors, and methodological rigor. We find the existence of a robust correlation between the quality of the estimates and their magnitude: more quality studies report more negative AARDs, ranging from -0.5% to -1.4% *per* day of the event window, 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 are afraid of being associated with alleged or condemned financial crimes. Consequently, the mere threat of sanctions (and their 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 theoretically for financial institutions by Chakravarty et al. (2021). Regulators may alternatively choose a lighter touch, possibly with anonymized sanction decisions or with confidential bilateral procedures. Another path for enforcers may be to rely on financial markets to complement or even substitute their enforcement actions by imposing reputational penalties. In such case, our results stress that regulatory transparency can be an efficient regulatory tool *per se* since significant negative abnormal returns follow the allegation of financial crimes, in particular when committed in the U.S. and other common-law countries and for accounting frauds, contrary to regulatory procedures. Enforcers could (for example) communicate during enforcement procedures to warn investors and encourage best practices and even substitute sanctions with "name and shame" strategies to punish and better financial crimes, more quickly and at lower cost than with long enforcement procedures. That way, market participants could better price financial crimes, should the enforcers' objective be that markets account for their communication 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, Abdullah, Jonathan M. Karpoff, Jennifer M. Koski, and Gerald D. Martin. 2020. "The Prevalence and Costs of Financial Misrepresentation." *Working paper*, available at SSRN 3532053.

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

Amini, Shahram M., and Christopher F. Parmeter. 2012. "Comparison of Model Averaging Techniques: Assessing Growth Determinants." *Journal of Applied Econometrics* 27(5):870-876.

Amiram, Dan, Zahn Bozanic, James D. Cox, Quentin Dupont, Jonathan M. Karpoff, and Richard Sloan. 2018. "Financial Reporting Fraud and Other Forms of Misconduct: A Multidisciplinary Review of the Literature." *Review of Accounting Studies* 23(2):732-783.

Andrews, Isaiah, and Maximilian Kasy. 2019. "Identification of and Correction for Publication Bias". *American Economic Review* 109(8):2766-2794.

Armour, John, Colin Mayer, and Andrea Polo. 2017. "Regulatory Sanctions and Reputational Damage in Financial Markets." *Journal of Financial and Quantitative Analysis* 52(4):1429-1448.

Ashton, John, Tim Burnett, Ivan Diaz-Rainey, and Peter Ormosi, P. 2021. "Known Unknowns: How Much Financial Misconduct is Detected and Deterred?." *Journal of International Financial Markets, Institutions and Money* 74:101389.

Aupperle, Kenneth E., Archie B. Carroll, and John D. Hatfield. 1985. "An Empirical Examination of the Relationship Between Corporate Social Responsibility and Profitability." *Academy of Management Journal* 28(2):446-463.

Bajzík, Josef, Tomáš Havránek, Zuzana Irsova, and Jiri Schwarz. 2020. "Estimating the Armington Elasticity: The Importance of Study Design and Publication Bias." *Journal of International Economics* 127:103383.

Balima, Hippolyte W., and Anna Sokolova. 2021. "IMF Programs and Economic Growth: A Meta-Analysis." *Journal of Development Economics* 153: 102741.

Ball, Ray, and Philip Brown. 1968. "An Empirical Evaluation of Accounting Income Numbers." *Journal of Accounting Research* 6(2):159-178.

Barber, Brad M., and Masako N. Darrough. 1996. "Product Reliability and Firm Value: The Experience of American and Japanese automakers, 1973-1992." *Journal of Political Economy* 104(5):1084-1099.

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

Barber, Brad M., and Terrance Odean. 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, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76:169-217.

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

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

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

Bhaskar, Krish, John Flower, and Rod Sellers. 2019. "Financial Failures and Scandals: From Enron to Carillion." Routledge, London.

Black, Barbara. 2000. "The Legal and Institutional Preconditions for Strong Securities Markets." *UCLA Law Review* 48:781-855.

Blanco-Perez, Cristina, and Abel Brodeur. 2020. "Publication Bias and Editorial Statement on Negative Findings." *The Economic Journal* 130(629):1226-1247.

Boehmer, Ekkehart, Jim Musumeci, and Annette B. Poulsen. 1991. "Event Study Methodology under Conditions of Event-Induced Variance." *Journal of Financial Econometrics* 30:252-273.

Bom, Pedro R. D., and Heiko Rachinger. 2019. "A Kinked Meta-Regression Model for Publication Bias Correction." *Research Synthesis Methods* 10(4):497-514.

Bonini, Stefano, and Diana Boraschi. 2012. "Corporate Scandals and Capital Structure. In Entrepreneurship, Governance and Ethics" (241-269). Springer, Dordrecht.

Brodeur, Abel, Amthias Lé, Marc Sangnier, and Yanos Zylberberg. 2016. "Star Wars: The Empirics Strike Cack." *American Economic Journal: Applied Economics* 8(1):1-32.

Brodeur, Abel, Nikolai Cook, and Anthony G. Heyes. 2020. "Methods Matter: P-Hacking and Causal Inference in Economics." *American Economic Review* 110(11):3634-60.

Bruns, Stephan B., and John P. A. Ioannidis. 2016. "p-Curve and p-Hacking in Observational Research." *PLoS One* 11(2):e0149144.

Chakravarty, Surajeet, Lawrence Choo, Miguel A. Fonseca, and Todd R. Kaplan. 2021. "Should Regulators Always Be Transparent? A Bank Run Experiment." *European Economic Review 136*:103764.

Choi, Stephen, and Marcel Kahan. 2007. "The Market Penalty for Mutual Fund Scandals." *Boston University Law Review* 87:1021-1057.

Christensen, Hans B., Luzi Hail, and Christian Leuz. 2016. "Capital-Market Effects of Securities Regulation: Prior Conditions, Implementation, and Enforcement." *The Review of Financial Studies* 29(11):2885-2924.

Cummins, J. David, Christopher M. Lewis, and Ran Wei. 2006. "The Market Value Impact of Operational Loss Events for US Banks and Insurers." *Journal of Banking and Finance* 30(10):2605-2634.

de Batz, Laure. (2020). "Financial Impact of Regulatory Sanctions on Listed Companies." *European Journal of Law and Economics* 49(2):301-337.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 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, Hemang, Chris E. Hogan, and Michael S. Wilkins. 2006. "The Reputational Penalty for Aggressive Accounting: Earnings Restatements and Management Turnover." *The Accounting Review* 81(1):83-112.

Djama, Constant. 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, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. "The Law and Economics of Self-Dealing." *Journal of Financial Economics* 88(3):430-465.

Doucouliagos, Hristos., and Tom D. Stanley. 2013. "Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity." *Journal of Economic Surveys* 27(2):316-339.

Dolley, James Clay. 1933. "Characteristics and Procedure of Common Stock Split-Ups." *Harvard Business Review* 11(3):316-326.

Dupont, Quentin, and Jonathan M. Karpoff. 2019. "The Trust Triangle: Laws, Reputation, and Culture in Empirical Finance Research." *Journal of Business Ethics* 163:1-22.

Dyck, Alexander, Adair Morse, and Luigi Zingales. 2010. "Who Blows the Whistle on Corporate Fraud?" *The Journal of Finance* 65(6):2213-2253.

Egger, Matthias, George Davey Smith, Martin Schneider, and Christopher Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *The British Medical Journal* 315(7109):629-634.

Eicher, Theo S., Chris Papageorgiou, and Adrian E. Raftery. 2011. "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants." *Journal of Applied Econometrics* 26(1):30-55.

Endrikat, Jan. 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, Peter Jan. 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, Peter Jan., and Marc Van Essen. 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, 55-74): Springer Netherlands.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25(2):383-417.

Fama, Eugene F. 1990. "Contract Costs and Financing Decisions." *The Journal of Business* 63(1)-S71-S91.

Fama, Eugene F., Lawrence Fisher, Michael C. Jensen, and Richard Roll. 1969. "The Adjustment of Stock Prices to New Information." *International Economic Review* 10(1):1-21.

Fang, Lily. H., and Joel Peress. 2009. "Media Coverage and the Cross-Section of Stock Returns." *The Journal of Finance* 64(5):2023-2052.

Fang, Lily H., Joel Peress, and Lu Zheng. 2014. "Does Media Coverage of Stocks affect Mutual Funds' Trading and Performance." *The Review of Financial Studies* 27(12):3441-3466.

Feldkircher, Martin. 2012. "Forecast Combination and Bayesian Model Averaging: A Prior Sensitivity Analysis." *Journal of Forecasting* 31(4):361-376.

Feldkircher, Martin, and Stefan Zeugner. 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, Carmen, Eduardo Ley, E., and Mark F. J. Steel. 2001. "Benchmark Priors for Bayesian Model Averaging." *Journal of Econometrics* 100(2):381-427.

Feroz, Ehsan H., Kyungjoo Park, and Victor S. Pastena. 1991. "The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases." *Journal of Accounting Research* 29:107-142.

Frooman, Jeff. 1997. "Socially Irresponsible and Illegal Behavior and Shareholder Wealth: A Meta-Analysis of Event Studies." *Business & Society* 36(3):221-249.

Furukawa, Chishio. 2019. "Publication Bias Under Aggregation Frictions: From Communication Model to New Correction Method." *Working Paper, Massachusetts Institute of Technology*.

Gande, Amar, and Craig M. Lewis. 2009. "Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers." *Journal of Financial and Quantitative Analysis* 44(4):823-850.

Garoupa, Nuno. 2001. "Optimal Magnitude and Probability of Fines." *European Economic Review 45*(9):1765-1771.
Gechert, Sebastian, Tomáš Havránek, Zuzana Irsova, and Dominika Kolcunova. 2022. "Measuring Capital-Labor Substitution: The Importance of Method Choices and Publication Bias." *Review of Economic Dynamics 45:55-82*.
George, Edward I. 2010. "Dilution Priors: Compensating for Model Space Redundancy." In Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown (158-165). *Institute of Mathematical Statistics*.
Geyer-Klingeberg, Jerome, Markus Hang, and Andreas Rathgeber. 2020. "Meta-Analysis in Finance Research: Opportunities, Challenges, and Contemporary Applications." *International Review of Financial Analysis* 71:101524.
Goldman, Eitan, Urs Peyer, and Irina Stefanescu. 2012. "Financial Misrepresentation and its Impact on Rivals." *Financial Management* 41(4):915-945.
Guiso, Luigo, Paola Sapienza, and Luigi Zingales. 2008. "Trusting the Stock Market." *The Journal of Finance* 63(6):2557-2600.
Hansen, Bruce E. 2007. "Least Squares Model Averaging." *Econometrica* 75(4):1175-1189.
Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2016. "... and the Cross-Section of Expected Returns." *The Review of Financial Studies* 29(1):5-68.
Harvey, Campbell R. 2017. "Presidential Address: The Scientific Outlook in Financial Economics." *The Journal of Finance* 72(4), 1399-1440.
Haslem, Bruce, Irena Hutton, and Aimee Hoffmann Smith. 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, Tomáš, and Zuzana Irsova. 2017. "Do Borders Really Slash Trade? A Meta-Analysis." *IMF Economic Review* 65(2):365-396.
Havránek, Tomáš, Tom D. Stanley, Hristos Doucouliagos, Pedro Bom, Jerome Geyer-Klingeberg, Ichiro Iwasaki, ... and R. C. M. van Aert. 2020. "Reporting Guidelines for Meta-Analysis in Economics." *Journal of Economic Surveys* 34(3):469-475.
Havránek, Tomáš, and Anna Sokolova. 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, Delvin D. 1991. "Business Ethics and Social Responsibility in Finance Instruction: An Abdication of Responsibility." *Journal of Business Ethics* 10(9):711-721.
Hennes, Karen M., Andrew J. Leone, and Brian P. Miller. 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, Jérôme, Christine Louargant, Patrice Laroche, and Jean-Noël Ory. 2019. "How Do Rating Agencies' Decisions Impact Stock Markets? A Meta-Analysis." *Journal of Economic Surveys* 33(4):1173-1198.
Ioannidis, John P., Tom D. Stanley, and Hristos Doucouliagos. 2017. "The Power of Bias in Economics Research." *The Economic Journal* 127(4):236-265.
Jackson, Howell E., and Mark J. Roe. 2009. "Public and Private Enforcement of Securities Laws: Resource-Based Evidence." *Journal of Financial Economics* 93(2):207-238.
Jo, Ara. 2021. "Culture and Compliance: Evidence from the European Union Emissions Trading Scheme." The Journal of Law and Economics 64(1): 18-205.
Jones, Kari, and Paul H. Rubin. 2001. "Effects of Harmful Environmental Events on Reputations of Firms. *In Advances in financial economics. Emerald Group Publishing Limited*.
Kahan, Dan M., and Eric A. Posner. 1999. "Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines." *The Journal of Law and Economics* 42(S1):365-392.
Karpoff, Jonathan 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, Jonathan M. 2020. "Financial Fraud and Reputational Capital." In Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation, Chapter 6, 153-177.
Karpoff, Jonathan M., Allison Koester, D. Scott Lee, and Gerald S. Martin. 2017. "Proxies and Databases in Financial Misconduct Research." *The Accounting Review* 92(6):129-163.
Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin. 2008. "The Cost to Firms of Cooking the Books." *Journal of Financial and Quantitative Analysis* 43(3):581-611.
Karpoff, Jonathan M., and John R. Lott, Jr. 1993. "The Reputational Penalty Firms Bear from Committing Criminal Fraud." *The Journal of Law and Economics* 36(2):757-802.
Karpoff, Jonathan M., John R. Lott, Jr, and Eric W. Wehrly. 2005. "The Reputational Penalties for Environmental Violations: Empirical Evidence." *The Journal of Law and Economics* 48(2):653-675.
Kočenda, Evžen, and Ichiro Iwasaki. 2021. "Bank Survival around the World: A Meta-Analytic Review." *Journal of Economic Surveys* 36(1), 108-156.
Kolari, James W., and Seppo Pynnönen. 2010. "Event Study Testing with Cross-Sectional Correlation of Abnormal Returns." *The Review of Financial Studies* 23(11):3996-4025.

Kothari, Sabino P., Susan Shu, and Peter D. Wysocki. 2009. "Do Managers Withhold Bad News?" *Journal of Accounting Research* 47(1):241-276.

Kothari, Sabino P., and Jerold B. Warner. 2008. "Econometrics of Event Studies. In Handbook of Empirical Corporate Finance 1," Elsevier, Chapter 1, 3-36.

Kothari, Sabino P., and Jerold B. Warner. 1997. "Measuring Long-Horizon Security Price Performance." *Journal of Financial Economics* 43(3):301-339.

Lane, Tom. 2016. "Discrimination in the Laboratory: A Meta-Analysis of Economics Experiments." *European Economic Review* 90:375-402.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2006. "What Works in Securities Laws?" *The Journal of Finance* 61(1):1-32.

Leuz, Christian, Dhananjay Nanda, and Peter D. Wysocki. 2003. "Earnings Management and Investor Protection: An International Comparison." *Journal of Financial Economics* 69:505-527.

Lev, Baruch, Stephen G. Ryan, and Min Wu. 2008. "Rewriting Earnings History." *Review of Accounting Studies* 13(4):419-451.

Lewis, Jeffrey B., and Drew A. Linzer. 2005. "Estimating Regression Models in Which the Dependent Variable is Based on Estimates." *Political Analysis* 13(4):345-364.

Ley, Eduardo, and Mark F.J. Steel. 2009. "On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression." *Applied Economics* 24:651-674.

Liang, Hao, and Luc Renneboog. 2017. "On the Foundations of Corporate Social Responsibility." *The Journal of Finance* 72(2):853-910.

Lins, Karl V., Henri Servaes, and Ane Tamayo. (2017). "Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility During the Financial Crisis." *The Journal of Finance*, 72(4):1785-1824.

Liu, Chelsea, and Alfred Yawson. 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. Craig. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35(1):13-39.

Marascuilo, Leonard A., and Ronald C. Serlin. 1988. "Statistical Methods for the Social and Behavioral Sciences." *WH Freeman/Times Books/Henry Holt & Co*.

Matousek, Jindrich, Tomáš Havránek, and Zuzana Irsova. 2021. "Individual Discount Rates: A Meta-Analysis of Experimental Evidence." *Experimental Economics*, 1-41.

Miller, Gregory S. 2006. "The Press as a Watchdog for Accounting Fraud." *Journal of Accounting Research* 44:1001-1033.

Mitchell, Mark L., and Michael T. Maloney. 1989. "Crisis in the Cockpit? The Role of Market Forces in Promoting Air Travel Safety." *The Journal of Law and Economics* 32(2-1): 329-355.

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, and The Prisma Group. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement." *PLoS med*, 6(7):e1000097.

Morris, Brandon C., Jared F. Egginton, and Kathleen P. Fuller. 2019. "Return and Liquidity Response to Fraud and SEC Investigations." *Journal of Economics and Finance* 43(2):313-329.

Murphy, Deborah L., Ronald E. Shrieves, and Samuel L. Tibbs. 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. Khalid, Atul Rai, and Semih Tartaroglu. 2014. "Market Reactions to Wells Notice: An Empirical Analysis." *International Journal of Disclosure and Governance* 11(2):177-193.

Nelson, Christine, Sara Gilley, and Garett Trombley Esq. 2009. "Disclosures of SEC Investigations Resulting in Wells Notices." *Securities Litigation Journal* 19(4):19-21.

Nourayi, Mahmoud M. 1994. "Stock Price Responses to the SEC's Enforcement Actions." *Journal of Accounting and Public Policy* 13(4):333-347.

Ormosi, Peter L. 2014. "A Tip of the Iceberg? The Probability of Catching Cartels." *Journal of Applied Econometrics* 29(4):549-566.

Ozeki, Norimasa. 2019. "Determinants of Market Reaction to Disclosure of Accounting Misconduct: Evidence from Japan." *Securities Analysts Journal* 57(3):72-84.

Palmrose, Zoe Vonna, Vernon J. Richardson, and Susan Scholz. 2004. "Determinants of Market Reactions to Restatement Announcements." *Journal of Accounting and Economics* 37:59-89.

Parsons, Christopher A., Johan Sulaeman, and Sheridan Titman. 2018. "The Geography of Financial Misconduct." *The Journal of Finance* 73(5):2087-2137.

Peltzman, Sam. 1976. "Toward A More General Theory of Regulation." *The Journal of Law and Economics* 19(2):211-240.

Peltzman, Sam. 1981. "The Effects of FTC Advertising Regulation." *Journal of Law and Economics*, 24(3):403-448.

Peress, Joel. 2014. "The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes." *The Journal of Finance* 69:2007-2043.

Pritchard, Adam C., and Stephen P. Ferris. 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, Adrian E., David Madigan, Jennifer Hoeting. 1997. “Bayesian Model Averaging for Linear Regression Models.” *Journal of the American Statistical Association* 92:179-191.

Reurink, Arjan. 2018. “Financial Fraud: A Literature Review.” *Journal of Economic Surveys* 32(5):1292-1325.

Ripken, Susanna Kim. 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, Marek, Tomáš Havránek, and Roman Horváth. 2013. “How to Solve the Price Puzzle? A Meta-Analysis.” *Journal of Money, Credit and Banking* 45(1):37-70.

Sapienza, Paola, and Luigi Zingales, 2012. “A Trust Crisis.” *International Review of Finance* 12(2):123-131.

Sharpe, William F. 1970. “Portfolio Theory and Capital Markets.” McGraw-Hill, New York.

Shleifer, Andrei. 2005. “Understanding Regulation.” *European Financial Management* 11(4):439-451.

Sokolova, Anna, and Todd Sorensen. 2021. “Monopsony in Labor Markets: A Meta-Analysis.” *ILR Review* 74(1):27-55.

Soltani, Bahram. 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, David H., and Eugene Soltes. 2019. “Is 'Not Guilty' the Same as 'Innocent'? Evidence from SEC Financial Fraud Investigations.” *Working paper*.

Stanley, Tom D. 2005. “Beyond Publication Bias.” *Journal of Economic Surveys* 19(3):309-345.

Stanley, Tom D., and Hristos Doucouliagos. 2010. “Picture this: A Simple Graph that Reveals Much Ado About Research.” *Journal of Economic Surveys* 24(1):170-191.

Stanley, Tom D., and Hristos Doucouliagos. 2012. “Meta-Regression Analysis in Economics and Business” (Vol. 5). Routledge.

Stanley, Tom D., and Hristos Doucouliagos. 2017. “Neither Fixed nor Random: Weighted Least Squares Meta-Regression.” *Research Synthesis Methods* 8(1):19-42.

Stanley, Tom D. and Hristos Doucouliagos. 2019. “Practical Significance, Meta-Analysis and the Credibility of Economics.” *IZA Discussion Papers n° 12458,* Institute of Labor Economics (IZA).

Stanley, Tom D., Evan C. Carter, and Hristos Doucouliagos. 2018. “What Meta-Analyses Reveal About the Replicability of Psychological Research.” *Psychological Bulletin* 144:1325-1346.

Stigler, George J., 1964. “Public Regulation of the Securities Market.” *Journal of Business* 37(2):117-142.

Taffler, Richard J., Jeffrey Lu, and Asad Kausar. 2004. “In Denial? Stock Market Underreaction to Going-Concern Audit Report Disclosures.” *Journal of Accounting and Economics* 38:263-296.

Tibbs, Samuel L., Deborah L. Harrell, and Ronald E. Shrieves. 2011. “Do Shareholders Benefit from Corporate Misconduct? A Long-Run Analysis.” *Journal of Empirical Legal Studies* 8(3):449-476.

Tukey, John W. 1977. “Exploratory Data Analysis” (Vol. 2, 131-160).

van den Broek, Sebastianus Petrus, Ron G.M. Kemp, Annemarie Charlotte de Vries, and Willem F.C. Verschoor. 2010. “The Reputational Penalties to Firms in Antitrust Investigations.” *Financial Management Association European Meeting, Hamburg, Germany:18.*

Veld, Chris, Patrick Verwijmeren, and Yuriy Zabolotnyuk. 2018. “Wealth Effects of Seasoned Equity Offerings: A Meta-Analysis.” *International Review of Finance* 20(1):77-131.

Zigraiova, Diana, Tomáš Havránek, Zuzana Irsova, and Jiri Novak. 2021. “How Puzzling is the Forward Premium Puzzle? A Meta-Analysis.” *European Economic Review* 134:103714.

Table 1: Variable Definitions and Descriptive Statistics

Table 1 describes most of the variables for the full sample of financial crimes (480 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 estimates. In fact, on average, four estimates are reported for each study. Some categories are not mutually exclusive.

Variables	Description	Mean	Std dev.	Min.	Max.	Weighted mean
Effect: Abnormal returns (AAR and CAAR)*	Average abnormal returns or cumulative average abnormal returns reported in the original studies, estimated with an event study methodology.	-4.77%	6.38%	-27.2%	+2.0%	-5.92%
Standard Error*	Reported standard error of the estimated abnormal returns, or estimated with the conservative *t*-statistic *(see below).*	1.88%	2.76%	0.01%	17.3%	2.33%
Effect: Average Abnormal Return per day (AARD)*	Average abnormal returns per day (of the event window), equal to the reported average abnormal return (for one-day event windows) or to the reported cumulative average abnormal return divided by the number of days of the event window of the CAAR (for longer event windows).	-1.81%	2.79%	-14.5%	+2.0%	-2.42%
Standard Error per day*	Standard error per day, equal to the reported or estimated standard error for one-day event window, or to the reported or estimated standard error divided by the number of days of the event window of the CAAR.	0.80%	1.26%	0.03%	7.44%	1.00%
1. Data characteristics						
Geographical scope:**	1 if only one country in the scope, and zero otherwise.	0.95	0.23	0	1	0.97
	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).	0.69	0.47	0	1	0.76
	1 if the estimate's sample is the U.S., and zero otherwise.	0.64	0.48	0	1	0.71
	1 if the estimate's sample is Europe., and zero otherwise.	0.12	0.33	0	1	0.09
	1 if the estimate's sample is emerging economies (China, Malaysia, South Korea, Thailand, Turkey), and zero otherwise	0.23	0.42	0	1	0.18
	1 if the estimate's sample is Asia., and zero otherwise.	0.23	0.42	0	1	0.19
	1 if the estimate's sample is China, and zero otherwise.	0.19	0.39	0	1	0.13
Period under review:	Beginning of period under review (oldest financial crime).	1994	10.23	1965	2014	1994
	End of period under review (most recent financial crime).	2005	8.11	1979	2018	2004
	Average year of the period under review.	2000	8.66	1973	2016	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	0	3.8	3.22
	Length of the period under review (in years).	11.39	6.26	2	35	10.77
Event types:						
Types of regulatory breaches:** (See Appendix C, Figure C.2 for graphical illustration)	1 if the scope of crimes covers all violations of securities laws (incl. accounting fraud), and zero otherwise. Typically, articles investigating all sanction decisions made by an authority.	0.48	0.48	0	1	0.44
	1 if the scope of crimes covers exclusively accounting frauds, and zero otherwise.	0.33	0.47	0	1	0.43
	1 if the scope of crimes covers non-accounting violations of securities laws, and zero otherwise. Typically, articles investigating other types of market abuses (insider trading or price manipulation).	0.20	0.39	0	1	0.13
	1 if the scope of crimes covers market abuses (insider dealing, price manipulation, and breach of public disclosure requirements), and zero otherwise.	0.50	0.50	0	1	0.53
Sources of the news/origins of the data under review:**	1 if the crimes disclosed in newspaper articles (typically WSJ in the U.S.), and zero otherwise.	0.42	0.49	0	1	0.37
	1 if the crimes disclosed by enforcers, through regulatory communication, and zero otherwise.	0.66	0.47	0	1	0.66
	1 if the crimes disclosed by firms, in corporate communication, and zero otherwise.	0.26	0.44	0	1	0.30
Steps of enforcement procedure:**	1 if the crimes were alleged (not convicted), and zero otherwise.	0.61	0.49	0	1	0.58
	1 if the crimes were being investigated, and zero otherwise.	0.11	0.31	0	1	0.08
	1 if the crimes went through settlement, and zero otherwise.	0.04	0.19	0	1	0.03
	1 if the crimes led to an accounting restatement, and zero otherwise.	0.14	0.34	0	1	0.23
	1 if the crimes were convicted by an authority/court (verdict of regulatory procedures, verdict of lawsuits or class-actions, accounting restatement), and zero otherwise.	0.41	0.49	0	1	0.45
Types of enforcement procedure:***	1 if the crimes led to a regulatory procedure, and zero otherwise.	0.53	0.50	0	1	0.53
	1 if the crimes led to a stock exchange procedure, and zero otherwise.	0.09	0.28	0	1	0.08
	1 if the crimes led to a class-action, and zero otherwise.	0.24	0.43	0	1	0.22
	1 if the crimes led to a private lawsuit, and zero otherwise.	0.10	0.30	0	1	0.10

Variables	Description	Mean	Std dev.	Min.	Max.	Weighted mean
Main sectors:[***]	1 if specified that the most frequent sector involved in financial misconducts is industry, and zero otherwise.	0.33	0.47	0	1	0.32
	1 if specified that the most frequent sector involved in financial misconducts is financial institutions, and zero otherwise.[3]	0.17	0.38	0	1	0.18
2. Estimation characteristics						
Model:	1 if market model used to estimate abnormal returns (not Fama-French models, CAPM, or market-adjusted model), and zero otherwise.	0.83	0.38	0	1	0.84
	1 if equally weighted market index, and zero otherwise.	0.50	0.50	0	1	0.54
	1 if CRSP dataset used for returns (Center for Research in Securities Prices), and zero otherwise.	0.44	0.50	0	1	0.56
	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.	7.20	4.21	1	16	4.32
Sample characteristics:	1 if the initial sample size of financial crimes is specified in the article (before cleaning the data), and zero otherwise.	0.78	0.42	0	1	0.74
	1 if confounding events are explicitly excluded from the final sample, and zero otherwise.	0.29	0.46	0	1	0.23
	Number of sampled events (financial crimes) in the sample (after excluding confounding events and events with data problems).	264	378	4	2,194	260
	Number of sampled observations (financial crimes), as the logarithm of the final number of events in the sample.	4.82	1.25	1.39	7.69	4.80
Estimation window:	1 if estimation window used to estimate the parameters of normal returns specified in the article, and zero otherwise.	0.72	0.45	0	1	0.64
	Beginning of the estimation window (in days, relative to the event in t = 0).	-155	129	-1080	0	-153
	End of the estimation window (in days, relative to the event in t = 0).	-20	30	-300	0	-21
Event window <u>of the event study</u>:	Beginning of the event window for which abnormal returns are disclosed in the article (in days, relative to the event in t = 0).	-15	35	-255	0	-17
	End of the estimation window for which abnormal returns are disclosed in the article (in days, relative to the event in t = 0).	18	39	0	300	19
	Length of the event window for which abnormal returns are disclosed in the article (in days).	32.8	60.4	1	511	36.4
	1 if event windows beyond $[\![-10;+10]\!]$ (*i.e.* "long term" event windows, for which estimates are not reported in the meta-dataset).	0.28	0.45	0	1	0.27
Event window <u>of the reported estimate</u>:	1 of one-day event window (*i.e.* AAR), and zero otherwise.	0.37	0.48	0	1	0.30
	Beginning of the event window of the reported estimate (in days, relative to the event in t = 0).	-1.6	3.0	-10	6	-1.3
	End of the estimation window of the reported estimate (in days, relative to the event in t = 0).	1.4	2.8	-6	10	1.3
	Length of the event window of the reported estimate (in days).	4.0	4.5	1	21	3.6
	1 if the event window is strictly before the event date (t = 0), and zero otherwise.	0.14	0.35	0	1	0.09
	1 if the event window is limited to the event date (t = 0), and zero otherwise.	0.17	0.38	0	1	0.16
	1 if the event window is around the event date (t = 0), and zero otherwise.	0.57	0.50	0	1	0.66
	1 if the event window is strictly after the event date (t = 0), and zero otherwise.	0.12	0.33	0	1	0.09
	1 if "exotic" event window (standing for less than 5% of the compounded event windows, *i.e.* less than 24 estimates, as opposed to more frequent and standard event windows), and zero otherwise.	0.26	0.44	0	1	0.22
Statistical significance:[**]	Conservative *t*-statistics: - *t*-stat when published. - when the *t*-stat not published (Frooman, 1997): 1) the statistical significance levels converted into conservative levels of significance; 2) the *z*s 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); 3) the *p* values converted into *t*s by using a *t* table and the appropriate degrees of freedom: and 4) conservative hypothesis of 10% statistical significance for three studies standing for seven estimates (Desai et al. (2006), Nelson et al. (2009), and Goldman et al. (2012)), with the significant abnormal returns but without including *t*-statistics nor the statistical significance level.	-3.28	4.85	-46.82	7.88	-3.43
	1 if abnormal returns are statistically significant, and zero otherwise.	0.73	0.45	0	1	0.80

Variables	Description	Mean	Std dev.	Min.	Max.	Weighted mean
	1 if statistical significance level ("stars"), and zero otherwise.	0.84	0.37	0	1	0.83
	1 if *t*-statistics, and zero otherwise.	0.50	0.50	0	1	0.46
	1 if *p*-value, and zero otherwise.	0.15	0.36	0	1	0.14
	1 if parametric *z*-statistics, and zero otherwise.	0.24	0.42	0	1	0.20
	1 if non-parametric tests (most frequently signed-rank tests such as Corrado and Zivney (1992), Kolari & Pynnonen (2011), Cowan (1992), Corrado (1989)).	0.18	0.38	0	1	0.15
	1 if non-parametric rank test, and zero otherwise.	0.11	0.32	0	1	0.11
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	1	0.66
	1 if additional estimates of reputational penalties.	0.09	0.28	0	1	0.08
	1 of mention of average cash fines set by authorities subsequent to financial crimes, and 0 otherwise.	0.14	0.34	0	1	0.12
3. Publication characteristics						
Characteristics of the article:	A number of authors of the paper.	2.32	0.86	1	4	2.37
	1 if multiple authorships in the sample, and zero otherwise.	0.29	0.46	0	1	0.33
	Year of publication.	2009	7.76	1984	2020	2009
	Publication year, as the logarithm of the year of publication.	3.20	0.45	0	3.61	3.20
Journal fields:[**, 4]	1 if finance journal, and zero otherwise.	0.38	0.48	0	1	0.41
	1 if economic journal, and zero otherwise.	0.28	0.45	0	1	0.28
	1 if accounting journal, and zero otherwise.	0.25	0.44	0	1	0.24
	1 if business (management, and organization) journal, and zero otherwise.	0.25	0.44	0	1	0.26
	1 if law journal, and zero otherwise.	0.20	0.40	0	1	0.18
	1 if cross-disciplinary journal, and zero otherwise. As stated in Amiram et al. (2018), studies on financial misconduct belongs to three perspectives: law, accounting, and finance.	0.53	0.50	0	1	0.50
Quality of the publication:	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	1	0.81
	1 if published in a Scopus journal.	0.63	0.48	0	1	0.67
	Number of citations in Google Scholar (as number).	135	351	0	5,007	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	0	5.30	1.60
	Scopus Cite Score in 2018.	1.65	1.89	0	7.34	1.64
	Scopus Cite Score of the year of publication (2011 to 2018, otherwise 2011).	1.10	1.32	0	5.58	1.17
	IDEAS/RePEc Recursive Discounted impact factor.	0.40	1.11	0	5.67	0.39

Sources: Studies, Authors' calculations [*] Winsorized at the 1% level. [**] Non-mutually exclusive. [***] Mutually exclusive.

[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] In three articles (Bonini and Boraschi (2010), Firth et al. (2011), and Ozeki (2019)), financial firms were excluded from the sample.

[4] The classification of journals echoes the Web of Science research areas. Each journal was classified based on their tittles and official descriptions.

Tables 2: Abnormal Returns for Different Subsets of Data

Table 2.A details, for the whole sample and different subsets, the **average abnormal returns (AARs) and cumulative abnormal returns (CAARs)**, as collected in the original studies, complemented by the number of observations, the standard errors (SE) and a 95% confidence interval. Means are simple averages ("Unweighted", reported in Columns [2] to [5]) or they are weighted by the inverse of the number of estimates reported *per* study ("Weighted", reported in Columns [6] to [9]). Some categories are not mutually exclusive. The definitions of the subsets are available in Table 1.

(Cumulative) Average Abnormal Returns ((C)AARs)	Nb. obs.	Unweighted				Weighted			
		Mean	SE	95% conf. int.		Mean	SE	95% conf. int.	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
1. Structural characteristics:									
Geographical specificities:									
U.S. only	305	-6.51%	0.41%	-7.32%	-5.70%	-7.43%	0.42%	-8.25%	-6.61%
Common-law countries	329	-6.14%	0.39%	-6.93%	-5.40%	-7.07%	0.40%	-7.79%	-6.23%
Code-law countries*	151	-1.72%	0.23%	-2.16%	-1.28%	-2.45%	0.31%	-3.06%	-1.83%
Emerging countries*	110	-0.50%	0.09%	-0.67%	-0.33%	-1.51%	0.18%	-1.86%	-1.16%
China only	89	-1.49%	0.21%	-1.90%	-1.07%	-1.54%	0.20%	-1.93%	-1.14%
Event under review:									
Exclusively accounting frauds*	156	-6.85%	0.55%	-7.94%	-5.76%	-8.90%	0.57%	-10.00%	-7.78%
Market abuses	238	-5.68%	0.42%	-6.51%	-4.85%	-7.82%	0.46%	-8.72%	-6.91%
Violations of securities laws (including accounting frauds)*	230	-3.83%	0.40%	-4.52%	-3.05%	-3.97%	0.40%	-4.75%	-3.19%
Exclusively violations of securities laws (excl. accounting fraud)*	94	-3.60%	0.56%	-4.71%	-2.49%	-2.78%	0.46%	-3.70%	-1.85%
Step of the enforcement:									
Alleged crimes (allegation in the press, initiation of regulatory procedures, investigation, class-action or lawsuit filing, etc.)*	293	-5.62%	0.39%	-6.38%	-4.86%	-6.17%	0.38%	-6.92%	-5.41%
Convicted crimes (verdict of regulatory procedures, verdict of lawsuits or class-actions, accounting restatement)*	187	-3.43%	4.19%	-4.26%	-2.60%	-5.60%	0.52%	-6.63%	-4.56%
Type of procedure:									
Public enforcement*	256	-4.57%	0.42%	-5.40%	-3.75%	-5.59%	0.43%	-6.42%	-4.76%
Private enforcement (stock market procedure, class action, lawsuit, settlement)*	188	-4.52%	0.43%	-5.37%	-3.67%	-5.03%	0.48%	-5.98%	-4.08%
Source of the news:									
Crimes disclosed in newspaper articles	202	-5.01%	0.47%	-5.94%	-4.08%	-5.78%	0.48%	-6.73%	-4.84%
Crimes disclosed by enforcers	319	-4.61%	0.36%	-5.32%	-3.89%	-5.80%	0.39%	-6.58%	-5.03%
Crimes disclosed by firms	123	-5.49%	0.60%	-6.69%	-4.30%	-7.53%	0.67%	-8.86%	-6.20%
2. Estimation characteristics:									
Estimation model:									
Market model*	399	-4.67%	0.32%	-5.29%	-4.02%	-5.62%	0.34%	-6.30%	-4.95%
Other models*	81	-5.34%	0.65%	-6.64%	-4.04%	-7.48%	0.74%	-8.95%	-6.01%
Event windows:									
Before the event ($t < 0$)*	67	-1.47%	0.30%	-2.07%	-0.87%	-1.22%	0.29%	-1.80%	-0.64%
On the event day ($t = 0$)*	83	-3.49%	0.55%	-4.58%	-2.40%	-4.80%	0.62%	-6.03%	-3.56%
Around the event day (including $t = 0$)*	272	-6.81%	0.43%	-7.67%	-5.96%	-7.56%	0.44%	-8.42%	-6.70%
After the event day ($t > 0$)*	59	-0.89%	0.23%	-1.35%	-0.42%	-0.81%	0.22%	-1.25%	-0.36%
"Exotic" event windows	123	-7.15%	0.69%	-8.50%	-5.79%	-6.86%	0.64%	-8.13%	-5.59%
Exclusion of confounding events	140	-3.91%	0.48%	-4.87%	-2.95%	-3.77%	0.47%	-4.70%	-2.84%
Complementary estimations:									
Reputational penalty estimation	42	-2.52%	0.54%	-3.61%	-1.42%	-3.80%	0.62%	-5.04%	-2.55%
Cross-sectional regressions of (C)AARs	299	-5.51%	0.39%	-6.28%	-4.73%	-6.87%	0.42%	-7.69%	-6.05%
3. Publication status:									
Published papers/chapters*	389	-4.88%	0.32%	-5.51%	-4.24%	-5.59%	0.33%	-6.23%	-4.94%
Unpublished papers*	91	-4.33%	0.67%	-5.67%	-2.99%	-7.37%	0.85%	-9.06%	-5.68%
All estimates	**480**	**-4.77%**	**0.29%**	**-5.34%**	**-4.20%**	**-5.92%**	**0.31%**	**-6.54%**	**-5.31%**

*Sources: Studies, Authors' calculations (winsorized at the 1% level) * Mutually exclusive categories by section.*

Table 2.B complements Table 1.A by detailing, for the whole sample and different subsets, the **normalized average abnormal returns *per* day (AARDs)**, complemented by the number of observations, the standard errors (SE) and a 95% confidence interval. Means are simple averages ("Unweighted", reported in Columns [2] to [5]) or they are weighted by the inverse of the number of estimates reported per study ("Weighted", reported in Columns [6] to [9]). Some categories are not mutually exclusive. The definitions of the subsets are available in Table 1.

Average Abnormal Returns *per* Day (AARDs)	**Nb. obs.**	**Unweighted**				**Weighted**			
		Mean	**SE**	**95% conf. int.**		**Mean**	**SE**	**95% conf. int.**	
	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**
1. Structural characteristics:									
Geographical specificities:									
U.S. only	305	-2.46%	0.18%	-2.82%	-2.11%	-3.04%	0.20%	-3.43%	-2.64%
Common-law countries*	329	-2.35%	0.17%	-2.68%	-2.01%	-2.88%	0.19%	-3.25%	-2.50%
Code-law countries*	151	-0.64%	0.11%	-0.85%	-0.42%	-0.95%	0.16%	-1.26%	-0.64%
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 frauds*	156	-2.88%	0.27%	-3.40%	-2.35%	-3.67%	0.28%	-4.22%	-3.13%
Market abuses	238	-2.38%	0.21%	-2.79%	-2.00%	-3.37%	0.24%	-3.84%	-2.90%
Violations of securities laws (including accounting frauds)*	230	-1.23%	0.15%	-1.52%	-0.94%	-1.51%	0.18%	-1.88%	-1.15%
Exclusively violations of securities laws (excl. accounting frauds)*	94	-1.45%	0.26%	-1.96%	-0.94%	-1.35%	0.27%	-1.89%	-0.81%
Step of the enforcement:									
Alleged crimes (allegation in the press, initiation of regulatory procedures, investigation, class-action or lawsuit filing, etc.)*	293	-2.21%	0.17%	-2.55%	-1.87%	-2.71%	0.20%	-3.11%	-2.31%
Convicted crimes (verdict of regulatory procedures, verdict of lawsuits or class-actions, accounting restatement)*	187	-1.19%	0.17%	-1.52%	-0.85%	-2.01%	0.21%	-2.43%	-1.60%
Type of procedure:									
Public enforcement*	256	-1.82%	0.18%	-2.18%	-1.46%	-2.41%	0.22%	-2.84%	-1.99%
Private enforcement (stock market procedure, class action, lawsuit, settlement)*	188	-1.56%	0.18%	-1.91%	-1.21%	-1.94%	0.21%	-2.35%	-1.54%
Source of the news:									
Crimes disclosed in newspaper articles	202	-1.72%	0.18%	-2.07%	-1.36%	-1.97%	0.18%	-2.32%	-1.62%
Crimes disclosed by enforcers	319	-1.77%	0.16%	-2.09%	-1.45%	-2.46%	0.20%	-2.85%	-2.06%
Crimes disclosed by firms	123	-2.05%	0.25%	-2.57%	-1.56%	-2.77%	0.26%	-3.29%	-2.25%
2. Estimation characteristics:									
Estimation model:									
Market model*	399	-1.82%	0.14%	-2.09%	-1.55%	-2.15%	0.14%	-2.44%	-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	-0.74%	0.15%	-1.05%	-0.44%	-0.63%	0.17%	-0.96%	-0.30%
On the event day ($t = 0$)*	83	-3.31%	0.48%	-4.27%	-2.35%	-4.60%	0.57%	-5.74%	-3.47%
Around the event day (including $t = 0$)*	272	-1.84%	0.15%	-2.13%	-1.55%	-2.37%	0.16%	-2.68%	-2.06%
After the event day ($t > 0$)*	59	-0.77%	0.22%	-1.20%	-0.34%	-0.56%	0.19%	-0.93%	-0.19%
"Exotic" event windows	123	-0.93%	0.14%	-1.20%	-0.65%	-1.05%	0.14%	-1.33%	-0.77%
Exclusion of confounding events	140	-1.21%	0.15%	-1.50%	-0.92%	-1.23%	0.14%	-1.51%	-0.94%
Complementary estimations:									
Reputational penalty estimation	42	-0.71%	0.14%	-0.99%	-0.43%	-1.27%	0.20%	-1.68%	-0.87%
Cross-sectional regressions of (C)AARs	299	-2.01%	0.17%	-2.33%	-1.68%	-2.50%	0.17%	-2.84%	-2.16%
3. Publication status:									
Published papers/chapters*	389	-1.85%	0.14%	-2.13%	-1.57%	-2.37%	0.16%	-2.64%	-2.01%
Unpublished papers*	91	-1.64%	0.28%	-2.21%	-1.08%	-2.80%	0.39%	-3.59%	-2.01%
All estimates	**480**	**-1.81%**	**0.13%**	**-2.06%**	**-1.56%**	**-2.42%**	**0.15%**	**-2.71%**	**-2.12%**

*Sources: Studies, Authors' calculations (winsorized at the 1% level) * Mutually exclusive categories by section.*

Table 3: Meta-Regression Analysis of Publication Selection Bias

Table 3 details the results of the publication selection bias analysis, based on the FAT-PET tests (Eq. (1)) for the original sample of reported estimates ((C)AARs, column 1) and for the normalized sample of estimates (AARDs, column 2). Additionally, four sub-samples of the normalized estimates (AARDs) are compared: the U.S. *versus* other countries (columns 3 and 4), and exclusively accounting frauds *versus* other securities law violations (columns 5 and 6).

The standard errors (SE or SED) 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 1 (OLS; study-level fixed effects and study-level between effects, to exploit respectively idiosyncratic study-level variations (methodology, samples) and differences in size of the 111 studies; hierarchical Bayes; and using the number of observations reported by the study as an instrument variable; 2) weighted least squares estimations in Panel 2 (by the precision, *i.e.* the inverse of the standard errors, and by the inverse of the number of estimates reported by the study); and 3) three recent non-linear estimations in Panel 3, 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).

	(C)AARs	**AARDs**				
	Full sample	**Full sample**	**U.S. only**	**Other countries**	**Account. frauds**	**Violations of sec. laws**
	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**
Panel 1. Unweighted estimations						
1. OLS						
SE or SED (*publication bias*)	-1.52***	-1.48***	-1.66***	-0.56	-1.34***	-1.53***
	(0.275)	(0.262)	(0.216)	(0.527)	(0.404)	(0.287)
Intercept (*effect beyond bias*)	-1.91%***	-0.63%***	-0.83%***	-0.41%**	-1.23%***	-0.40%***
	(0.005)	(0.002)	(0.002)	(0.002)	(0.004)	(0.001)
2. Study-level fixed effects						
SE or SED (*publication bias*)	-1.86***	-1.45***	-1.67***	-0.43***	-1.43***	-1.72***
	(0.142)	(0.120)	(0.157)	(0.145)	(0.188)	(0.172)
Intercept (*effect beyond bias*)	-1.27%***	-0.66%***	-0.82%***	-0.47%***	-1.11%***	-0.29%**
	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
3. Study-level between effects						
SE or SED (*publication bias*)	-1.50***	-1.59***	-1.83***	-0.73***	-1.46***	-1.50***
	(0.143)	(0.148)	(0.174)	(0.187)	(0.198)	(0.269)
Intercept (*effect beyond bias*)	-2.45%***	-0.82%***	-0.93%***	-0.45%	-1.29%***	-0.55%*
	(0.005)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
4. Hierarchical Bayes						
SE or SED (*publication bias*)	-1.61***	-1.47***	-1.82***	-0.95***	-1.14***	-1.59***
	(0.213)	(0.219)	(0.280)	(0.320)	(0.328)	(0.264)
Intercept (*effect beyond bias*)	-1.80%***	-0.74%***	-0.67%***	-0.29%***	-1.80%***	-0.33%***
	(0.027)	(0.027)	(0.040)	(0.070)	(0.063)	(0.039)
5. IV number of observations reported by study						
SE or SED (*publication bias*)	-3.08***	-1.68***	-1.87***	-0.76	-1.83***	-2.11***
	(0.618)	(0.436)	(0.533)	(0.483)	(0.404)	(0.703)
Intercept (*effect beyond bias*)	1.03%	-0.47%*	-0.62%	-0.32%**	-0.62%*	-0.006%
	(0.010)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
Panel 2. Weighted least square estimations						
1. Weighted by the precision (inverse of the standard error)						
SE or SED (*publication bias*)	-2.23***	-1.93***	-1.97***	-1.17***	-2.02***	-1.79***
	(0.228)	(0.204)	(0.211)	(0.393)	(0.296)	(0.249)
Intercept (*effect beyond bias*)	-0.58%***	-0.27%***	-0.52%***	-0.13%*	-0.39%*	-0.25%***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)

	(C)AARs			AARDs								
	Full sample		Full sample		U.S. only		Other countries		Account. frauds		Violations of sec. laws	
	[1]		[2]		[3]		[4]		[5]		[6]	
2. Weighted by the inverse of the number of estimates reported by study												
SE or SED (*publication bias*)	-1.50	***	-1.55	***	-1.81	***	-0.65		-1.43	***	-1.52	***
	(0.305)		(0.287)		(0.202)		(0.590)		(0.391)		(0.294)	
Intercept (*effect beyond bias*)	-2.43%	***	-0.86%	***	-0.96%	***	-0.49%	**	-1.45%	***	-0.57%	**
	(0.006)		(0.003)		(0.003)		(0.002)		(0.004)		(0.003)	
Panel 3. Non-linear estimations												
1. Weighted average of adequately powered (Ioannidis et al., 2017)[2]												
Effect beyond bias	-0.21%	***	-0.15%	***	-0.27%	***	-0.08%	***	-0.13%	***	-0.16%	***
	(0.009)		(0.003)		(0.007)		(0.002)		(0.006)		(0.005)	
2. Selection model (Andrews and Kasy, 2019)[3]												
Effect beyond bias	-0.56%	*	-0.39%	***	-0.73%	***	-0.07%	***	-1.34%	***	-0.43%	***
	(0.132)		(0.076)		(0.087)		(0.011)		(0.366)		(0.062)	
3. Stem-based method (Furukawa, 2019)												
Effect beyond bias	-0.36%		-0.13%		-0.52%		-0.01%		-0.14%		-0.14%	
	(0.008)		(0.004)		(0.006)		(0.002)		(0.006)		(0.005)	
Number of observations[1]	**480**		**480**		**305**		**175**		**156**		**324**	

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] 40 (C)AARs and 70 AARDs estimates are adequately powered under the conventional definition that an estimate is adequately powered if it has Cohen's 80% or higher power if its standard error is less than the absolute value of Unrestricted Weighted Least Square divided by 2.8 (Ioannidis et al., 2017; Stanley et al., 2018).
[3] Complementary results of the selection model are displayed in Appendix H, Figure H.1, with funnel plots and histograms of Z-Statistics for the whole sample and sub-samples.

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

Table 4 details the results for the Bayesian Model Averaging (BMA) [column 1], the frequentist check (OLS) [column 2], and the Frequentist Model Averaging (FMA) [column 3]. 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 12 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). The definitions of the variables are available in Table 1.

Response variable:	**AARDs**	**Bayesian Model Averaging [1]**			**Frequentist Check (OLS) [2]**			**Frequentist Model Averaging [3]**		
Observations:	**480**	Post. Mean	Post. std. dev.	PIP	Coef.	Std. er.	*p*-value	Coef.	Std. er.	*p*-value
Constant		2.61%	NA	100%	2.04%	0.007	0.00	3.75%	0.009	0.00
Standard error *per* day		-1.40	0.074	100%	-1.41	0.179	0.00	-1.36	0.073	0.00
Structural characteristics:										
Geographical scope:	Only U.S.	-0.05%	0.003	16%				-0.33%	0.002	0.15
Period under review:	Mid-point of sampled events*	-0.20%	0.002	43%				-0.44%	0.002	0.07
Types of event:	Exclusively accounting fraud	-0.70%	0.003	97%	-0.68%	0.002	0.00	-0.73%	0.002	0.00
	Alleged frauds	-1.31%	0.002	100%	-1.28%	0.002	0.00	-1.25%	0.002	0.00
	Crimes disclosed in newspaper articles	0.02%	0.001	11%				0.13%	0.002	0.52
Estimation characteristics:										
Sample characteristics	Initial sample size specified	0.00%	0.001	8%				0.10%	0.002	0.63
	Confounding events excluded	-0.02%	0.001	10%				-0.22%	0.002	0.30
	Nb sampled events/financial crimes*	-0.21%	0.001	89%	-0.26%	0.001	0.00	-0.22%	0.001	0.01
Model	Estimation window specified	-0.88%	0.002	99%	-0.83%	0.002	0.00	-1.00%	0.002	0.00
Event window of the reported AARDs	Length of the event window	0.02%	0.000	42%				0.03%	0.000	0.32
	Event window strictly before the event	0.91%	0.003	99%	0.97%	0.002	0.00	0.80%	0.003	0.00
	Event window = event	-0.23%	0.003	46%				-0.47%	0.002	0.04
	"Exotic" event windows	0.28%	0.003	51%	0.59%	0.002	0.00	0.33%	0.003	0.29
Statistical significance indicators	Statistical significance level ("stars")	0.87%	0.002	99%	0.90%	0.003	0.01	0.87%	0.002	0.00
	Non-parametric tests	1.18%	0.003	100%	1.18%	0.004	0.00	1.15%	0.003	0.00
Complementary results	Cross-sectional regression	-0.70%	0.002	98%	-0.67%	0.003	0.01	-0.71%	0.002	0.00
	Reputational penalty estimation	0.09%	0.002	20%				0.44%	0.003	0.16
Publication characteristics:										
Characteristics of the article	Number of authors	-0.33%	0.001	96%	-0.33%	0.002	0.04	-0.33%	0.001	0.00
	Multiple authorships	-0.10%	0.002	29%				-0.32%	0.002	0.09
	Business journals	0.81%	0.002	99%	0.91%	0.003	0.00	0.67%	0.002	0.00
Quality of the publication	Nb Google quotes per year since pub.*	-0.15%	0.001	61%	-0.12%	0.001	0.24	-0.27%	0.001	0.01
	Scopus cite score	0.06%	0.001	39%				0.10%	0.001	0.15

Source: Authors' estimations. * *Log*

Notes: Std. dev. = standard deviation; PIP = posterior inclusion probability (as %); Std. er. = standard error.

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 (Panel 1) and on subsamples by geographies (Panel 2), types of event (Panel 3), and characteristics of the methodology (Panel 4). 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 1). 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 non-parametric statistical tests). 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 2-4, 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.

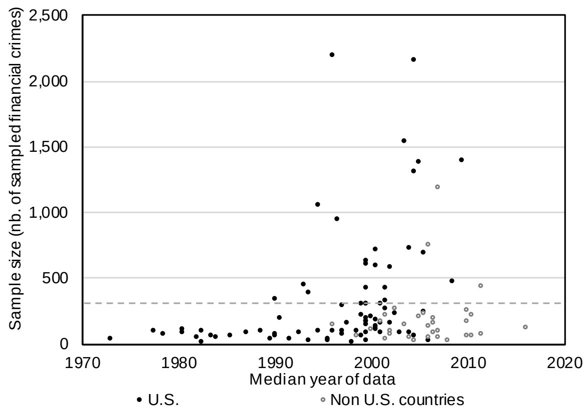
Panel 1. Levels of confidence	**Confidence 1**		**Confidence 2**		**Confidence 3**		**Confidence 4**	
Constant (*corrected AARDs*)	-1.13%	*	-1.37%	***	0.43%		-0.53%	***
	(0.007)		(0.005)		(0.003)		(0.002)	
Standard error (*publication bias*)	-1.48	***	-1.48	***	-1.51	***	-1.50	***
	(0.262)		(0.256)		(0.232)		(0.264)	
Lower confidence	0.007		0.008	*	-0.012	***	-0.005	
	(0.007)		(0.005)		(0.004)		(0.005)	
Observations	480		480		480		480	
Panel 2. Geography	**Only U.S.[2]**		**Common-law countries[2]**		**European Countries[2]**		**Emerging countries[2]**	
Constant (*corrected AARDs*)	-1.55%	***	-1.44%	***	-0.83%	***	-0.51%	***
	(0.005)		(0.004)		(0.000)		(0.002)	
Standard error (*publication bias*)	-1.67	***	-1.67	***	0.07		-0.83	
	(0.208)		(0.207)		(0.0620)		(0.607)	
Lower confidence	0.008	*	0.008	*	0.003	**	0.003	**
	(0.005)		(0.005)		(0.001)		(0.001)	
Observations	305		329		58		110	
Panel 3. Types of events	**Exclusively accounting frauds[1]**		**Other crimes[1]**		**Alleged frauds[2]**		**Sanctioned frauds[2]**	
Constant (*corrected AARDs*)	-2.29%	***	-0.23%	**	-1.82%	***	-0.43%	***
	(0.004)		(0.001)		(0.002)		(0.001)	
Standard error (*publication bias*)	-1.34	***	-1.53	***	-1.89	***	-0.99	***
	(0.403)		(0.287)		(0.185)		(0.339)	
Lower confidence	0.011	**	0.002		0.012	***	0.000	
	(0.003)		(0.002)		(0.001)		(0.005)	
Observations	156		324		293		187	
Panel 4. Methodology	**Exotic ev. windows[2]**		**Stat. signif. star[1]**		**Non param. tests[1]**		**Cross sectional reg.[2]**	
Constant (*corrected AARDs*)	-0.54%	***	-0.24%	**	-1.79%	**	-1.04%	***
	(0.001)		(0.001)		(0.008)		(0.004)	
Standard error (*publication bias*)	-1.84	***	-1.49	***	-0.81	**	-2.04	**
	(0.100)		(0.266)		(0.410)		(0.143)	
Lower confidence	0.003	**	0.002	**	0.015	*	0.007	*
	(0.001)		(0.002)		(0.008)		(0.143)	
Observations	123		401		84		299	

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*

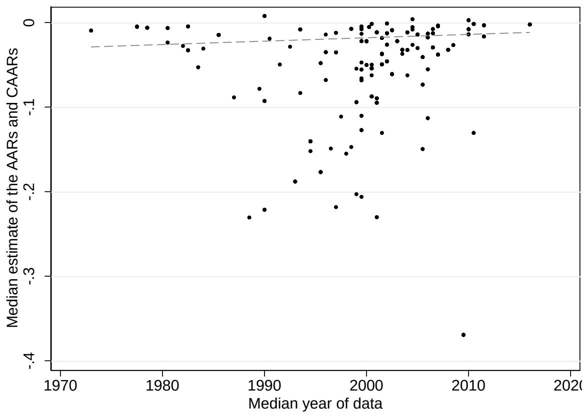
Figure 1: Chronological Ordering of Sample Sizes and Abnormal Returns

Figure 1 shows in Panel A the chronological distribution of the 32,500 financial crimes investigated by each of the 111 individual studies, with a split between U.S. and the rest of the world, and in Panel B the chronological distribution of median (cumulative) average abnormal returns reported in primary studies. The median year of the data used in the corresponding study ranges from 1973 to 2016. The dotted lines stand for the average number of financial crimes by study (264) in Panel A and for the time trend of abnormal returns in Panel B (AARDs also trend positively along time). On average, the financial crimes occurred in 2000, and each study covers 11 years.

Panel A. Chronological Ordering of Sample Sizes.



Panel B. Chronological Ordering of Median Abnormal Returns.

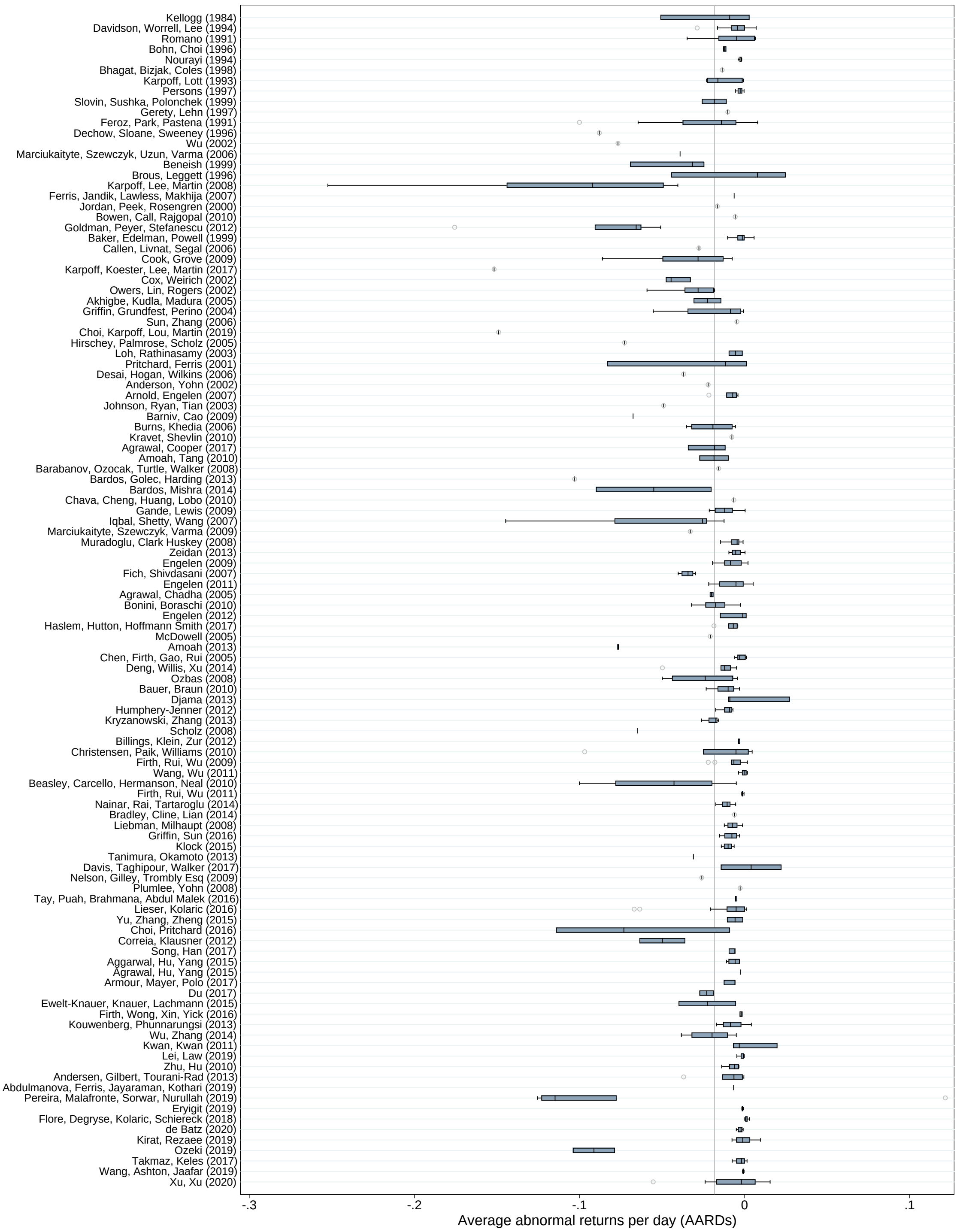


Source: Authors' calculations.

Figures 2: Distribution of Abnormal Returns Across Studies

Figure 2.A shows a box plot of the estimated average abnormal returns (AARs) and cumulative average abnormal returns (CAARs) 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 vertical dashed line denotes the *naïve* average (-4.8%). Studies are sorted by the median year of the sampled data, in ascending order.

Figure 2.B shows a box plot of the estimated average abnormal returns per day (AARDs) 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 vertical dashed line denotes the *naïve* average (-1.8%). Studies are sorted by the median year of the sampled data, in ascending order.



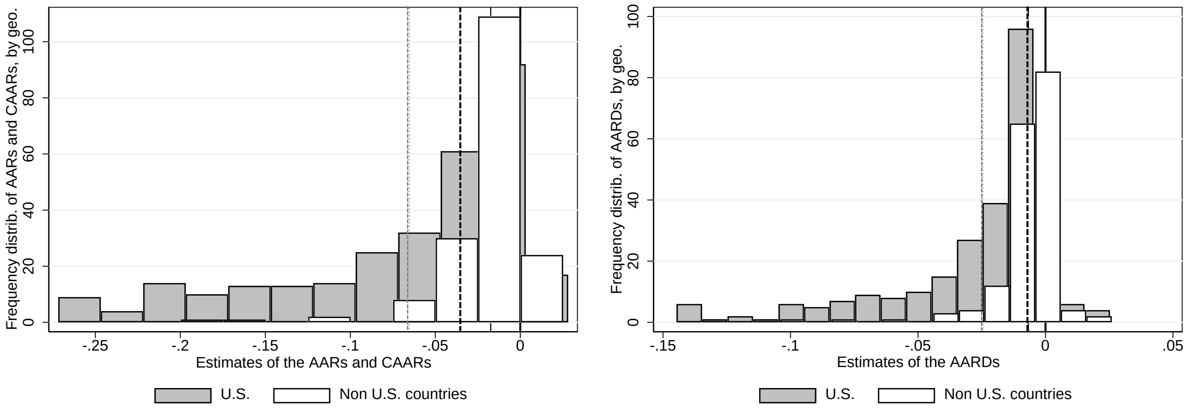
Source: Authors' calculations.

Figures 3: Frequency Distributions of Abnormal Returns

Figures 3 show the histograms of the estimates of average abnormal returns AARs and CAAR ((C)AARs) as reported in the individual studies (on the left hand-side) and normalized average abnormal returns AARDs by the length of the event windows (on the right hand-side). Each sample is split between the U.S. and non-U.S. countries (Panel 1) and between exclusively accounting fraud and other violations of securities laws (Panel 2). Outliers are excluded from the figures but included in all the tests.

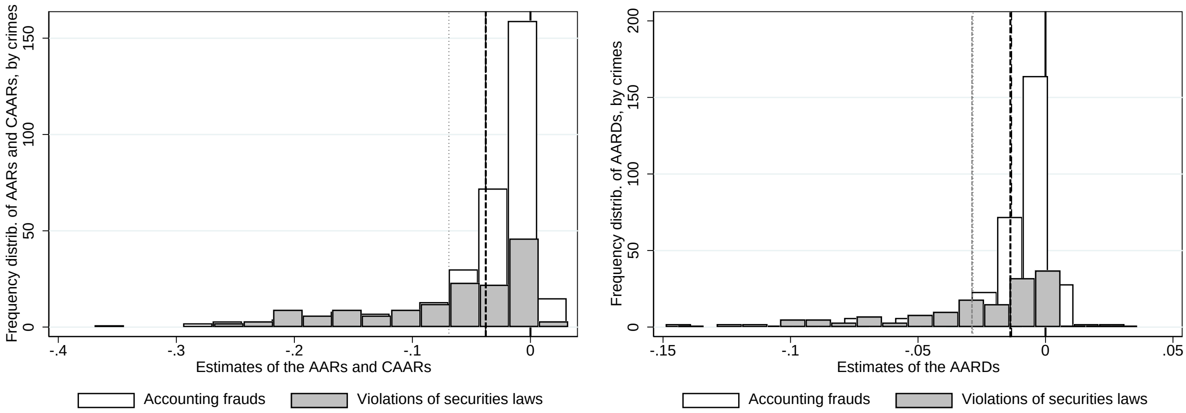
Figures 3.1: Financial Crimes Committed in the U.S. or in Other Countries.

Dotted grey lines depict the average abnormal returns ((C)AARs with AARs and CAARs, *versus* AARDs) for the U.S. (respectively -6.5% and -2.5%) and dashed black lines the averages for non-U.S. countries (respectively -1.2% and -0.7%).



Figures 3.2: Exclusively Accounting Frauds versus Other Violations of Securities Laws.

Dotted grey lines depict the average abnormal returns ((C)AARs with AARs and CAARs, versus AARDs) for the accounting frauds (respectively -6.9% and -2.9%) and dashed black lines the averages for violations of securities laws (respectively -3.8% and -1.2%).



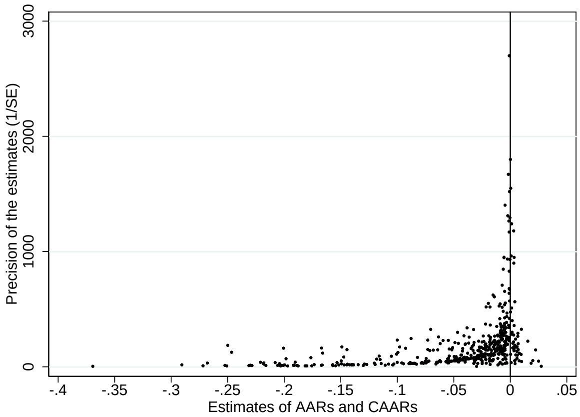
Source: Authors' calculations.

Figures 4: Funnel Graphs of the Impact of Financial Crimes

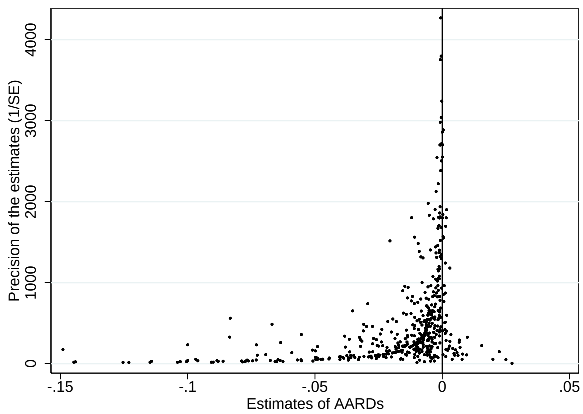
The following funnel graphs scatter the estimated abnormal returns of the disclosure of financial crimes against these estimates' precision (*i.e.* the inverse of the estimated standard errors). The distribution is expected to be symmetrical around the true value of the estimate, in the absence of publication bias. For each Figure, the left hand-side funnel plot depicts the full sample of 480 estimates as published in original studies ((C)AARs; with all AARs and CAARs, Panels A), while the right hand-side funnel plot depicts the normalized sample by the length of the event windows (AARDs, Panels B). Figures 4.1 plot the full sample. Figures 4.2 split estimates by geography, to compare estimates from U.S. with estimates from other jurisdictions. Figures 4.3 split estimates between exclusively accounting frauds and violations of securities laws.

Figures 4.1: Full Sample of Abnormal Returns.

Panel A. (C)AARs

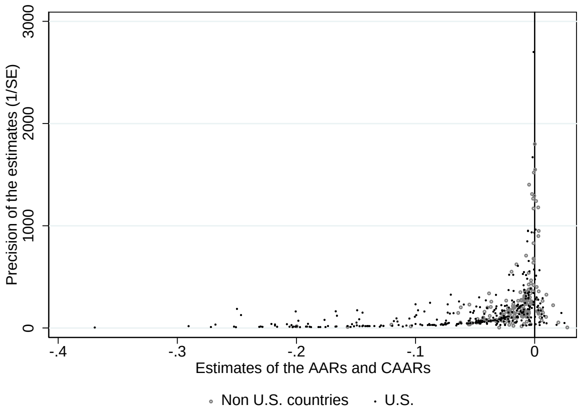


Panel B. AARDs

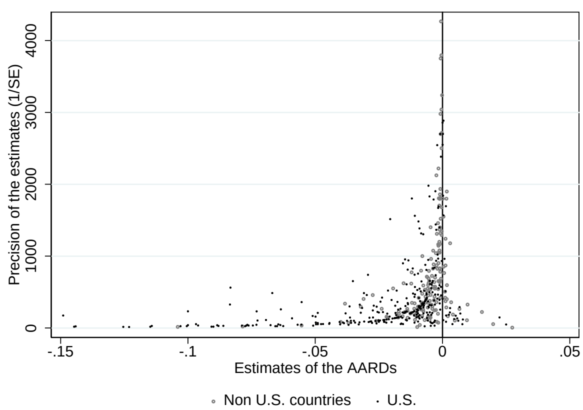


Figures 4.2: Financial Crimes Committed in the U.S. versus in Non U.S. Countries.

Panel A. (C)AARs

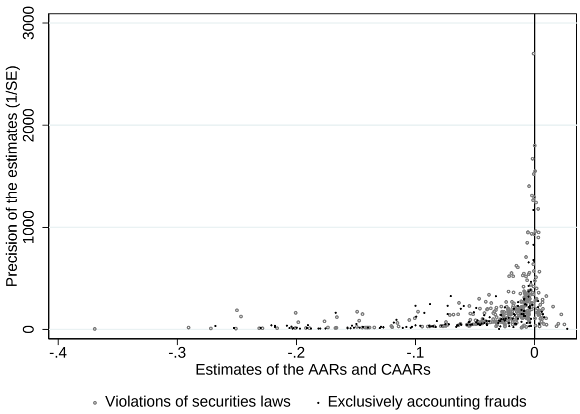


Panel B. AARDs

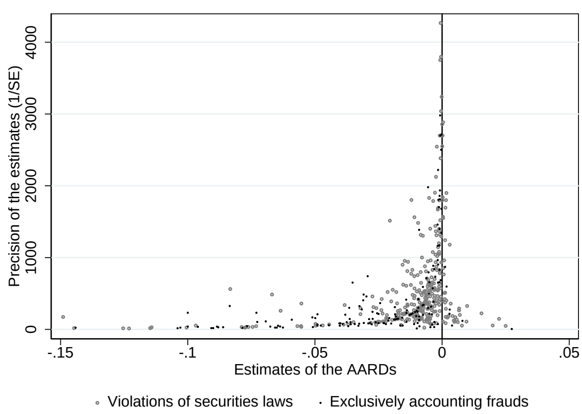


Figures 4.3: Exclusively Accounting Frauds versus Violations of Securities Laws.

Panel A. (C)AARs



Panel B. AARDs



Source: Authors' calculations.

Figures 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. Figure 5.1 is a histogram of the conservative *t*-statistics (*i.e.* 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 Figure 5.2 with the Kernel density estimate distribution for the estimated *t*-statistics of the sample of estimates from the 111 individual studies.

Figure 5.1: Frequency Distribution of *t*-Statistics.

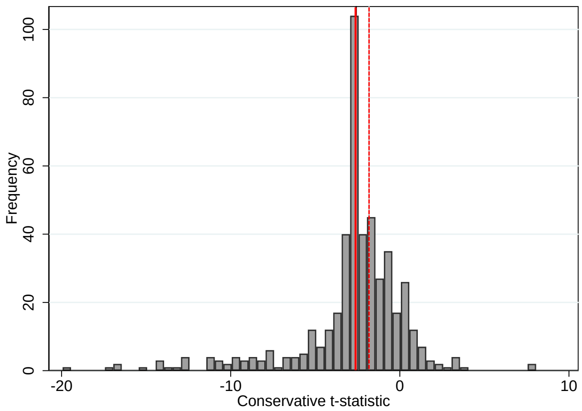
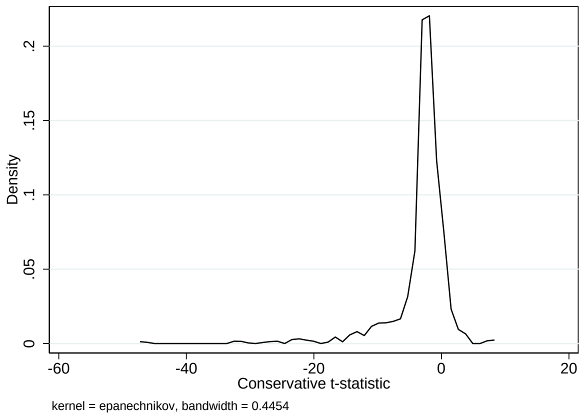


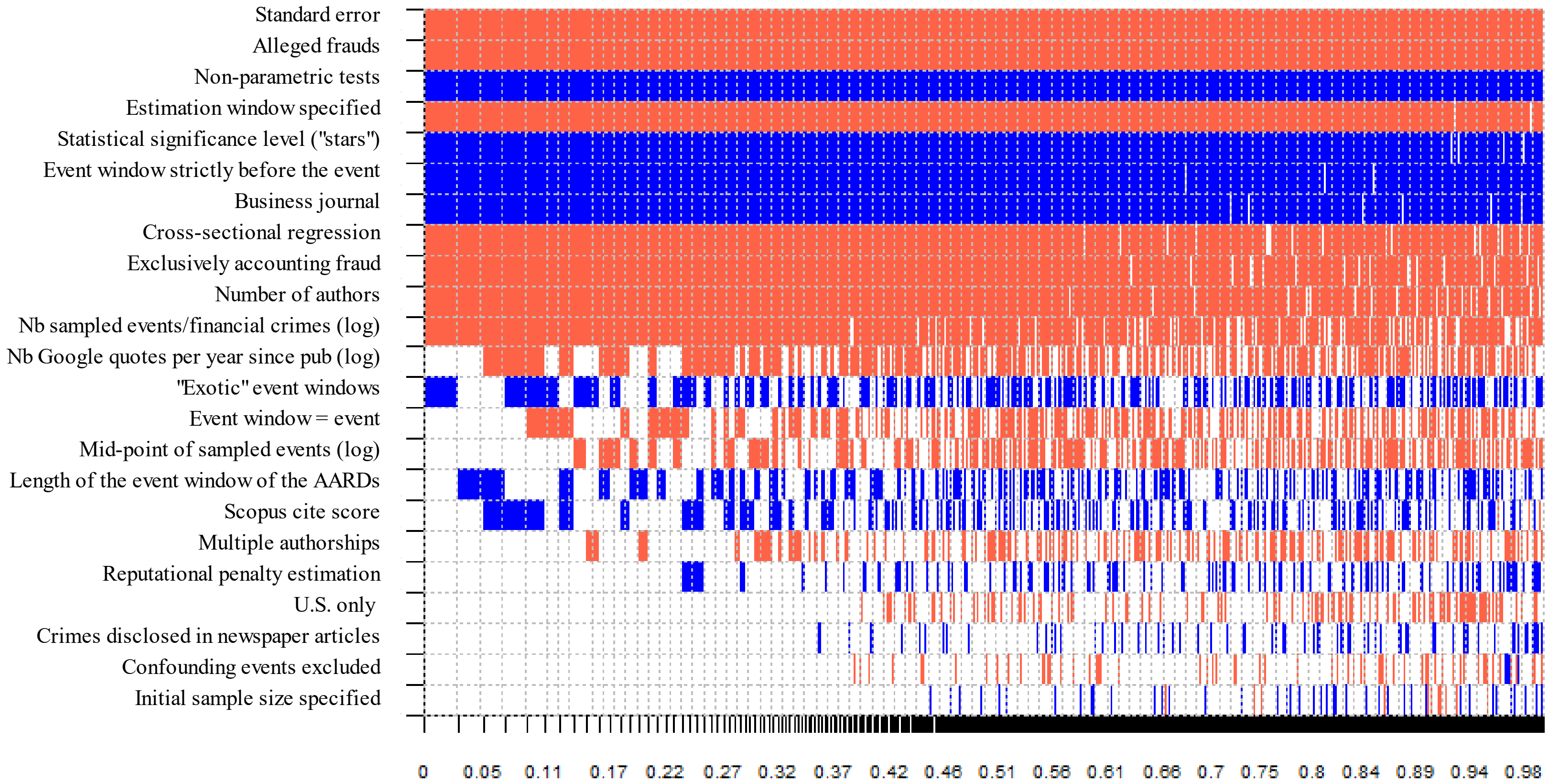
Figure 5.2: Kernel Density Estimate of *t*-Statistics.



Source: Authors' calculations.

Figure 6: Model Inclusion in Bayesian Model Averaging

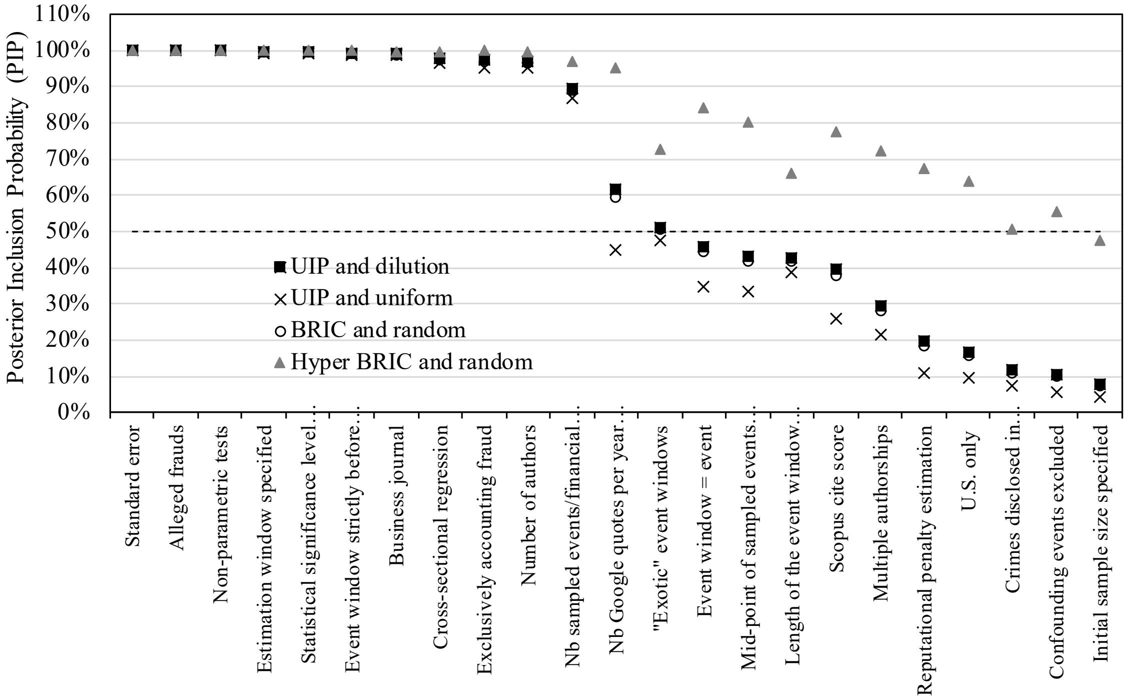
Figure 6 depicts the model inclusion in Bayesian Model Averaging, with the average abnormal returns *per* day (AARDs) as the response variable. Each column denotes an individual model. Our baseline specification uses the unit information prior recommended by Eicher et al. (2011) and the dilution prior 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 1. 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 J.



Source: Authors' calculations.

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' calculations.

Appendix A

STUDIES SUBJECT TO META-ANALYSIS – KEY FACTS (FOR ONLINE PUBLICATION)

Table A.1: The Meta Dataset

Table A.1 describes the main features of the studies included in the meta-analysis. Detailed references are listed in Appendix J. 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/CAAR" is the average of all abnormal returns reported in original studies and included in the meta-analysis, together with its standard deviation. The variable "AARD" is the normalized abnormal returns published per days of the event window, whatever the event window (between -10 to +10 days around the event day). 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/ CAAR	Std. dev.	AARD	Stat. signif.	Nb. est.
Abdulmanova, Ferris, Jayaraman, Kothari	2019	WP	Regulatory securities fraud	US	2004	2013	462	-2.64%	1.05%	-0.65%	yes	2
Aggarwal, Hu, Yang	2015	Journal of Portfolio Management	Regulatory securities fraud (incl. accounting fraud)	CN	2001	2011	750	-1.65%	1.15%	-0.56%	yes/no	5
Agrawal, Chadha	2005	Journal of Law and Economics	Accounting fraud	US	2000	2001	119	-4.93%	1.01%	-1.99%	yes/no	2
Agrawal, Cooper	2017	Quarterly Journal of Finance	Accounting fraud	US	1997	2002	419	-11.9%	1.38%	-2.13%	yes	3
Akhigbe, Kudla, Madura	2005	Applied Financial Economics	Accounting fraud	US	1991	2001	77	-6.74%	3.55%	-2.25%	yes	1
Amoah	2013	Advances in Public Interest Accounting	Regulatory securities fraud	US	1996	2006	301	-22.98%	0.19%	-7.66%	yes	2
Amoah, Tang	2010	Advances in Accounting	Accounting fraud	US	1997	2002	143	-5.54%	3.77%	-1.85%	yes/no	2
Andersen, Gilbert, Tourani-Rad	2013	JASSA	Regulatory securities fraud	AU	2004	2012	18	-2.89%	1.47%	-1.08%	yes	7
Anderson, Yohn	2002	WP	Accounting fraud	US	1997	1999	4	-15.45%	-	-2.21%	yes	1
Armour, Mayer, Polo	2017	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	UK	2001	2011	40	-1.37%	0.27%	-0.80%	yes	3
Arnold, Engelen	2007	Management & Marketing	Regulatory securities fraud (incl. accounting fraud)	BE, NL	1994	2003	57	-0.93%	0.66%	-0.93%	yes/no	6
Baker, Edelman, Powell	1999	Business and Professional Ethics Journal	Regulatory securities fraud	US	1991	1996	14	-1.13%	1.40%	-0.19%	yes/no	8
Barabanov, Ozocak, Turtle, Walker	2008	Financial Management	Regulatory securities fraud	US	1996	2003	623	-4.70%	-	-1.57%	yes	1
Bardos, Golec, Harding	2013	Journal of Financial Research	Accounting fraud	US	1997	2002	166	-20.58%	-	-10.29%	yes	1
Bardos, Mishra	2014	Applied Financial Economics	Accounting fraud	US	1997	2002	24	-11.0%	9.89%	-5.50%	yes	2
Barniv, Cao	2009	Journal of Accounting and Public Policy	Accounting fraud	US	1995	2003	61	-20.26%	-	-6.75%	yes	1
Bauer, Braun	2010	Financial Analytical Journal	Regulatory securities fraud (incl. accounting fraud)	US	1996	2007	648	-5.89%	3.98%	-1.12%	yes	15

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/ CAAR	Std. dev.	AARD	Stat. signif.	Nb. est.
Beasley, Carcello, Hermanson, Neal	2010	COSO	Accounting fraud	US	1998	2007	213	-6.33%	5.32%	-4.83%	yes/no	8
Beneish	1999	The Accounting Review	Accounting fraud	US	1987	1993	50	-23.23%	3.16%	-4.18%	yes	3
Bhagat, Bizjak, Coles	1998	Financial Management	Regulatory securities fraud	US	1981	1983	46	-2.71%	-	-1.36%	yes	1
Billings, Klein, Zur	2012	WP	Regulatory securities fraud (incl. accounting fraud)	US	1996	2008	408	-2.31%	1.05%	-0.34%	yes	3
Bohn, Choi	1996	University of Pennsylvania Law Review	Regulatory securities fraud	US	1975	1986	103	-2.32%	1.44%	-1.21%	yes	2
Bonini, Boraschi	2010	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	US	1996	2005	686	-8.07%	6.60%	-1.78%	yes	11
Bowen, Call, Rajgopal	2010	The Accounting Review	Regulatory securities fraud (incl. accounting fraud)	US	1989	1996	78	-2.84%	-	-0.57%	yes	1
Bradley, Cline, Lian	2014	Journal of Corporate Finance	Regulatory securities fraud (incl. accounting fraud)	US	1996	2011	1530	-3.68%	-	-0.61%	yes	1
Brous, Leggett	1996	Journal of Financial Research	Regulatory securities fraud (incl. accounting fraud)	US	1989	1991	62	-0.39%	3.61%	-0.39%	yes/no	3
Burns, Khedia	2006	Journal of Financial Economics	Accounting fraud	US	1997	2001	215	-9.34%	1.92%	-1.98%	yes	4
Callen, Livnat, Segal	2006	Journal of Investing	Accounting fraud	US	1986	2001	385	-8.30%	-	-2.77%	yes	1
Chava, Cheng, Huang, Lobo	2010	International Journal of Law and Management	Regulatory securities fraud	US	1995	2004	85	-1.30%	-	-0.70%	yes	1
Chen, Firth, Gao, Rui	2005	Journal of Accounting and Public Policy	Regulatory securities fraud	CN	1999	2003	169	-0.92%	1.09%	-0.21%	yes/no	10
Choi, Karpoff, Lou, Martin	2019	WP	Regulatory securities fraud (incl. accounting fraud)	US	1978	2015	942	-14.90%	-	-14.90%	yes	1
Choi, Pritchard	2016	Journal of Legal Studies	Regulatory securities fraud	US	2004	2007	231	-6.54%	5.31%	-6.54%	yes	3
Christensen, Paik, Williams	2010	Journal of Forensic & Investigative Accounting	Regulatory securities fraud (incl. accounting fraud)	US	2001	2003	151	-7.18%	11.47%	-2.09%	yes/no	6
Cook, Grove	2009	Journal of Forensic & Investigative Accounting	Regulatory securities fraud (incl. accounting fraud)	US	1984	2005	88	-12.36%	6.20%	-3.44%	yes	14
Correia, Klausner	2012	WP	Accounting fraud	US	2000	2011	683	-14.94%	5.90%	-4.98%	yes	2
Cox, Weirich	2002	Managerial Auditing Journal	Accounting fraud	US	1992	1999	27	-5.25%	1.11%	-4.16%	yes	3
Davidson, Worrell, Lee	1994	Journal of Business Ethics	Accounting fraud	US	1965	1990	34	-0.80%	1.12%	-0.57%	yes/no	16
Davis, Taghipour, Walker	2017	Managerial Finance	Regulatory securities fraud	US	1996	2013	2153	0.40%	2.60%	0.40%	yes	2
de Batz	2020	European Journal of Law and Economics	Regulatory securities fraud (incl. accounting fraud)	FR	2004	2016	52	-0.80%	0.33%	-0.25%	yes/no	10
Dechow, Sloane, Sweeney	1996	Contemporary Accounting Research	Accounting fraud	US	1982	1992	78	-8.80%	-	-8.80%	yes	1
Deng, Willis, Xu	2014	Journal of Financial and Quantitative Analysis	Regulatory securities fraud (incl. accounting fraud)	US	1996	2006	156	-10.00%	5.49%	-1.69%	yes	6
Desai, Hogan, Wilkins	2006	The Accounting Review	Accounting fraud	US	1997	1998	146	-11.07%	-	-3.69%	yes	1
Djama	2013	Revue Française de Gestion	Accounting fraud	FR	1995	2008	36	-0.67%	3.02%	0.29%	yes/no	3
Du	2017	Journal of Business Finance & Accounting	Accounting fraud	US	2001	2011	17	-5.50%	0.06%	-2.29%	yes	2

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/ CAAR	Std. dev.	AARD	Stat. signif.	Nb. est.
Engelen	2009	WP	Regulatory securities fraud	BE, DE, FR, LU, NL, UK	1995	2005	83	-2.39%	1.94%	-0.77%	yes/no	12
Engelen	2011	Book chapter	Regulatory securities fraud	BE, DE, FR, LU, NL, UK	1995	2005	101	-0.71%	1.02%	-0.71%	yes/no	6
Engelen	2012	CESifo Economic Studies	Regulatory securities fraud	US	1993	2008	122	-0.49%	0.87%	-0.49%	yes/no	3
Eryiğit	2019	Journal of Financial Crime	Accounting fraud	TR	2005	2015	160	-1.39%	0.71%	-0.12%	yes/no	4
Ewelt-Knauer, Knauer, Lachmann	2015	Journal of Business Economics	Regulatory securities fraud	DE	1998	2014	126	-11.29%	#DIV/0!	-0.26%	yes	2
Feroz, Park, Pastena	1991	Journal of accounting research	Accounting fraud	US	1982	1989	58	-0.32%	4.50%	-2.64%	yes/no	11
Ferris, Jandik, Lawless, Makhija	2007	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	US	1982	1999	194	-1.89%	-	-0.63%	yes	1
Fich, Shivdasani	2007	Journal of Financial Economics	Regulatory securities fraud	US	1998	2002	200	-5.11%	1.80%	-3.46%	yes	4
Firth, Rui, Wu	2009	Journal of Accounting and Public Policy	Regulatory securities fraud	CN	1999	2005	61	-1.18%	0.82%	-0.75%	yes/no	10
Firth, Rui, Wu	2011	Journal of Corporate Finance	Accounting fraud	CN	2000	2005	267	-0.90%	0.58%	-0.11%	yes/no	8
Firth, Wong, Xin, Yick	2016	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	CN	2003	2010	75	-0.75%	0.21%	-0.21%	yes	2
Flore, Degryse, Kolaric, Schiereck	2018	WP	Regulatory securities fraud (incl. accounting fraud)	DE, ES, FR, NL, SW, UK, US	2005	2015	251	0.20%	0.16%	0.14%	yes/no	5
Gande, Lewis	2009	Journal of Financial and Quantitative Analysis	Regulatory securities fraud	US	1996	2003	605	-4.48%	5.43%	-1.21%	yes/no	7
Gerety, Lehn	1997	Managerial and Decision Economics	Accounting fraud	US	1981	1987	37	-3.05%	-	-1.02%	yes	1
Goldman, Peyer, Stefanescu	2012	Financial Management	Accounting fraud	US	1976	2010	444	-18.90%	1.12%	-8.91%	yes	5
Griffin, Grundfest, Perino	2004	Abacus	Regulatory securities fraud	US	1990	2002	2133	-6.00%	7.22%	-1.83%	yes/no	4
Griffin, Sun	2016	Accounting and Finance Research	Regulatory securities fraud	US	2001	2007	80	-3.12%	1.68%	-0.84%	yes/no	4
Haslem, Hutton, Hoffmann Smith	2017	Financial Management	Regulatory securities fraud	US	1995	2006	594	-8.33%	4.52%	-0.84%	yes	6
Hirschey, Palmrose, Scholz	2005	WP	Accounting fraud	US	1995	1999	405	-21.80%	-	-7.27%	yes	1
Humphery-Jenner	2012	Journal of Financial Intermediation	Regulatory securities fraud	US	1996	2007	416	-4.56%	2.47%	-1.07%	yes	5
Iqbal, Shetty, Wang	2007	Journal of Financial Research	Regulatory securities fraud	US	1996	2003	298	-10.04%	8.64%	-5.04%	yes	10
Johnson, Ryan, Tian	2003	WP	Accounting fraud	US	1992	2005	87	-14.70%	-	-4.90%	yes	1
Jordan, Peek, Rosengren	2000	Journal of Financial Intermediation	Regulatory securities fraud (incl. accounting fraud)	US	1989	1994	35	-4.96%	-	-1.65%	yes	1
Karpoff, Koester, Lee, Martin	2017	The Accounting Review	Accounting fraud	US	1978	2011	1052	-15.17%	-	-15.17%	yes	1
Karpoff, Lee, Martin	2008	Journal of financial and quantitative analysis	Accounting fraud	US	1978	2002	371	-11.17%	7.83%	-11.17%	yes	6
Karpoff, Lott	1993	Journal of Law and Economics	Accounting fraud	US	1978	1987	4	-2.55%	2.24%	-1.27%	yes/no	5

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/ CAAR	Std. dev.	AARD	Stat. signif.	Nb. est.
Kellogg	1984	Journal of Accounting and Economics	Accounting fraud	US	1967	1979	26	-1.90%	2.84%	-1.90%	yes/no	3
Kirat, Rezaee	2019	Applied Economics	Regulatory securities fraud (incl. accounting fraud)	FR	2004	2017	54	-0.24%	0.86%	-0.02%	yes	7
Klock	2015	Journal of Business & Securities Law	Regulatory securities fraud (incl. accounting fraud)	US	1996	2012	714	-1.28%	0.69%	-1.01%	yes	4
Kouwenberg, Phunnarungsi	2013	Pacific-Basin Finance Journal	Regulatory securities fraud (incl. accounting fraud)	TH	2003	2010	111	-1.20%	1.32%	-0.75%	yes/no	4
Kravet, Shevlin	2010	Review of Accounting Studies	Accounting fraud	US	1997	2001	299	-5.40%	-	-0.77%	yes	1
Kryzanowski, Zhang	2013	Journal of Multinational Financial Management	Accounting fraud	CA	1997	2006	210	-3.57%	1.67%	-1.91%	yes	4
Kwan, Kwan	2011	International Review of Business Research Papers	Regulatory securities fraud	MY	2005	2009	41	0.32%	1.46%	0.32%	yes/no	3
Lei, Law	2019	WP	Regulatory securities fraud (incl. accounting fraud)	CN	1999	2015	1188	-0.51%	0.41%	-0.14%	yes/no	8
Liebman, Milhaupt	2008	Columbia Law Review	Regulatory securities fraud	CN	2001	2006	68	-2.73%	1.40%	-0.72%	yes/no	8
Lieser, Kolaric	2016	WP	Regulatory securities fraud (incl. accounting fraud)	US	1996	2014	1377	-5.11%	7.75%	-1.26%	yes/no	15
Loh, Rathinasamy	2003	Review of Pacific Basin Financial Markets and Policies	Regulatory securities fraud (incl. accounting fraud)	US	1996	1998	290	-1.20%	1.03%	-0.54%	yes	2
Marciukaityte, Szewczyk, Uzun, Varma	2006	Financial Analysts Journal	Regulatory securities fraud (incl. accounting fraud)	US	1978	2001	28	-7.81%	-	-3.91%	yes	1
Marciukaityte, Szewczyk, Varma	2009	Financial Analysts Journal	Accounting fraud	US	1997	2002	187	-6.58%	-	-3.29%	yes	1
McDowell	2005	WP	Accounting fraud	US	1998	2003	174	-6.23%	-	-2.08%	yes	1
Muradoglu, Clark Huskey	2008	WP	Regulatory securities fraud (incl. accounting fraud)	US	1995	2004	296	-0.58%	0.39%	-0.58%	yes/no	12
Nainar, Rai, Tartaroglu	2014	International Journal of Disclosure and Governance	Regulatory securities fraud	US	1999	2007	77	-1.98%	1.07%	-1.10%	yes/no	6
Nelson, Gilley, Trombley	2009	Securities Litigation Journal	Regulatory securities fraud	US	2002	2007	58	-2.59%	-	-2.59%	yes	1
Nourayi	1994	Journal of Accounting and Public Policy	Regulatory securities fraud (incl. accounting fraud)	US	1977	1984	82	-0.72%	0.31%	-0.24%	yes	4
Owers, Lin, Rogers	2002	International Business and Economics Research Journal	Accounting fraud	US	1994	1997	13	-15.92%	6.06%	-3.14%	yes	6
Ozbas	2008	WP	Accounting fraud	US	1999	2003	75	-9.76%	8.53%	-2.55%	yes/no	4
Ozeki	2019	Securities Analysts Journal	Accounting fraud	JP	2005	2016	218	-13.05%	3.75%	-9.13%	yes/no	2
Pereira, Malafronte, Sorwar, Nurullah	2019	Journal of Financial Services Research	Regulatory securities fraud (incl. accounting fraud)	US	2004	2015	1387	-32.30%	28.11%	-6.38%	yes/no	5
Persons	1997	Journal of Business Research	Regulatory securities fraud	US	1972	1993	95	-0.35%	0.26%	-0.28%	yes	4
Plumlee, Yohn	2008	WP	Accounting fraud	US	2003	2006	1303	-0.80%	-	-0.27%	yes	1
Pritchard, Ferris	2001	WP	Regulatory securities fraud	US	1995	1999	89	-9.35%	13.69%	-3.12%	yes/no	3
Romano	1991	Journal of Law, Economics, and Organization	Regulatory securities fraud	US	1970	1987	66	-1.09%	1.82%	-0.79%	yes/no	6

Author(s)	Pub. year	Publication outlet	Financial crimes	Countries	Sample period		Sample size	AAR/ CAAR	Std. dev.	AARD	Stat. signif.	Nb. est.
Scholz	2008	U.S. Department of Treasury	Accounting fraud	US	1997	2006	264	-13.00%	-	-6.50%	yes	1
Slovin, Sushka, Polonchek	1999	Journal of Financial Economics	Regulatory securities fraud (incl. accounting fraud)	US	1975	1992	61	-5.29%	0.16%	-1.83%	yes	2
Song, Han	2017	Journal of Business Ethics	Regulatory securities fraud (incl. accounting fraud)	KR	2001	2010	220	-4.39%	1.69%	-0.70%	yes	3
Sun, Zhang	2006	WP	Regulatory securities fraud	CN	1990	2002	144	-1.40%	-	-0.47%	yes	1
Takmaz, Keles	2017	Journal of Business Research Turk	Regulatory securities fraud	TR	2007	2016	72	-1.49%	1.42%	-0.24%	yes/no	4
Tanimura, Okamoto	2013	Asian Economic Journal	Accounting fraud	JP	2000	2008	39	-6.20%	-	-3.10%	yes	1
Tay, Puah, Brahmana, Abdul Malek	2016	Journal of Financial Crime	Regulatory securities fraud (incl. accounting fraud)	MY	1996	2013	17	-0.53%	0.04%	-0.53%	no	3
Wang, Ashton, Jaafar	2019	The British Accounting Review	Accounting fraud	CN	2007	2016	433	-0.30%	0.17%	-0.08%	yes/no	7
Wang, Wu	2011	China Journal of Accounting Research	Accounting fraud	CN	1999	2005	67	-0.14%	0.81%	-0.05%	yes/no	5
Wu	2002	WP	Accounting fraud	US	1977	2000	932	-23.00%	-	-7.67%	yes	1
Wu, Zhang	2014	China Journal of Accounting Studies	Regulatory securities fraud	CN	2002	2011	157	-3.24%	2.39%	-2.08%	yes	6
Xu, Xu	2020	International Review of Law and Economics	Regulatory securities fraud (incl. accounting fraud)	CN	2014	2018	107	-1.75%	4.15%	-0.72%	yes/no	10
Yu, Zhang, Zheng	2015	Financial Management	Accounting fraud	CN	1999	2011	195	-3.01%	3.30%	-0.58%	yes	2
Zeidan	2013	Journal of Business Ethics	Regulatory securities fraud	US	1990	2009	163	-0.75%	0.65%	-0.50%	yes/no	4
Zhu, Hu	2010	WP	Accounting fraud	CN	2006	2008	88	-3.03%	1.63%	-0.70%	yes/no	7
Overall	**2009**				**1994**	**2004**	**293**	**-4.77%***	**6.38%***	**-1.81%***		**4.3****

*Source: Authors * Winsorized at the 1% level. ** Average between studies.*

Appendix B

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)).[1] 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:

- 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.
- 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.

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.

---

[1] 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.).

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 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:*

$$AR_{i,t} = R_{i,t} - \widehat{\alpha_i} - \widehat{\beta_i}\, R_{m,t} \ . \tag{III}$$

$R_{i,t}$ 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.

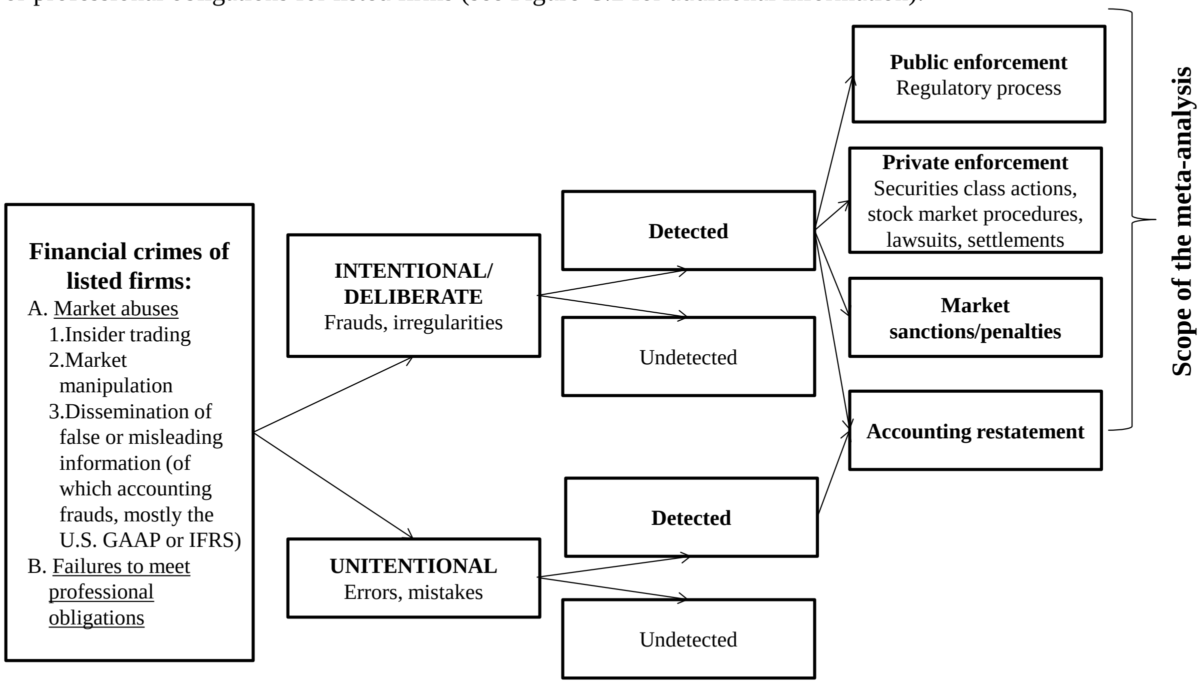
$$CAAR_{[t_1;t_2]} = \frac{1}{n}\sum_{i=1}^{n} CAR_{i,[t_1;t_2]} = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=t_1}^{t_2} AR_{i,t} \ . \tag{IV}$$

Appendix C

SCOPE OF THE META-ANALYSIS (FOR ONLINE PUBLICATION)

Figure C.1: Graphical Presentation of the Scope of the Meta-Analysis

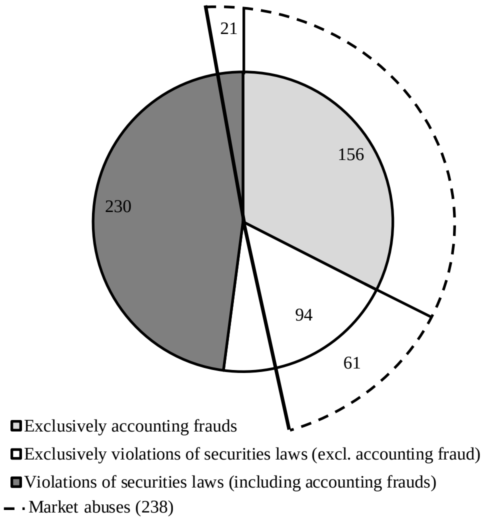
This figure C.1 graphically describes the inclusion criteria of the meta-analysis. From a wide range of studies on financial crimes by listed firms, the scope was reduced to the literature investigating detected and intentional crimes and the subsequent market reactions that are based on an event-study methodology. Financial crimes cover the following range of misconduct: market abuses with insider trading, price manipulation, 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 (see Figure C.2 for additional information).



Source: Authors

Figure C.2: Scope of Financial Crimes in the Sample

Financial crimes committed by listed firms are defined as follows, in line with the academic, practitioner, and policy literature: insider dealing,[2] price manipulation,[3] breach of public disclosure requirements[4] (*i.e.* the three market abuses), and more generally breaches of financial regulations.[5] Based on this definition, the scope of 480 estimates collected for this meta-analysis can be graphically depicted as in the following figure. The 21 estimates of violations of securities laws which are market abuses were collected in 3 articles on China and the U.S. The 61 estimates of exclusively (non-accounting) violations of securities laws were collected in 11 articles, most frequently regarding insider trading (complemented with two articles on price manipulations and breaches to disclosures to the SEC).



Source: Authors

[2] Divulgence and/or use of insider information for investment decisions, at the expense of shareholders. Typical examples include: an insider recommends (or encourages) to carry out an operation based on a financial instrument related to (or based on) the insider information; a person knowingly uses this recommendation or incentive (i.e. acknowledging that this information is based on insider information); a person knowingly shares this recommendation or incentive; a person front-runs client's orders.

[3] Deliberate misconduct to influence securities prices and fair price formation. Price manipulation happens when a person carries out an operation, places an order, or behaves in a way that gives – or is likely to give – false or misleading signals on the supply, demand or price of financial instruments, or that fixes – or is likely to fix – the price of a financial instrument at an abnormal or artificial level. Examples include: end-of-day manipulation, matched orders, circular trading, reference price influence, improper order handling (churning, wash trades, spoofing), or boiler-room operation.

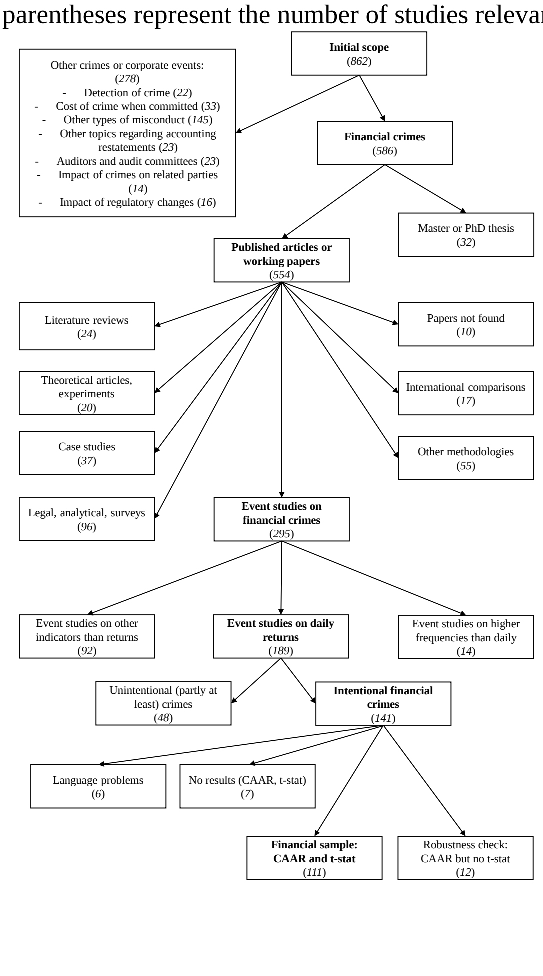
[4] When a person discloses information (whatever the medium) likely to give false or misleading indications on the health and the perspectives of an issuer or on the demand, supply or price of a financial instrument (in particular breaches to financial reporting). For example, the failure to comply with financial reporting laws and regulations, most frequently leading to the issuance of a financial restatement.

[5] Failure to comply with professional obligations, such as failures of internal control, of the management of conflicts of interest, of the obligations relative to anti-money laundering and countering the financing of terrorism, etc.

Figure C.3: Data Collection process and PRISMA Statement

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 crimes 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.[6]

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

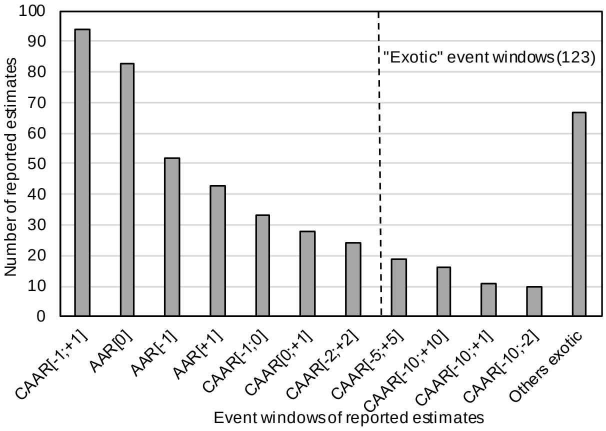
[6] We tried to circumvent the fact that language issues can act as a constraint on the scope of meta-analyses. We extended searches to the following languages: English, French, German, Portuguese, and Spanish. Some articles in Chinese, Japanese, and Turkish could not be included in the literature review, 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.

Appendix D

ADDITIONAL ANALYSES BY EVENT WINDOWS (FOR ONLINE PUBLICATION)

Figures D.1: Event Windows of the Estimates included in the sample

Figure D.1 depicts, by declining order of frequency, the most frequent event windows of the estimates included in the sample. Some granular information is also given for the “exotic” event windows: the 123 estimates with event windows standing for less than 5% of the sample of 480 estimates (*i.e.* less than 24 estimates). Exotic event windows account for ¼ of the sample.

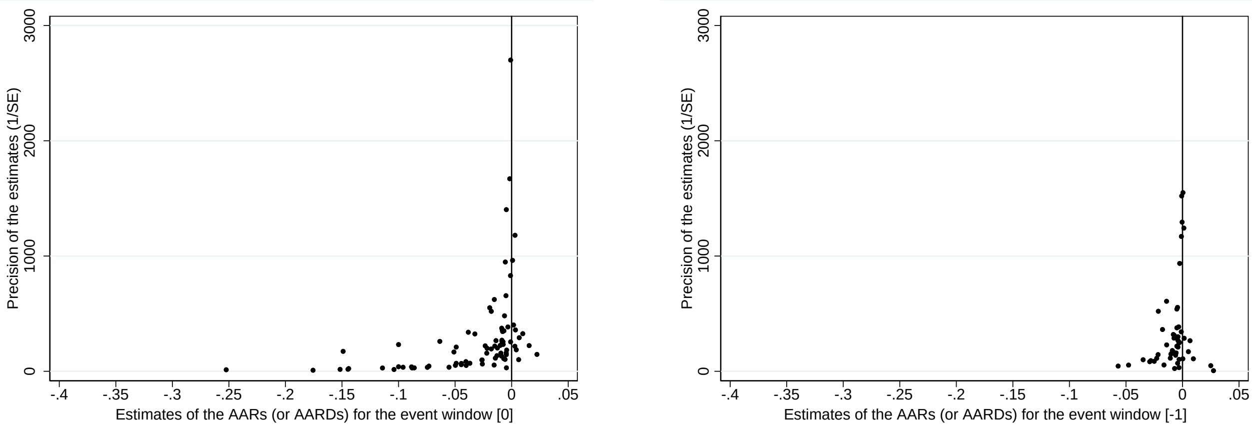


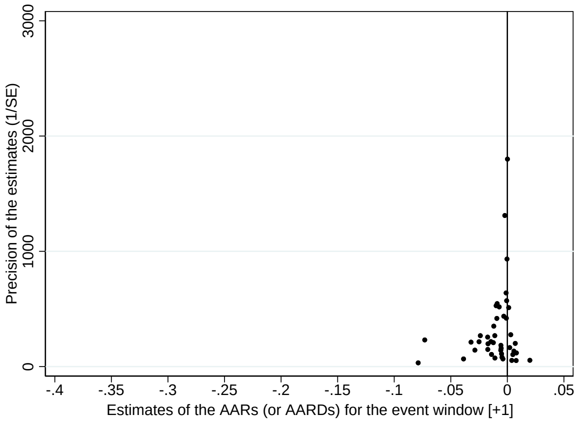
Source: Authors

Figures D.2: Funnel Plots by Event Windows

Figures I.2 detail 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 480 estimates: at least 24 estimates). The following event windows are deemed “frequent”, 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 to maximize their probability of publication. The first Figures I.2.1 display funnel plots for one-day event windows, *i.e.* for which abnormal returns are the same for the original and the normalized samples. The second Figures I.2.2 compare Panels A ((C)AARs) with Panel B (AARDs) for longer event windows ([-1;+1], [-1;0], [0;+1], and [-2;+2]) and exotic event windows.

Figures D.2.1: One-Day Event Windows ([0], [-1], and [+1]): Average Abnormal Returns

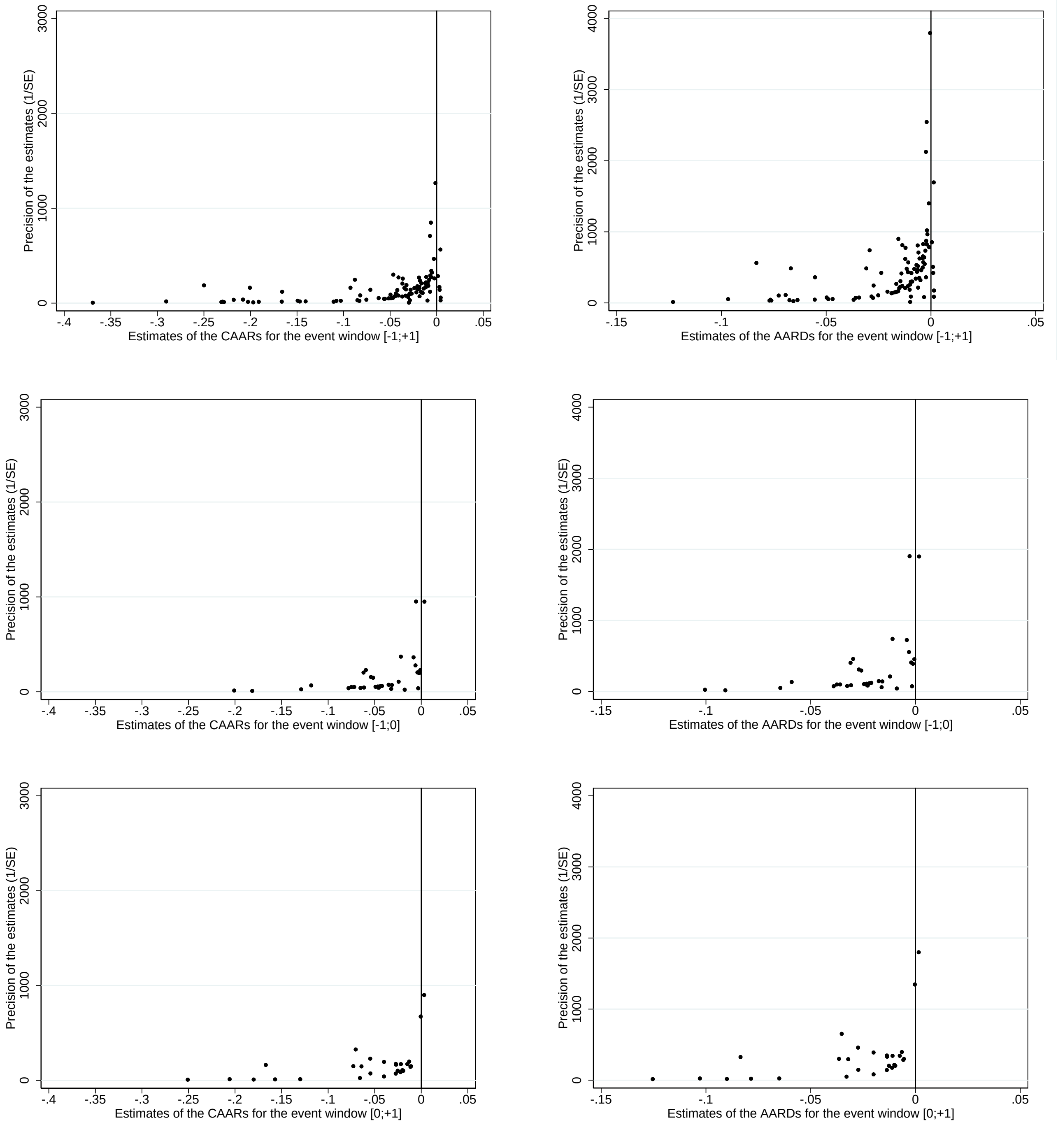


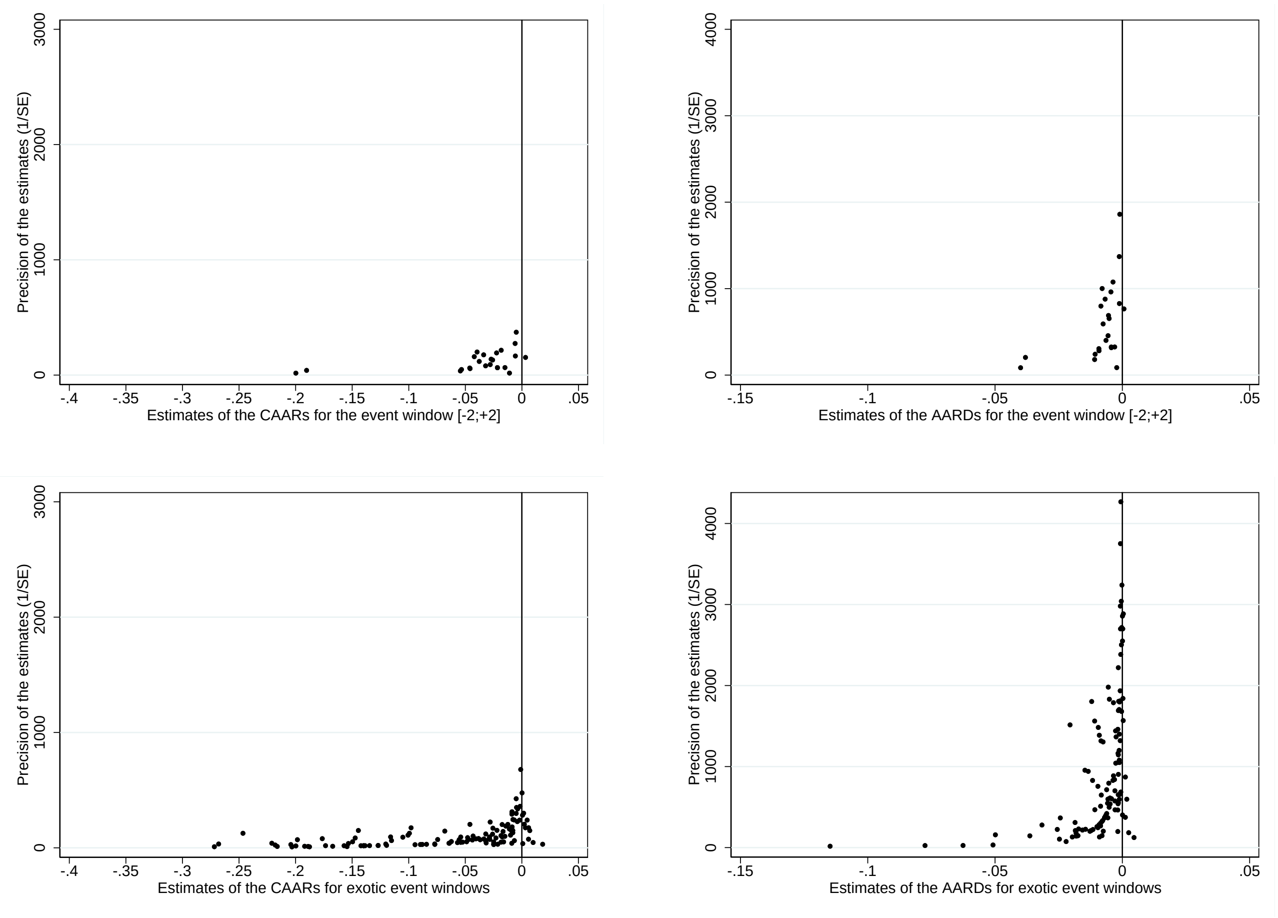


Precision of the estimates (1/SE)
0
1000
2000
3000
-.4
-.35
-.3
-.25
-.2
-.15
-.1
-.05
0
.05
Estimates of the AARs (or AARDs) for the event window [+1]

Figures D.2.2: Longer ([-1;+1], [-1;0], [0;+1], and [-2;+2]) and "Exotic" Event Windows. Cumulative Average Abnormal Returns (Panel A on the Left Handside) versus Cumulative Average Abnormal Returns per Day (Panel B on the Right Handside)

Panel B. CAARs | **Panel B. AARDs**





Source: Authors

Table D.1: Meta-Regression Analysis of Publication Selection Bias by Event Windows

Table D.1 details the results of the publication selection bias analysis, based on the FAT-PET tests (Eq. (1)) for the full sample of reported estimates ((C)AARs, column 1), for four sub-samples reflecting the four most frequent event windows ($\text{CAAR}[-1;+1]$, $\text{AAR}[0]$, $\text{AAR}[-1]$, and $\text{AAR}[+1]$, columns 2 to 5), and for the normalized abnormal returns (AARDs, column 6). 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). Most of the results (except for the $[+1]$ event window, but with a smaller sample size) concur with strong and significant publication bias leading to lower though negative and significant abnormal returns in the aftermath of the publication of a financial crime.

	All (C)AARs [1]	CAAR[-1;+1] [2]	AAR[0] [3]]	AAR[-1] [4]	AAR[+1] [5]	All AARDs [6]
<u>Panel A.</u> Unweighted estimations						
1. OLS						
SE (*publication bias*)	-1.52 ***	-1.42 ***	-1.21 *	0.11 *	-0.71	-1.48 ***
	(0.275)	(0.414)	(0.682)	(0.063)	(0.867)	(0.262)
Intercept (*effect beyond bias*)	-1.91% ***	-3.09% ***	-1.69% **	-0.95% ***	-0.53%	-0.63% ***
	(0.005)	(0.009)	(0.008)	(0.002)	(0.005)	(0.002)
2. Study-level fixed effects						
SE (*publication bias*)	-1.86 ***	-1.33	-2.71 ***	-0.23	-1.07	-1.45 ***
	(0.142)	(0.905)	(0.256)	(0.282)	(0.690)	(0.120)
Intercept (*effect beyond bias*)	-1.27% ***	-3.29%	0.54%	-0.59% *	-0.28%	-0.66% ***
	(0.003)	(0.021)	(0.004)	(0.003)	(0.005)	(0.001)
3. Study-level between effects						
SE (*publication bias*)	-1.50 ***	-1.24 ***	-0.71 ***	0.14 *	-0.26	-1.59 ***
	(0.143)	(0.192)	(0.213)	(0.077)	(0.581)	(0.148)
Intercept (*effect beyond bias*)	-2.45% ***	-3.68% ***	-2.18% ***	-0.93% ***	-0.76%	-0.82% ***
	(0.005)	(0.009)	(0.007)	(0.002)	(0.005)	(0.002)
4. Hierarchical Bayes						
SE (*publication bias*)	-1.61 ***	-1.11 ***	-1.16 ***	0.06 ***	-0.04 ***	-1.47 ***
	(0.213)	(0.625)	(0.470)	(0.580)	(0.560)	(0.219)
Intercept (*effect beyond bias*)	-1.80% ***	-4.10% ***	-1.50% ***	-0.96% ***	-1.00% ***	-0.74% ***
	(0.027)	(0.084)	(0.091)	(0.130)	(0.150)	(0.027)
5. IV number of observations reported by study						
SE (*publication bias*)	-3.08 ***	-1.34	-2.24 *	-0.06	2.75 *	-1.68 ***
	(0.618)	(1.082)	(1.146)	(0.422)	(1.581)	(0.436)
Intercept (*effect beyond bias*)	1.03%	-3.25%	-0.15%	-0.78%	-2.92% **	-0.47% *
	(0.010)	(0.025)	(0.012)	(0.005)	(0.014)	(0.003)
<u>Panel B</u>. Weighted least square estimations						
1. Weighted by the precision (inverse of the standard error)						
SE (*publication bias*)	-2.23 ***	-2.15 ***	-2.05 ***	-0.48	-0.91 *	-1.93 ***
	(0.228)	(0.326)	(0.431)	(0.330)	(0.497)	(0.204)
Intercept (*effect beyond bias*)	-0.58% ***	-1.44% ***	-0.44% *	-0.33% **	-0.39% *	-0.27% ***
	(0.002)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)

	All (C)AARs [1]		CAAR[-1;+1] [2]		AAR[0] [3]]		AAR[-1] [4]		AAR[+1] [5]		All AARDs [6]	
2. Weighted by the inverse of the number of estimates reported by study												
SE (*publication bias*)	-1.50	***	-1.54	***	-1.28	**	0.14	***	-0.08		-1.55	***
	(0.305)		(0.458)		(0.619)		(0.028)		(0.868)		(0.287)	
Intercept (*effect beyond bias*)	-2.43%	***	-3.67%	***	-2.45%	**	-0.88%	***	-0.67%		-0.86%	***
	(0.006)		(0.017)		(0.012)		(0.002)		(0.005)		(0.003)	
Panel C. Non-linear estimations												
1. Weighted average of adequately powered (Ioannidis et al., 2017)												
Effect beyond bias	-0.21%	***	-0.69%	***	-0.28%	***	-0.26%	***	-0.03%		-0.15%	***
	(0.009)		(0.003)		(0.001)		(0.001)		(0.001)		(0.003)	
2. Selection model (Andrews and Kasy, 2019)[2]												
Effect beyond bias	-0.56%	*	-0.50%	***	-0.45%	*	-0.73%	***	-0.90%	***	-0.43%	***
	(0.132)		(0.175)		(0.172)		(0.207)		(0.241)		(0.062)	
3. Stem-based method (Furukawa, 2019)												
Effect beyond bias	-0.36%		-0.38%		-0.50%		-0.03%		-0.08%		-0.14%	
	(0.008)		(0.014)		(0.006)		(0.004)		(0.004)		(0.005)	
Number of observations[1]	**480**		**94**		**83**		**52**		**43**		**480**	

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 H, Figures H.1.

Appendix E

MAIN FEATURES OF FINANCIAL CRIMES AND ENFORCEMENT (FOR ONLINE PUBLICATION)

Table E.1: Main Features of Some Securities Enforcers

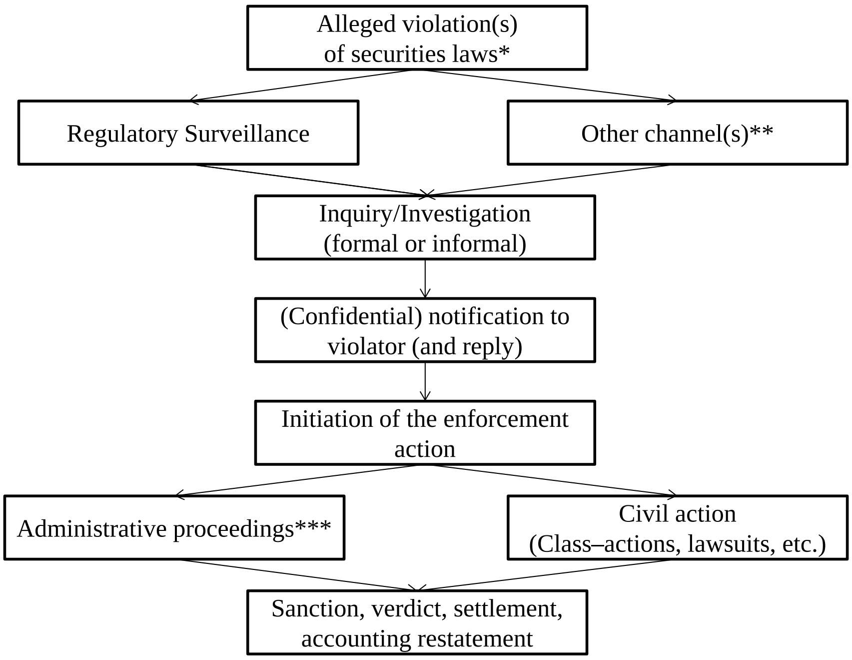
Table E.1 compares the main features of securities law enforcement in the four most frequent countries in the sample: the U.S., China, the UK, and France. Each country has its own enforcement mix, with different weights given to public (higher in code-law countries) or private (higher in common-law countries, 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 E.1: Common Features of Financial Crime Prosecution

Figure E.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.).

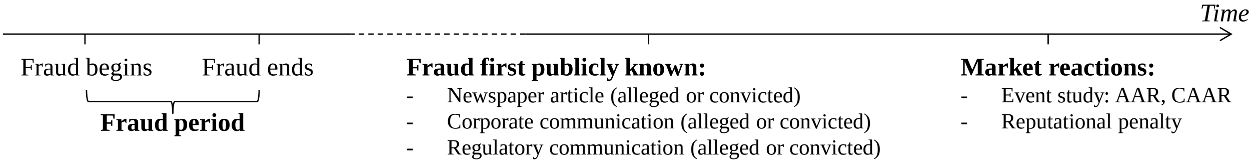
** Self-regulatory organizations (stock exchanges, Justice Ministries, etc.), media, external auditors, complaints from shareholders or stakeholders, whistleblowing, etc.

*** 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 E.2: Chronology of Market Reactions to Financial Crimes

This figure E.2 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 F

CORRELATION MATRIX OF EXPLANATORY VARIABLES (FOR ONLINE PUBLICATION)

Table F.1: Correlation Matrix

Table F.1 depicts the correlation matrix of the 22 explanatory variables included in the BMA. A detailed description of all variables is available in Table 1.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]
Only U.S. [1]	1																					
Mid-point year [2]	-0.42*	1																				
Exclusively accounting frauds [3]	0.12*	-0.17*	1																			
Alleged frauds [4]	0.47*	-0.33*	-0.08	1																		
Crimes disclosed by firms [5]	0.15*	-0.27*	-0.10*	0.30*	1																	
Initial sample size published [6]	0.09	-0.10*	0.08	0.12*	-0.12*	1																
Confounding events excluded [7]	-0.13*	0.00	-0.17*	0.08	0.16*	-0.01	1															
Number of observations (log events in the sample) [8]	0.09	0.33*	-0.17*	-0.08	-0.19*	0.09*	-0.36*	1														
Estimation window specified [9]	-0.09*	-0.04	-0.20*	-0.03	0.21*	-0.09	0.06	-0.09	1													
Length of the event window [10]	0.00	0.16*	-0.02	-0.05	0.01	-0.06	0.05	0.12*	0.08	1												
Event window strictly before the event [11]	-0.04	-0.04	-0.06	0.04	0.08	0.04	0.01	0.00	-0.01	-0.13*	1											
Event window = event [12]	0.03	-0.07	-0.01	0.06	-0.03	0.05	-0.01	0.01	-0.06	-0.31*	-0.18*	1										
"Exotic" event window [13]	0.01	0.17*	0.03	-0.05	0.02	-0.05	0.01	0.16*	-0.02	0.79*	-0.03	-0.27*	1									
Signif. level ("stars") [14]	-0.02	0.00	-0.21*	0.06	0.12*	-0.01	0.04	0.06	0.08	0.01	0.00	-0.02	-0.06	1								
z-statistics [15]	-0.35*	0.19*	-0.07	-0.15*	0.04	0.12*	0.16*	0.03	0.25*	-0.07	0.00	-0.01	-0.07	0.07	1							
Cross-sectional regression [16]	-0.03	0.27*	0.02	-0.23*	-0.07	-0.05	-0.13*	0.31*	-0.25*	0.06	-0.05	-0.02	0.09*	-0.01	0.09	1						
Reputational penalty estimation [17]	-0.15*	0.08	-0.15*	-0.15*	0.00	-0.08	0.03	-0.02	0.14*	0.07	0.00	-0.06	0.02	-0.02	-0.08	0.07	1					
Nb authors [18]	0.03	0.05	0.04	-0.10*	0.03	-0.10*	-0.17*	0.11*	-0.25*	-0.02	-0.06	0.05	-0.02	0.03	0.00	0.08	-0.13*	1				
Multiple authorships [19]	-0.15*	0.12*	-0.06	-0.08	0.10*	0.05	-0.02	0.09*	-0.11*	-0.08	-0.05	0.06	-0.08	0.06	0.10*	0.17*	0.01	0.06	1			
Business journal [20]	0.10*	-0.25*	0.01	0.18*	0.07	0.06	-0.03	-0.10*	0.20*	0.01	0.03	0.00	0.01	0.09*	-0.17*	-0.25*	-0.05	0.00	-0.21*	1		
Nb Google quotes *per* year (log) [21]	0.20*	-0.29*	0.32*	0.01	-0.15*	0.15*	-0.25*	-0.05	-0.19*	-0.06	-0.07	0.04	-0.08	-0.11*	-0.13*	0.11*	0.03	0.17*	0.03	-0.13*	1	
Scopus cite score [22]	0.13*	-0.42*	0.20*	0.01	-0.04	0.22*	-0.23*	0.02	-0.08	-0.07	-0.01	0.05	-0.06	-0.08	0.06	0.24*	-0.12*	0.16*	-0.12*	0.13*	0.57*	1

*Source: Authors * Statistical significance at the 5% level.*

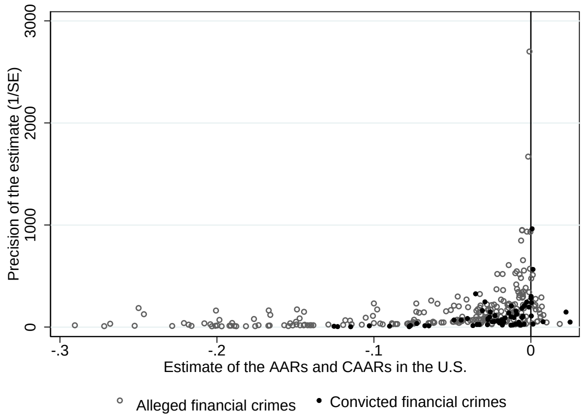
Appendix G

ADDITIONAL FUNNEL PLOTS BY GEOGRAPHY AND FINANCIAL CRIMES (FOR ONLINE PUBLICATION)

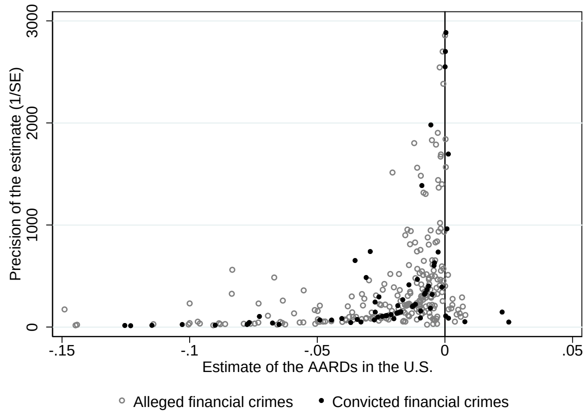
Figures G: Funnel Plots for the U.S. versus Other Countries

Figures G are funnel plots abnormal returns ((C)AARs in Panel A or AARDs in Panel B) specifically for the U.S. for Panel 1 or other countries for Panel 2, depending on whether the financial crime is alleged or convicted.

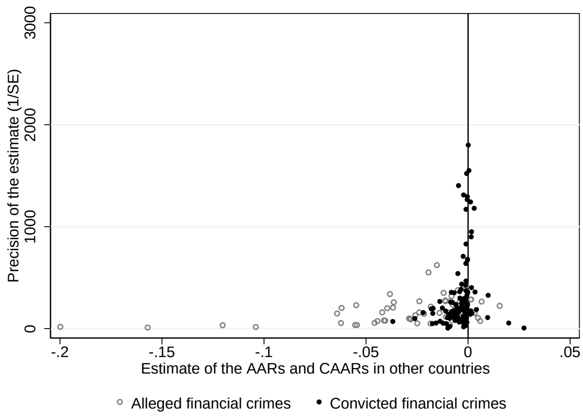
Panel A.1. (C)AARs in the U.S.



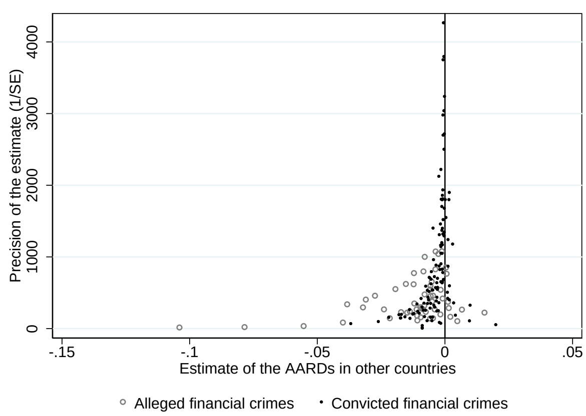
Panel B.1. AARDs in the U.S.



Panel A.2. (C)AARs in Other Countries



Panel B.2. AARDs in Other Countries



Source: Authors

Appendix H

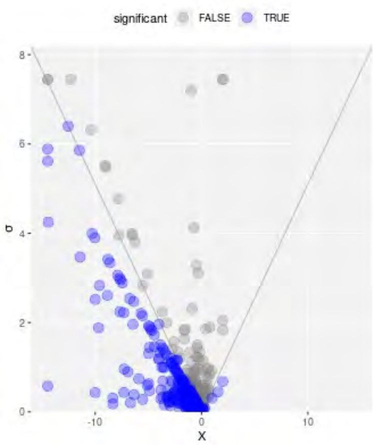
ADDITIONAL RESULTS FOR THE NON-LINEAR APPROACH BY ANDREWS AND KASY (2019)

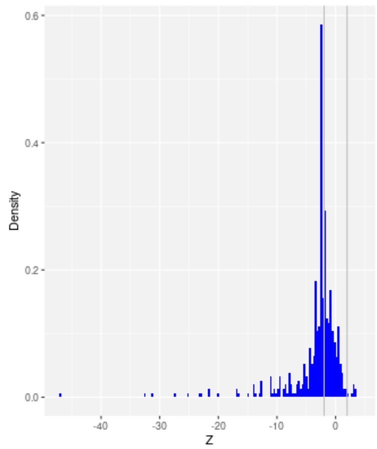
(FOR ONLINE PUBLICATION)

Figures H.1: Funnel Plots and Histograms of Z-Statistics (Non-Linear Approach by Andrews and Kasy, 2019)

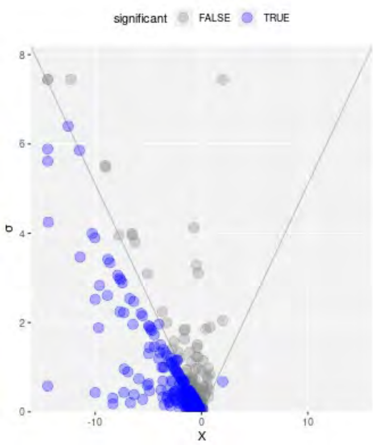
The following figures depict the funnel plots and the histograms of z-statistics for the full sample of average abnormal returns *per* day (Panel 1) and sub-samples limited to the U.S. (Panel 2.1.) or any other countries (Panel 2.2. and to exclusively accounting frauds (Panel 3.1.) or any violation of securities laws (Panel 3.2.)). The calculations are done using Maximilian Kasy's online application, which allows the estimation of models of selection publication using meta-studies.

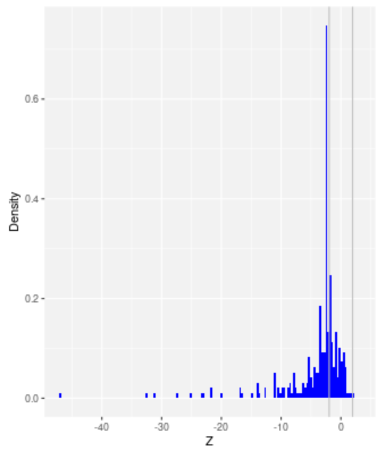
Panel 1. Full Sample of AARDs



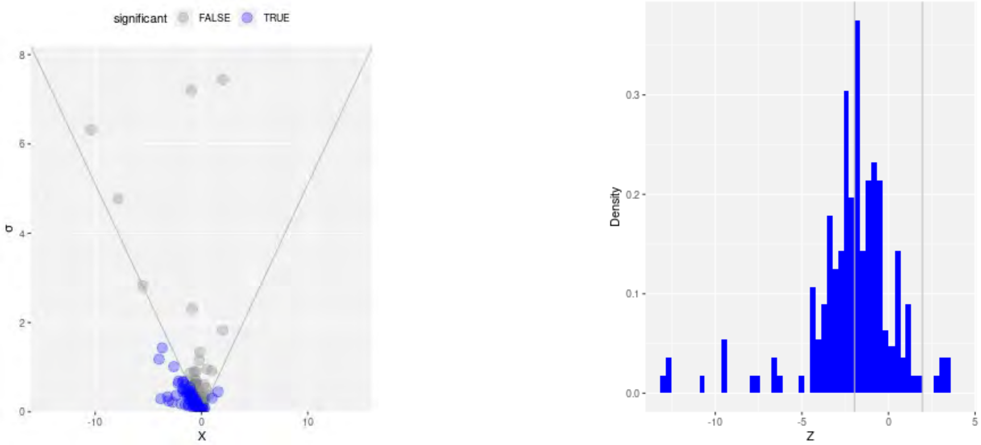


Panel 2.1. Financial Crimes Committed in the U.S.

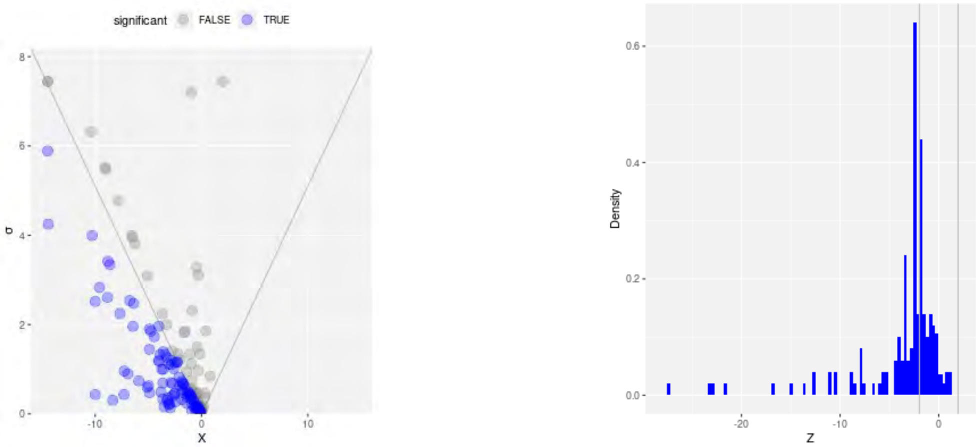




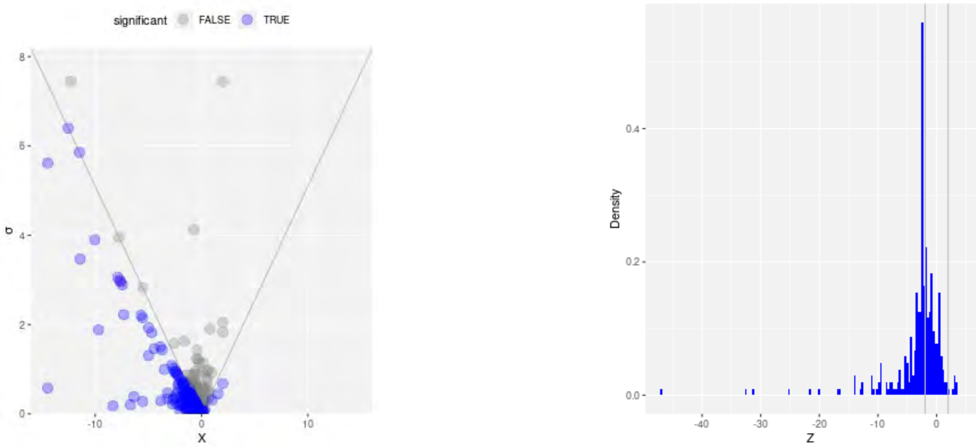
Panel 2.2. Financial Crimes Committed in Other Countries



Panel 3.1. Exclusively Accounting Frauds



Panel 3.2. Any Violation of Securities Laws



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

Appendix I

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)[7] and the augmented model by Ashenfelter et al. (1999).[8] 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),[9] 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 I.1 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 I.1: 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	-8.807	6.734			-4.799	3.936		
ω_3	-11.003	9.615			-5.571	5.201		
ω_4	-84.801	35.743			-41.335	16.095		
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	0.042	0.017	-0.035	0.002	0.042	0.013	0.007	0.011
σ	-0.109	0.007	-0.042	0.002	-0.037	0.002	-0.039	0.002
Log likelihood	1093.42		1178.17		1458.22		1203.27	
Observations	480		480		480		480	
	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.

[7] Hedges, Larry V. 1992. "Meta-Analysis." *Journal of Educational Statistics* 17(4):279-296.

[8] Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek. 1999. "A Review of Estimates of the Schooling/Earnings Relationship, With Tests for Publication Bias." *Labour Economics* 6(4):453-470.

[9] Havránek, Tomáš, and Anna Sokolova. 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.

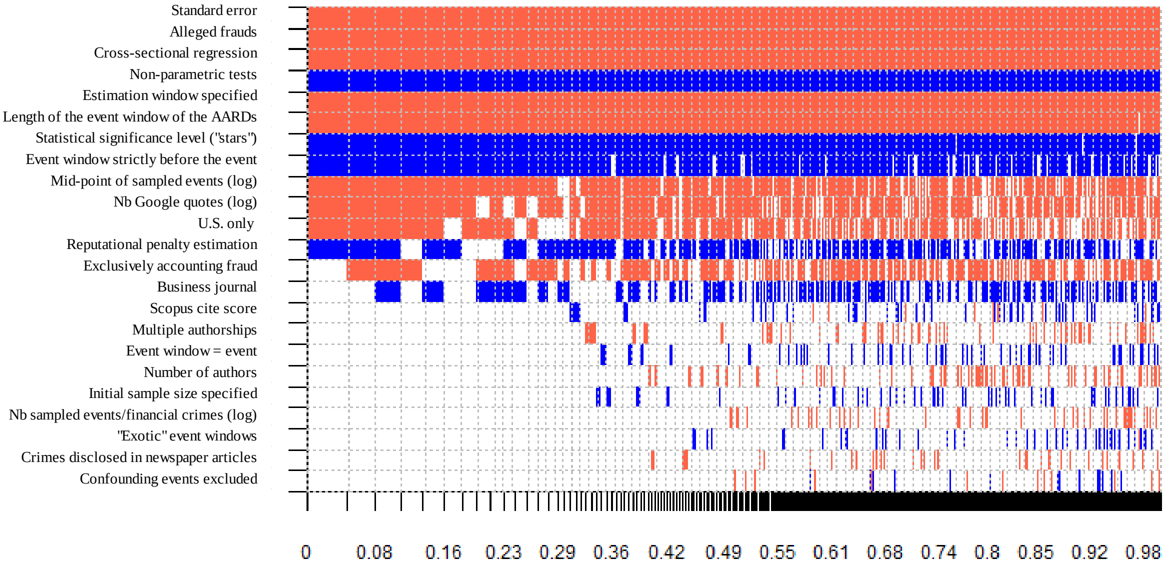
Appendix J

Robustness Checks of the BMA (for Online Publication)

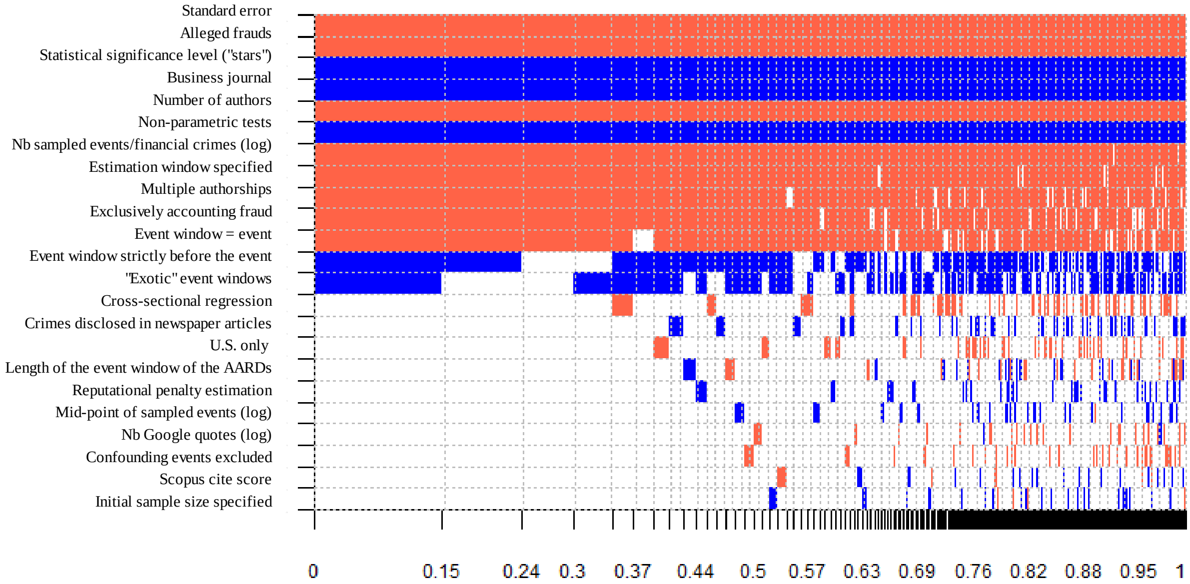
Figures J.1: Model Inclusion in Bayesian Model Averaging

Figures J.1 depict robustness checks supporting the robustness of the BMA results displayed in Figure 6. These figures depict the model inclusion in Bayesian Model Averaging with two response variables: the original sample of estimates ((C)AARs, with AARs and CAARs) in Panel 1, and normalized average abnormal returns per day (AARDs) in Panels B and C. Panel A tests the robustness of the results with the AARDs when using the original set of estimates collected from the studies ((C)AARs). Complementarily, all variables are weighted by the inverse of the number of estimates reported per study (Panel 2) or by the inverse of the standard errors (Panel 3).
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 1. 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.

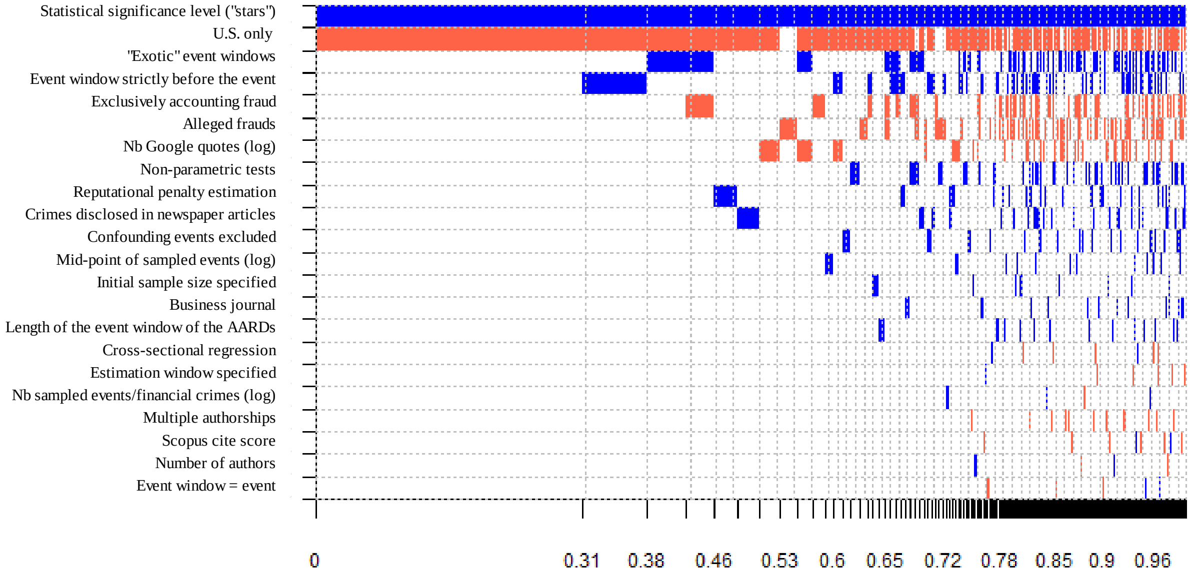
Panel 1. Model Inclusion in Bayesian Model Averaging for Average Abnormal Returns (AARs) and Cumulative Average Abnormal Returns (CAARs)



Panel 2. Average Abnormal Returns per Day (AARDs) Weighted by the Number of Estimates Reported per Study



Panel 3. Average Abnormal Returns per Day (AARDs) Weighted by the Inverse of the Standard Errors



Source: Authors

Appendix K

STUDIES INCLUDED IN THE META-ANALYSIS DATASET (FOR ONLINE PUBLICATION)

Abdulmanova, Anna, Stephen P. Ferris, Narayanan Jayaraman, and Patrik Kothari. 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, Reena, May Hu, and Jingjing Yang, 2015. "Fraud, Market Reaction, and the Role of Institutional Investors in Chinese Listed Firms." *The Journal of Portfolio Management* 41(5):92-109.

Agrawal, Anup, and Sahiba Chadha. 2005. "Corporate Governance and Accounting Scandals." *The Journal of Law and Economics* 48(2):371-406.

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

Akhigbe, Aigbe, Ronald J. Kudla, and Jeff Madura. 2005. "Why are Some Corporate Earnings Restatements More Damaging?" *Applied Financial Economics* 15(5):327-336.

Amoah, Nana Y., and Alex P. Tang. 2010. "Board, Audit Committee and Restatement-Induced Class Action Lawsuits." *Advances in Accounting* 26(2):155-169.

Amoah, Nana Y. 2013. "What is Fraud in Private Securities Lawsuit?" *Advances in Public Interest Accounting* 16:39-63.

Andersen, Angela, Aaron Gilbert, and Alireza Tourani-Rad. (2013). Breach of Continuous Disclosure in Australia. *JASSA* 4:21-26.

Anderson, Kirsten L., and Teri Lombardi Yohn. 2002. "The Effect of 10-K Restatements on Firm Value, Information Asymmetries, and Investors' Reliance on Earnings." *Working paper, Georgetown University*.

Armour, John, Colin Mayer, and Andrea Polo. 2017. "Regulatory Sanctions and Reputational Damage in Financial Markets." *Journal of Financial and Quantitative Analysis* 52(4):1429-1448.

Arnold, Monique, and Peter Jan Engelen. 2007. "Do Financial Markets Discipline Firms for Illegal Corporate Behaviour." *Management & Marketing* 2(4):103-110.

Baker, H. Kent, Richard B. Edelman, and Gary E. Powell. 1999. "The Effect of Announcements of Corporate Misconduct and Insider Trading on Shareholder Returns." *Business and Professional Ethics Journal* 18(1):47-64.

Barabanov, Sergey S., Onem Ozocak., H.J. Turtle, and Thomas J. Walker. 2008. "Institutional Investors and Shareholder Litigation." *Financial Management* 37(2):227-250.

Bardos, Katsiaryna Salavei, Joseph Golec, and John P. Harding. 2013. "Litigation Risk and Market Reaction to Restatements." *Journal of Financial Research* 36(1):19-42.

Bardos, Katsiaryna Salavei, and Dev Mishra. 2014. "Financial Restatements, Litigation and Implied Cost of Equity." *Applied Financial Economics* 24(1):51-71.

Barniv, Ran R., and Jian Cao. 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, Rob, and Robin Braun. 2010. "Misdeeds Matter: Long-Term Stock Price Performance After the Filing of Class-Action Lawsuits." *Financial Analysts Journal* 66(6):74-92.

Beasley, Mark S., Dana R. Hermanson, Joseph V. Carcello, and Terry L. Neal. 2010. "Fraudulent Financial Reporting: 1998-2007: An Analysis of U.S. Public Companies." *COSO, Committee of Sponsoring Organizations of the Treadway Commission*.

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

Bhagat, Sanjai, John Bizjak, and Jeffrey L. Coles. 1998. "The Shareholder Wealth Implications of Corporate Lawsuits." *Financial Management* 27:5-27.

Billings, Mary Brook, April Klein, and Emanuel Zur. 2012. "Shareholder Class Action Suits and the Bond Market." *Working paper* available at SSRN 1984666.

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

Bonini, Stefano, and Diana Boraschi. 2010. "Corporate Scandals and Capital Structure." *Journal of Business Ethics* 95:241-269.

Bowen, Robert M., Andrew C. Call, and Shiva Rajgopal. 2010. "Whistle-Blowing: Target Firm Characteristics and Economic Consequences." *The Accounting Review* 85(4):1239-1271.

Bradley, Daniel, Brandon N. Cline, and Qin Lian. 2014. "Class Action Lawsuits and Executive Stock Option Exercise." *Journal of Corporate Finance* 27:157-172.

Brous, Peter A., and Keith Leggett. 1996. "Wealth Effects of Enforcement Actions Against Financially Distressed Banks." *Journal of Financial Research* 19(4):561-577.

Burns, Natasha, and Simi Kedia. 2006. "The Impact of Performance-Based Compensation on Misreporting." *Journal of Financial Economics* 79(1):35-67.

Callen, Jeffrey L., Joshua Livnat, and Dan Segal. 2006. "Accounting Restatements: Are They Always Bad News for Investors?" *The Journal of Investing* 15(3):57-68.

Chava, Sudheer, C.S. Agnes Cheng, Henry Huang, and Gerald J. Lobo. 2010. "Implications of Securities Class Actions for Cost of Equity Capital." *International Journal of Law and Management* 52(2):144-161.

Chen, Gongmeng, Michael Firth, Daniel N. Gao, and Oliver M. Rui. 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, Hae Mi, Jonathan M. Karpoff, Xiaoxia Lou, and Gerald S. Martin. 2019. "Enforcement Waves and Spillovers." *Working paper*, available at SSRN 3526555.

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

Christensen, Theodore E., Daniel Gyung H. Paik, and Christopher D. Williams. 2010. "Market Efficiency and Investor Reactions to SEC Fraud Investigations." *Journal of Forensic & Investigative Accounting* 2(3):1-30.

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

Correia, Maria, and Michael Klausner. 2012. "Are Securities Class Actions "Supplemental" to SEC Enforcement? An Empirical Analysis." *Working Paper Stanford Law School Stanford.*

Cox, Raymond A., and Thomas R. Weirich. 2002. „The Stock Market Reaction to Fraudulent Financial Reporting." *Managerial Auditing Journal* 17(7):374-382.

Davidson, Wallace N., Dan L. Worrell, and Chun I. Lee. 1994. “Stock Market Reactions to Announced Corporate Illegalities.” *Journal of Business Ethics* 13(12):979-987.

Davis, Frederik, Behzad Taghipour, and Thomas J. Walker. 2017. “Insider Trading Surrounding Securities Class Action Litigation and Settlement Announcements.” *Managerial Finance* 43(1):124-140.

de Batz, Laure. 2020. “Financial Impact of Regulatory Sanctions on Listed Companies.” *European Journal of Law and Economics* 49(2):301-337.

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 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, Saiying, Richard H. Willis, and Li Xu. 2014. “Shareholder Litigation, Reputational Loss, and Bank Loan Contracting.” *Journal of Financial and Quantitative Analysis* 49(4):1101-1132.

Desai, Hemang, Chris E. Hogan, and Michael S. Wilkins. 2006. “The Reputational Penalty for Aggressive Accounting: Earnings Restatements and Management Turnover.” *The Accounting Review* 81(1):83-112.

Djama, Constant. 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, Lijing. 2017. “The CDS Market Reaction to Restatement Announcements.” *Journal of Business Finance & Accounting* 44(7-8):1015-1035.

Engelen, Peter Jan. 2009. “The Reputational Penalty for Illegal Insider Trading by Managers.” *Working paper presented at the Academy of Management Conference, Chicago U.S.*

Engelen, Peter Jan. 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, Peter Jan. 2012. “What is the Reputational Cost of a Dishonest CEO? Evidence from U.S. Illegal Insider Trading.” *CESifo Economic Studies* 58:140-163.

Eryiğit, Mehmet. 2019. “Short-Term Performance of Stocks after Fraudulent Financial Reporting Announcement.” *Journal of Financial Crime* 26(2):464-476.

Ewelt-Knauer, Corinna, Thorsten Knauer, and Maik Lachmann. 2015. “Fraud Characteristics and Their Effects on Shareholder Wealth.” *Journal of Business Economics* 85(9):1011-1047.

Feroz, Ehsan H., Kyungjoo Park, and Victor S. Pastena. 1991. “The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases.” *Journal of Accounting Research* 29:107-142.

Ferris, Stephen P., Tomas Jandik, Robert M. Lawless, and Anil Makhija. 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, Eliezer M., and Anil Shivdasani. 2007. “Financial Fraud, Director Reputation, and Shareholder Wealth.” *Journal of Financial Economics* 86(2):306-336.

Firth, Michael, Oliver M. Rui, and Wenfeng Wu. 2011. “Cooking the Books: Recipes and Costs of Falsified Financial Statements in China.” *Journal of Corporate Finance* 17(2):371-390.

Firth, Michael, Oliver M. Rui, and Xi Wu. 2009. “The Tand Consequences of Disseminating Public Information by Regulators.” *Journal of Accounting and Public Policy* 28(2):118-132.

Firth, Michael, Sonia Wong, Qingquan Xin, and Ho Yin Yick. 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, Christian, Hans Degryse, Sascha Kolaric, and Dirk Schiereck. 2018. “Forgive Me All My Sins: How Penalties Imposed on Banks Travel Through Markets.” *Working Paper, available at SSRN 3178589*.

Gande, Amar, and Craig M. Lewis. 2009. “Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers.” *Journal of Financial and Quantitative Analysis* 44(4):823-850.

Gerety, Mason, and Kenneth Lehn. 1997. “The Causes and Consequences of Accounting Fraud.” *Managerial and Decision Economics* 18(7-8):587-599.

Goldman, Eitan, Urs Peyer, and Irina Stefanescu. 2012. “Financial Misrepresentation and its Impact on Rivals.” *Financial Management* 41(4):915-945.

Griffin, Paul A., Joseph A. Grundfest, and Michael A. Perino. 2004. “Stock Price Response to News of Securities Fraud Litigation: An Analysis of Sequential and Conditional Information.” *Abacus* 40(1):21-48.

Griffin, Paul A., and Yuan Sun. 2011. “Troublesome Tidings? Investors’ Response to a Wells Notice.” *Accounting and Finance Research* 5(1):99-120.

Haslem, Bruce, Irena Hutton, and Aimee Hoffman Smith. 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, Mark, Zoe-Vonna Palmrose, and Susan Scholz. 2005. “Long-Term Market Underreaction to Accounting Restatements.” *Working Paper, University of Kansas, School of Business*.

Humphery-Jenner, Mark L. 2012. “Internal and External Discipline Following Securities Class Actions.” *Journal of Financial Intermediation* 21(1):151-179.

Iqbal, Zahid, Shekar Shetty, and Kun Wang. 2007. “Further Evidence on Insider Trading and the Merits of Securities Class Actions.” *Journal of Financial Research* 30(4):533-545.

Johnson, Shane A., Harley E. Ryan, Jr., and Yisong S. Tian. 2003. “Executive Compensation and Corporate Fraud.” *Working Paper, Lousiana State University, Baton Rouge, LA*.

Jordan, John S., Joe Peek, and Eric S. Rosengren. 2000. “The Market Reaction to the Disclosure of Supervisory Actions: Implications for Bank Transparency.” *Journal of Financial Intermediation* 9(3):298-319.

Karpoff, Jonathan M., Allison Koester, D. Scott Lee, and Gerald S. Martin. 2017. “Proxies and Databases in Financial Misconduct Research.” *The Accounting Review* 92(6):129-163.

Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin. 2008. “The Cost to Firms of Cooking the Books.” *Journal of Financial and Quantitative Analysis* 43(3):581-611.

Karpoff, Jonathan M., and John R. Lott, Jr. 1993. “The Reputational Penalty Firms Bear from Committing Criminal Fraud.” *The Journal of Law and Economics* 36(2):757-802.

Kellogg, Robert L. 1984. “Accounting Activities, Security Prices, and Class Action Lawsuits.” *Journal of Accounting and Economics* 6(3):185-204.

Kirat, Thierry, and Amir Rezaee. 2019. “How Stock Markets React to Regulatory Sanctions? Evidence from France.” *Applied Economics* 51(60):6558-6566.

Klock, Mark. 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, Roy, and Visit Phunnarungsi. 2013. “Corporate Governance, Violations and Market Reactions.” *Pacific-Basin Finance Journal* 21(1):881-898.

Kravet, Todd, and Terry J. Shevlin. 2010. “Accounting Restatements and Information Risk.” *Review of Accounting Studies* 15(2):264-294.

Kryzanowski, Lawrence, and Ying Zhang. 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, Jing Hui, and Shiang Shen Kwan. 2011. “Violation of Listing Requirements and Company Value: Evidence from Bursa Malaysia.” *International Review of Business Research Papers* 7(2):257-268.

Lei, Adrian C. H., and Philip K. F. Law. 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, Benjamin L., and Curtis J. Milhaupt. 2008.“ Reputational Sanctions in China's Securities Market.” *Business Finance & Accounting, Columbia Law Review* 108(4):929-983.

Lieser, Patrick, and Sascha Kolaric. 2016. “Securities Class Action Litigation, Defendant Stock Price Revaluation, and Industry Spillover Effects.” *Working paper, European Financial Management Association Conference paper n° 0388*.

Loh, Charmen, and R. S. Rathinasamy. 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, Dalia, Samuel H. Szewczyk, Hatice Uzun, and Raj Varma. 2006. “Governance and Performance Changes after Accusations of Corporate Fraud.” *Financial Analysts Journal* 62(3):32-41.

Marciukaityte, Dalia, Samuel H. Szewczyk, and Raj Varma. 2009. “Voluntary vs. Forced Financial Restatements: The Role of Board Independence.” *Financial Analysts Journal* 65(5):51-65.

McDowell, John. 2005. “A Look at the Market’s Reaction to the Announcements of SEC Investigations.” *Working paper, The Leonard N. Stern School of Business.*

Muradoglu, Gulnur, and Jennifer Clark Huskey. 2008. “The Impact of SEC Litigation on Firm Value.” *Working Paper*.

Nainar, S. M. Khalid, Atul Rai, and Semih Tartaroglu. 2014. “Market Reactions to Wells Notice: An Empirical Analysis.” *International Journal of Disclosure and Governance* 11(2):177-193.

Nelson, Christine, Sara Gilley, and Garett Trombley Esq. 2009. “Disclosures of SEC Investigations Resulting in Wells Notices.” *Securities Litigation Journal* 19(4):19-21.

Nourayi, Mahmoud M. 1994. “Stock Price Responses to the SEC's Enforcement Actions.” *Journal of Accounting and Public Policy* 13(4):333-347.

Owers, James E., Chen-Miao Lin, and Ronald C. Rogers. 2002. “The Information Content and Valuation Ramifications of Earnings Restatements.” *International Business and Economics Research Journal* 1(5):71-83.

Ozbas, Oguzhan. 2008. “Corporate Fraud and Real Investment.” *Working paper*.

Ozeki, Norimasa. 2019. “Determinants of Market Reaction to Disclosure of Accounting Misconduct: Evidence from Japan.” *Securities Analysts Journal* 57(3):72-84.

Pereira, John, Irma Malafronte, Ghulam Sorwar, and Mohamed Nurullah. 2019. „Enforcement Actions, Market Movement and Depositors’ Reaction: Evidence from the U.S. Banking System.” *Journal of Financial Services Research* 55(2-3):143-165.

Persons, Obeua S. 1997. “SEC's Insider Trading Enforcements and Target Firms' Stock Values.” *Journal of Business Research* 39(3):187-194.

Plumlee, Marlene, and Teri Lombardi Yohn. 2008. “Restatements: Investors Response and Firm Reporting Choices.” *Working Paper*.

Pritchard, A. C., and Stephen P. Ferris. 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, Roberta. 1991. “The Shareholder Suit: Litigation Without Foundation?” *Journal of Law, Economics, and Organization* 7:55-87.

Scholz, Susan. 2008. “The Changing Nature and Consequences of Public Company Financial Restatements.” The United States Department of the Treasury, 1-54.

Slovin, Myron B., Marie E. Sushka, and John A. Polonchek. 1999. “An Analysis of Contagion and Competitive Effects at Commercial Banks.” *Journal of Financial Economics* 54:197-225.

Song, Chanhoo, and Seung Hun Han. 2017. “Stock Market Reaction to Corporate Crime: Evidence from South Korea.” *Journal of Business Ethics* 143(2):323-351.

Sun, Peng, and Yi Zhang. 2006. “Is There Penalty for Crime: Corporate Scandal and Management Turnover in China?” *EFA 2006 Zurich Meetings Paper*.

Takmaz, Sefa, and Emel Mihriban Keles. 2017. “Do Stock Prices React to Illegal Corporate Behaviors? The Turkish Case.” *Journal of Business Research Turk* 9(4):245-258.

Tanimura, Joseph K., and M. Gary Okamoto. 2013. “Reputational Penalties in Japan: Evidence from Corporate Scandals.” *Asian Economic Journal* 27(1):39-57.

Tay, Liang-Mui, Chin-Hong Puah, Rayenda Khresna Brahmana, and Nurul Izza Abdul Malek. 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, Yang, John K. Ashton, and Aziz Jaafar. 2019. “Money Shouts! How Effective are Punishments for Accounting Fraud?” *The British Accounting Review* 51(5):100824.

Wang, Xia, and Min Wu. 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, Min. 2002. “Earnings Restatements: A Capital Market Perspective.” *Working paper, New York University, New York, NY*.

Wu, Xi, and Junsheng Zhang. 2014. “Stock Market Reaction to Regulatory Investigation Announcements.” *China Journal of Accounting Studies* 2(1):37-52.

Xu, Wenming, and Guangdong Xu. 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, Xin, Peng Zhang, and Ying Zheng. 2015. “Corporate Governance, Political Connections, and Intra-Industry Effects: Evidence from Corporate Scandals in China.” *Financial Management* 44(1):49-80.

Zeidan, Mohamad Jamal. 2013. “Effects of Illegal Behavior on the Financial Performance of U.S. Banking Institutions.” *Journal of Business Ethics* 112:313-324.

Zhu, Zhaohui, and Chengwei Hu. 2010. “Market Reactions to Financial Restatements-Evidence from Chinese Stock Market.” *2010 IEEE International Conference on Industrial Engineering and Engineering Management,* 2527-2530.