PREVENTION IS BETTER THAN CURE: FORECASTING FUTURE MISREPORTING

ABSTRACT
Our study introduces a unique model that forecasts future financial misreporting. By monitoring evolving patterns of human intervention in financial statements in real-time and applying Benford's Law in an innovative manner, we identify early signs of the slippery slope—a precursor to misreporting. This approach allows us to pinpoint firms at a heightened risk of misreporting in the future. Using a hold-out sample our model achieves an overall accuracy of 73.24 percent, correctly forecasting 90.58 percent, 83.61 percent, and 74.14 percent of firms one, two and three or more years respectively, in advance of the misreporting occurring. This pioneering model serves as an effective early warning system and offers a unique tool for auditors and boards of directors to proactively intervene before misreporting occurs, marking a significant improvement over existing models that primarily focus on post-misreporting interventions.

I. INTRODUCTION

"Huge ethical lapses often begin with small steps" (Pulliam 2005)

The primary objective of this paper is to provide stakeholders with an effective tool to proactively prevent financial misreporting, or at the very least, mitigate its costs should it occur. Misreporting refers to any significant accounting misconduct, fraud, misstatement or financial misrepresentation outside of GAAP.[1] Despite rigorous enforcement of securities regulations, such as the Sarbanes-Oxley Act (SOX) and the Dodd-Frank Act, alongside the use of severe financial penalties, financial misreporting remains a pervasive issue. This persistent problem raises concerns about the effectiveness of both the current audit practices and regulatory frameworks (SEC 2022; IAASB 2022; PCAOB 2022).

In contrast to existing misreporting prediction models which predominantly react to incidents of misreporting *after* they occur, our paper introduces a novel forecasting model that *forecasts* the likelihood of future misreporting *before* it happens. This model aims to provide decision-makers with anticipatory insights, much like bankruptcy forecasting models, enabling

[1]This definition aligns with the existing literature and includes occurrences of Accounting and Auditing Enforcement Releases (AAERs), financial restatements, and class action lawsuits. Each metric carries distinct advantages and limitations as detailed by Ham, Lang, Seybert, and Wang (2017), Karpoff, Koester, Lee, and Martin (2017), and Malenko, Grundfest, and Shen (2022). For a comprehensive description of our methods for identifying misreporting, see Section III.

them to make informed early interventions. By shifting from a reactive to a preventative strategy, this model aims to transform how stakeholders manage the risk of financial misreporting by addressing the gap in current methodologies and highlighting the need for a more forward-thinking approach.

Traditional misreporting models strive to retrospectively identify the specific year(s) misreporting took place before it is confirmed by regulators (e.g., Beneish 1999; Dechow, Ge, Larson, and Sloan 2011; Chakrabarty, Moulton, Pugachev, and Wang 2024), while our model forecasts the likelihood of misreporting *before* it occurs. Figure 1 illustrates the differing objectives between our model and the traditional *F-Score* model (Dechow et al 2011). Assume a firm misreports in time *t*; the *F-Score*, derived using financial variables observed in time *t*, would produce a high *F-Score* indicating a greater likelihood the accounts in time *t* are misreported.

In contrast, our model's objective is to forecast the misreporting event *t* before it occurs. For instance, our model could forecast in *t-3* that the firm is at a high risk of misreporting in the future, based on information available up to and including *t-3*. This alert would expect to be reiterated in every subsequent year we run the model (*t-2; t-1*), projecting the likelihood of potential misreporting in advance of *t*. A practical example of this is HP Inc., which overstated revenues and closing inventory from 2015Q1 to 2016Q4. While the *F-Score*, applied to their 2015 and 2016 financial statements, successfully identified those years as containing misreporting activity, our model had already signaled a significant risk of future misreporting in the fourth quarter of 2010. This warning persisted as the misreporting period approached, highlighting our model's ability to provide crucial early and ongoing risk assessments. Predicting the exact year of misreporting is crucial for corrective actions, but it does not offer an opportunity to prevent the initial misreporting. The application of our model, depicted as the 'prevention period' in Figure 1, allows stakeholders to intervene early, thereby directly reducing the costs associated with financial misreporting. In the case of HP Inc., our model effectively provides a four-year window for preventative measures to be taken against future misreporting.

Our model exploits the slippery slope behavior which has been identified in the prior literature as a precursor to misreporting. Behavior of the slippery slope encompasses non-egregious actions which are characterized by gradual and increasingly aggressive within-GAAP accounting interventions aiming to meet financial benchmarks (Dechow et al. 2011; Amiram, Bozanic, and Rouen 2015; Chu, Dechow, Hui, and Wang 2019). Such interventions eventually lead to misreporting with the adoption of questionable and egregious practices outside standard accounting guidelines (Schrand and Zechman 2012; Brown 2014; Chen, Lee, and Zapatero 2022).

To date the slippery slope has only been identified forensically by looking backwards following the confirmation of the misreporting event (Dechow et al. 2011; Amiram et al. 2015; Chu et al. 2019). For instance, Amiram et al. (2015) observed a slippery slope but only by using perfect hindsight i.e. that the firm had misreported and the date it took place. This retrospective approach however does not offer a mechanism to observe the slippery slope as it happens.

We use Benford Law (BL) to measure the slippery slope as it occurs and directly track the average escalating trend in human intervention in the accounting process. We do so by continuously monitoring any abnormal increases in a firm's BL deviations as they occur based on the numerical data in their quarterly accounts. BL is particularly advantageous as it is independent of underlying firm characteristics, business model, and economic performance (Amiram et al. 2015; Chakrabarty et al. 2024). According to BL, the first digits of naturally occurring data, such as financial statements, should exhibit a logarithmic distribution, for instance the digit '1' should appear as the first digit about 30.1 percent of the time, and '9' appear about 4.6 percent of the time. However, human intervention will affect the data and the distribution of these first digits will deviate from expected frequencies, regardless of the nature of the intention of such interventions. Prior research appears to assume that these motivations purely relate to an intention to misreport. Consequently, any observed deviations from BL have been leveraged as an explicit tool to pinpoint the specific year(s) of financial misreporting (Amiram et al. 2015; Chakrabarty et al. 2024). In contrast, our approach monitors deviations from BL over time, not aiming to pinpoint the year of misreporting but to identify the escalating human intervention in the accounts prior to the misreporting year.

We equate any observed abnormal increase in a firm's BL deviations to an increase in human intervention in that specific quarter. If there are continuous increases in human intervention quarter-on-quarter then we posit that the pattern reflects the characteristics of a slippery slope. We continuously update and calculate misreporting hazard scores for each firm by implementing our model annually with our slippery slope as the main predictor variable using quarterly data. These scores predict the probability of an imminent misreporting event. A red flag is assigned if the hazard score indicates a high likelihood that the firm will misreport in the future.

Our sample includes publicly traded US companies from 1961 to 2020, utilizing data from COMPUSTAT and various databases to identify misreporting firms (Karpoff, Koester, Lee and Martin 2017; Donelson, Kartapanis, McInnis, and Yust 2021). We validate our human intervention indicator by examining its correlation with accounting judgments and assumptions, finding significant positive correlations with accrual measures and motivations linked to achieving financial benchmarks (Chu et al. 2019).

Our forecasting model offers strong accuracy in predicting financial misreporting and shows effective capabilities in practical applications. For our hold-out sample, the model effectively red flags misreporting firms one year in advance of the misreporting event with a

90.58 percent accuracy and maintains strong performance over longer forecast horizons—83.61 percent two years ahead and 74.14 percent three or more years in advance. Once the model flags a firm as being at a high-risk of misreporting in the future this warning remains consistent and continues until the misreporting occurs, reinforcing the model's reliability. Moreover, it also successfully identifies firms that will not misreport with 69.99 percent accuracy. Consequently, the model's overall accuracy stands at 73.24 percent with a precision of 36.29 percent.[2] For our in-sample tests, the model achieves a notable accuracy rate of 87.68 percent with a precision of 60.79 percent, forecasting 99.86 percent of future misreporting firms and correctly identifying 84.81 percent of non-misreporting firms. These metrics underscore our model's advanced analytical capabilities and its substantial potential to significantly enhance stakeholders' ability to pre-empt potential future financial misreporting. We also explored alternative model specifications with various earnings management measures from previous research to measure the slippery slope. However, these measures lowered the model's accuracy by adding noise resulting in significantly higher false positives.

In summary our paper introduces an innovative model to forecast future misreporting and we believe it represents a pioneering contribution to both academic literature and practical application. This model adopts a forward-looking approach which is crucial for timely decision-making by uniquely capturing the slippery slope of misreporting as it unfolds. This contrasts with previous studies that have only retrospectively identified such patterns *after* the occurrence of misreporting events (Dechow et al. 2011; Amiram et al. 2015; Chu et al. 2019).

The practical implications of our model are extensive, spanning corporate governance, investor relations, auditing and academic research. For instance, it empowers corporate boards to recognize early signs of escalating human intervention and enables them to take proactive

[2] We employ Beneish et al. (2020) definitions whereby accuracy is the ratio of true positive and negatives to the sum of all firms (the number of correctly identified misreporting and no misreporting firms to the sum of all observations), while precision is the ratio of true positives to the sum of true and false positives (the number of misreporting firms flagged divided by all flagged observations).

measures to mitigate risks. Auditors can leverage this model to enhance their scrutiny and improve audit quality by pre-emptively identifying a firm's potential for misreporting. Investors and lenders, meanwhile, obtain valuable insights that can inform their decisions and potentially steer them away from high-risk entities. Our model significantly enriches the academic understanding of the early stages of misreporting by providing researchers with tools to investigate the initial triggers and subsequent responses to the slippery slope of financial misreporting. Furthermore, it introduces a novel metric specifically designed to analyze earnings management through the lens of human intervention, independent of firm fundamentals. This development addresses a critical gap in existing methodologies, offering a robust tool previously unattainable, which remains unaffected by underlying business metrics. Consequently, this innovation not only facilitates a potential for deeper exploration of earnings management and its complex dynamics, but also enables a potential reassessment of the existing financial misreporting literature.

The remainder of the paper is organized as follows: Section II describes the motivation and prior literature. Section III details the research design and sample, with Section IV presenting our main result and additional analysis, while Section V concludes.

II. MOTIVATION AND PRIOR LITERATURE

The ability to prevent financial statement misreporting remains limited despite the enactment of stringent securities regulations. Misstatements remain prevalent with the SEC penalties of $27.054 billion from 2017 to 2022 attesting to this (SEC, 2022). This ongoing issue not only underscores the regulatory challenges in curbing misreporting but also highlights the apparent shortcomings of current audit practices (Brazel, Jones, Thaye and Warne 2015; SEC 2022; IAASB 2022; PCAOB 2022). This issue is compounded by managerial attitudes that either underestimate the likelihood of detection (Dichev, Graham, Harvey, and Rajgopal 2013) or assume the benefits of misreporting outweigh the consequences, which is likely since approximately 33 percent of perpetrators gain a net benefit of $6.85 million even if caught (Amiram, Huang, and Rajagopal 2020). Such misconduct imposes heavy burdens on stakeholders and the economy with recent estimations suggesting that two-thirds of corporate misreporting remains uncovered (Dyck, Morse, and Zingales 2024; Beneish, Farber, Glendening, and Shaw 2023). It is therefore crucial that stakeholders be equipped with tools that can clearly signal a firm's risk of future misreporting. This foresight allows for timely interventions that can prevent misreporting or minimize its potential benefits, thus redirecting managerial decisions from focusing on net benefits to ones that recognize net costs.

To date, existing misreporting prediction models such as those by Beneish (1999), Cecchini, Aytug, Koehler, and Pathak (2010), Dechow et al. (2011), and Chakrabarty et al. (2024), while effective at detecting misreporting when it occurs, lack the foresight needed for forecasting future misreporting activities. For instance, the *F-Score* and its subsequent variations assess whether current financial statements are manipulated, but they do not, and were never intended to forecast whether the firms will be misreporting in future periods (Dechow et al. 2011; Chakrabarty et al. 2024).

The Antecedent to Misreporting: The Slippery Slope

Prior research has identified that material accounting misreporting and other forms of corporate misconduct typically follow a slippery slope rather than a sudden 'cliff edge process (Schrand and Zechman, 2012; Brown, 2014). This incremental process involves minor manipulations initially permissible under GAAP, which gradually necessitate more aggressive earnings management practices to meet expectations. Over time, these adjustments escalate, leading to significant compliance violations and potentially severe misreporting infractions (Chu et al. 2019; Chen et al. 2022). Managers face escalating pressures to sustain earnings momentum and surpass analysts' forecasts, motivating them to employ progressively more aggressive accounting strategies that may ultimately lead to misreporting and sanctions (Chu et al. 2019; Dechow, Sloan, and Sweeney 1996; Richardson, Tuna, and Wu 2002; Chen et al. 2022).

Historically, the misreporting literature has examined the slippery slope phenomenon with the advantage of hindsight. For instance, higher reported accruals have been observed in the years leading to the manipulation period (Beneish, 1997; Dechow et al. 2011; Chu et al. 2019), and strategic rounding up of EPS up to five years prior to misreporting has been noted (Malenko, Grundfest, and Shen, 2022). The *F-Score* for misreporting firms increases for up to three years prior to the misstatement period, often coinciding with a high frequency of beating analyst forecasts (Chu et al. 2019). Additionally, Amiram et al. (2015) found a positive association between prior year deviations from BL and the misreporting year. These findings suggest that increasingly aggressive accounting behavior is strongly associated with future misconduct, spanning several years and reducing the likelihood of early detection by gatekeepers. Gino and Bazerman (2009) argue that unethical behavior evolving gradually is more likely to be accepted by observers, such as auditors, who may not detect slow escalation in earnings management, thus allowing the slippery slope to go undetected and unprevented.

Building on these insights, if the slippery slope is indeed a common precursor to misreporting, then detecting it as it occurs could enable the anticipation and forecasting of future misreporting risks. However, until now, ex-ante detection of the slippery slope has not been attempted. Although Amiram et al. (2015) established an association between deviations from BL in periods *t-2* and *t-1* and the occurrence of misreporting in period *t*, this analysis was inherently retrospective, conditioned on the presence of misreporting and its timing. Likewise, the upward trend in the *F-Score* is only apparent in hindsight. While such insights are invaluable for understanding the pervasive nature of the slippery slope among firms that misreport, they do not aid in measuring and forecasting the slippery slope as it unfolds.

Using Benford Law to Measure the Slippery Slope.

To capture the slippery slope as it occurs, we refine the prior use of BL by emphasizing that deviations from BL capture all human interventions in financial statements, not solely misreporting ones. BL describes the statistical frequency distribution of leading digits across naturally occurring datasets.[3] However, this distribution is often disrupted by human intervention, causing deviations from the expected logarithmic pattern observed in such datasets (Hsü 1948, Kubovy 1977, Hill 1998).

In financial reporting, deviations from BL indicate various levels of human intervention, ranging from benign adjustments to overt manipulations.[4] These deviations encompass a broad spectrum of human activities impacting the data, thus extending beyond mere instances of misreporting (Nigrini 1996; Amiram et al. 2015; Chakrabarty et al. 2024). Prior literature, which has largely focused on identifying specific years of misreporting, may not have fully acknowledged this broader interpretation.[5] This oversight could explain why models based on BL deviations exhibit high levels of false positives, resulting from the misclassification of normal human interventions as misreporting (Beneish and Vorst 2022).[6]

Our refined approach utilizes the temporal dynamics of deviations from BL to detect patterns indicative of the slippery slope in financial reporting. By examining the progression of these deviations from one period to the next, we aim to identify escalating human

[3] $P_d = log_b(d+1) - log_b(d) = log_b\left(\frac{d+1}{d}\right)$. Where P is the probability of the occurrence of first digit d, and b is the logarithmic base. The expected probability of occurrence for leading digits 1 through 9, results in the theoretical distribution which today is referred to as BL.

[4]We refer the readers to Amiram et al. (2015) Appendix 1 to 5 which provides several resources that help strengthen the intuition regarding 'why' and 'how' BL can be used to detect human intervention in accounting numbers.

[5]For example, Amiram et al. (2015) subjected financial data to various types of human intervention in their simulation to test the sensitivity and specificity of BL-based tests. Their simulation assumed interventions were manipulative, such as revenues or understating inventory. However, if these changes were for legitimate reasons, like adjustments in revenue recognition policy or inventory provisions, the observed deviation from BL would remain unchanged.

[6]Beneish and Vorst (2022) find utilizing Amiram et al. (2015)'s BL deviation score – *FSD score* – to pinpoint the year of misreporting results in a false positive rate of 42.83 percent and utilizing Chakrabarty et al. (2024) *ABF-score* (a combination of BL deviation and the *F-Score*) results in a false positive rate of 58.77 percent which renders these models as not very useful in ascertaining whether misreporting took place or not.

interventions within financial statements. Importantly, our focus is not on whether these deviations indicate compliance or non-compliance with GAAP; instead, we concentrate on discerning any escalating patterns that could suggest future manipulative behavior.

All firms operating under GAAP inherently exercise a degree of discretion and judgment, which involves various levels of human intervention. Typically, these interventions should display stability over time unless there is an increasing pattern of manipulation. For example, a significant accounting policy adjustment may initially cause a deviation from BL, but if the change is legitimate, this deviation is likely to stabilize in subsequent periods, indicating no ongoing escalation.

In summary, the strength of BL's utility in our model lies in its ability to track human interventions dynamically, without any prior bias linked to the firm's inherent characteristics or economic performance, offering a clearer insight into managerial behaviors over time (Amiram et al. 2015; Chakrabarty et al. 2024). This approach contrasts with other empirical measures like accruals or earnings benchmarks, which are significantly influenced by external economic factors and may not reliably indicate the presence of the slippery slope (Owens, Wu, and Zimmerman 2017; Chu et al. 2019).

III. RESEARCH DESIGN

Identifying Escalating Human Intervention

We employ deviations from BL to identify human interventions in financial statements, akin to the approach used in Amiram et al. (2015) with their Financial Statement Divergence score (*FSD*). However, we introduce two modifications to mitigate biases associated with the size of the digit pool. Although *FSD* is scale invariant, as it is unaffected by measurement units (Pinkham 1961), it is susceptible to variations in the size of the digit pool—the aggregate number of digits in the financial statements—which can affect the sensitivity and visibility of the expected BL distribution, particularly when the digit pool is small (Barney and Schulzke 2016).

To address this 'small digit bias', we apply Johnson and Weggnmann (2013)'s modified version of *FSD* (*Mod_FSD*), which reduces variability induced by a smaller number of digits (Horton, Krishna Kumar, and Wood 2020). *Mod_FSD* is calculated as follows:

$$Mod\_FSD_{it} = \frac{\sum_{j=1}^{9} |(X_{jit})\text{-}FSD_{it}|}{9} \tag{1}$$

where X_{jit} equals the absolute value of the difference between the actual observed frequency of leading digit j (AF_{jit}) and its expected frequency (EF_j) under BL for firm i in quarter t.[7] Furthermore, we address a mechanistic bias where changes in the size of the digit pool can inadvertently affect the *FSD*; as the digit pool size increases, the *FSD* typically decreases, and vice versa, irrespective of any actual human intervention (Barney and Schulzke 2016; Chakrabarty et al. 2024). To correct for this bias, particularly given our focus on BL deviation trends over time, we adjust the *Mod_FSD* by multiplying it by the pool of digits in the financial statements for that firm-quarter:

$$Adj\_FSD_{it} = Mod\_FSD_{it} \times N_{it} \tag{2}$$

where N_{it} is the aggregate number of items in the financial statement (i.e., balance sheet, cash flow, and income statement) of firm i during quarter t.[8]

To determine whether a firm exhibits on average an escalating trend in human intervention in quarter t we measure the abnormal *Adj_FSD* (*Abn_FSD*). Specifically, we determine the difference between firm i's *Adj_FSD* score in quarter t and its average of all

[7] *FSD* is calculated as $\frac{\sum_{j=1}^{9} |AF_{jit}\text{-}EF_{jit}|}{9}$ Where AF_j is the actual observed frequency of digit j in the dataset, and EF_j is the expected frequency of digit j under BL.

[8] We conducted a Monte Carlo simulation to examine *FSD* vs. *Mod_FSD* vs. *Adj_FSD* properties for simulated numbers that follow BL distribution with different pools of digit sizes. Overall, we find that our *Adj_FSD* is less mechanically biased and more precise than *FSD* and that *Mod_FSD* significantly removes the small pool of digits bias.

available preceding quarters *Avg_Adj_FSD* up to quarter $t-1$.[9] If the financial statement's deviation from BL in quarter t is larger than the average of all the prior quarters, so *Abn_FSD* > 0, this indicates an abnormal increasing level of human intervention in quarter t. We therefore create an indicator variable for increasing human intervention, HI_{it}, which takes value one if $Abn\_FSD_{it} > 0$, and zero otherwise. Moreover, since the slippery slope is not merely isolated to single instances of increasing levels of human intervention (i.e., when $HI_{it} = 1$), but rather the sustained sequences of such incidents, we also use the number of consecutive $HI_{it} = 1$'s to capture the uninterrupted sequences of increasing human intervention. We define this measure as $String_{it}$ which serves as a measure of the slippery slope period for firm i.[10] All variables are defined in Appendix 1.

Forecasting Misreporting Model

We develop a Cox proportional hazard model consistent with the forecast bankruptcy literature (Shumway 2001).[11] The hazard model is advantageous because it allows the incorporation of both time-dependent and time-independent variables into the analysis, thereby providing a comprehensive understanding of the risks associated with misreporting. Unlike a logistic regression model which gives a flat probability, the hazard model adjusts for the fact that the probability may change over time (Shumway 2001), depending on how long the firm has been at risk of future misreporting. It therefore analyses the time intervals preceding the misreporting event, as opposed to the misreporting period itself and thus its application is well-suited to forecasting future misreporting. The model determines a firm's specific hazard score which signifies the instantaneous probability of an event occurring (in our case misreporting) at a

---

[9]Our results are robust to using the past 5 or 10 years to calculate *Avg_Adj_FSD* thus we do not suffer potential biases introduced by the mean calculation over a long time series.
[10]To illustrate, consider a firm that had increased its human intervention in the financial statements for the past four quarters up to $t-1$ as reflected by the increasing BL deviations, then we have $String_{it-1} = 4$. If a firm also increased human intervention in quarter t then $HI_{it} = 1$, otherwise $HI_{it} = 0$. In the case when $HI_{it} = 1$, then we will have $String_{it} = 5$. However, if a firm does not increase its human intervention in quarter t and $HI_{it} = 0$, the string of consecutive increases in human intervention is broken and hence $String_{it} = 0$.
[11] Using a discrete-time Cox model we obtain similar forecasting accuracy.

particular time *t*, given the event has not yet occurred, conditional on the covariates (in our case the slippery slope variable). The 'proportional' aspect of the model implies that the ratio of the hazard event for two different firms remains constant over time, thereby allowing for the comparison of hazard scores across firms.

We design our model where the hazard event takes the value of one if a firm misreports in a particular quarter for the first time and zero otherwise. Since our primary focus is forecasting contemporaneously, we exclude any information not available at the time when we determine the hazard score i.e., time *t*.[12] A firm contributes an observation in the estimation equation for every quarter available in COMPUSTAT until the first misreporting quarter. We estimate the model employing several predictive variables which capture the slippery slope. The first two reflect the *String*, specifically the lag of *String* ($String_{it-1}$) which captures a firm's consecutive human intervention in the quarterly financial statements at time *t-1* and the indicator variable HI_{it}, capturing the current human intervention in the financial statements in the current quarter *t*. By integrating both the lagged string ($String_{it-1}$) and the current *HI* it ensures our model not only reflects the present state but also considers the firm's immediate historical behavior. Neglecting to account for this lag *String* could risk overlooking significant past variations which may be substantial. The third characteristic is Abn_FSD_{it} capturing an abnormal change in human intervention in the financial statements in the current quarter *t* relative to prior periods. Consistent with Chava and Jarrow (2004), we include industry fixed effects, based on two-digit SIC code, to control for the fact that the likelihood of misreporting

[12]Thus, the hazard event only takes the value of one if the misreporting event has already been revealed by time *t*, the date of the forecast. For example, if a firm misreports in quarter *t-1*, but this is only later revealed in quarter *t+3*, then the hazard event is zero as we do not know about this misreporting behaviour at time *t*. However, when real-time progresses to *t+3*, this fresh information will be incorporated into the model leading to the hazard event being assigned as a one in *t-1* the date of the misreporting event. In the case of firms that never misreport the hazard event always equals zero.

can differ for firms in different industries.[13] Finally, because in a hazard model multiple time observations for each firm are unlikely to be independent, we also estimate the model using robust standard errors clustered at the firm level.

Thus, we estimate the specific hazard model as follows:

$$h(t|X)=h_0(t)exp(\beta_1 \times String_{it-1}+\beta_2 \times HI_{it}+\beta_3 \times Abn\_FSD_{it}+Industry\ FE) \quad (3)$$

where $h(t|X)$ is the hazard score at time t conditional on the covariates *String, HI, Abn_FSD, Industry FE*; $h_0(t)$ is the baseline hazard function at time t, representing the hazard when all covariates are equal to zero (Shumway 2001). The baseline hazard rate $h_0(t)$ is assumed to be arbitrary, and no distribution assumptions are needed for estimating the model.[14] We expect both $\widehat{\beta_1}$ and $\widehat{\beta_2}$ to be positive as both coefficients capture a pattern of the slippery slope. With respect to the sign of $\widehat{\beta_3}$ we have no clear expectations. Consistent with our understanding of the slippery slope behavior, when firms exhaust their options to manipulate financial statements within-GAAP when getting closer to the misreporting date, (Dechow, Ge and Schrand 2010; Dechow et al. 2011; Amiram et al. 2015), then the magnitude of *Abn_FSD* may decline, expecting that the smaller *Abn_FSD* the more likely misreporting will occur in the future. However, in practice the magnitude may not be so linear in its decline, even under the slippery slope phenomenon.

With the estimated coefficients from our model (equation 3) a misreporting hazard score ($h(t|X)$) is estimated in each quarter t for each firm i to determine the likelihood at time t that a firm will misreport in the future. For example, for firm i in our dataset we first compute the linear predictor at time t. We do so by multiplying the estimated coefficient of each

[13]Evidence in Dechow et al. (2011) supports the importance of industry effects on misreporting. They observe that the frequency of misreporting firms varies by industry. For example, in Dechow et al. (2011)'s sample, approximately 20 percent of firms in 'Durable Manufactures' industries misreport while these firms only represent approximately 11 percent of firms in the COMPUSTAT population.

[14]Details regarding the estimation of this type of model can be found in Shumway (2001) and Cox and Oakes (1984).

predictor at time t by the value of that predictor for firm i at time t and summing across all predictors:

$$Linear\_predictor_{it}=\widehat{\beta}_1 \times String_{it-1}+\widehat{\beta}_2 \times HI_{it}+\widehat{\beta}_3 Abn\_FSD_{it}+Industry\ FE. \quad (4)$$

Then, to determine the misreporting hazard score for each firm i at time t we multiply the estimated baseline hazard ($h_0(t)$) by the exponential of the linear predictor (exp ($Linear\_predictor_{it}$)). A higher misreporting hazard score indicates a greater risk of the misreporting occurring at time t, given that it has not occurred up to that time.

To discriminate between firms with a high propensity for future misreporting and those less likely to engage in such actions, we compare a firm's misreporting hazard score to the likelihood ratio of misreporting (which reflects the general trend in the population to misreport). Similar to Dechow et al (2011) the likelihood ratio is derived from the ratio of the number of firms identified as *Misreport* within our sample up to quarter t to the total number of firms observed up to and inclusive of the same quarter t. Thus, the ratio mirrors the general propensity within the sampled population to engage in misreporting behaviors over time. However, we diverge from Dechow et al. (2011) by focusing on real-time data, as our approach only counts firms in the numerator of the likelihood ratio if their misreporting has been revealed by the end of quarter t. This modification ensures a real-time assessment of misreporting risk.[15]

Firms are assessed for risk based on whether their misreporting hazard score exceeds the dynamically calculated likelihood ratio of misreporting. If a firm's score surpasses this threshold, it is flagged as high-risk, indicating a strong likelihood of future misreporting. We assign a 'red flag' to such firms as a predictive alert. Conversely, if the firm's score falls below this threshold, it is categorized as unlikely to misreport in the foreseeable future.

[15] In the case of HP Inc., who committed accounting egregious misreporting, namely fraud in 2015 and 2016, we consistently include the firm in the denominator but only include it in the numerator following the revelation of the misreporting event. For instance, the 2020 unconditional probability is 0.1907 (5,676/29,751).

The forecasting model follows an interactive learning approach which progressively widens the temporal window of data acquisition. It is crucial for the model to be privy to an extensive number of misreporting events to effectively discern underlying patterns. Thus, forecasting commences only once an adequate accumulation of misreporting events are available. After this initial training stage, the forecast model is operationalized and produces a set of hazard scores for each firm. This facilitates the identification of 'red flags' for those firms exhibiting a high risk of misreporting in future periods. We then follow an iterative process of annual updates.[16] As every year elapses, the model assimilates new revelations from the recently concluded year into its training dataset. To illustrate, at the culmination of 2007, the training data incorporated information regarding additional events spanning from 1961 through to 2007. As we transition to the end of 2008 the model has absorbed additional revealed misreporting cases, thus extending its training data to encompass events from 1961 through to 2008. This perpetually augmented dataset serves as the foundation upon which the misreporting hazard scores are determined at the conclusion of each year using quarterly data. These hazard scores then allow for the identification of annual red flags and hence steer our forecasts regarding potential misreporting in the subsequent periods, e.g., red flags in 2008 forecasting misreporting events in 2009 and following years.

Sample Selection

Our sample selection begins with all quarterly financial statement data from COMPUSTAT for the period 1961 to 2020.[17] We calculate deviations from BL using all quarterly COMPUSTAT data items appearing in the financial statements, taking the first non-zero digit for data items

[16]For brevity we present the results updating the model annually however when updating the model quarterly we obtain similar results.

[17]We find changing the start date of our sample period to either 1980 or 1990, instead of 1961, does not change the accuracy of the forecast model. We started in 1961 as the first AAER was related to a set of financial statements reported in 1971.

with an absolute value lower than one.[18] Unreported analysis confirms the quarterly data conforms to BL.

To identify misreporting firms, we collect data on AAERs, SCAs, and restatements, following Karpoff et al. (2017) and Donelson et al. (2021). Table 1 summarizes the sample selection process. We gather all AAERs issued between January 1, 1982, and December 31, 2020, during which the SEC issued 4,200 AAERs. We exclude AAERs unrelated to the first chronological fraud, those that do not identify the firm, involve non-financial statement wrongdoing, or are unlinked to a specific reporting period (Dechow et al. 2011). This results in a final sample of 642 unique firms with AAERs after removing overlapping SCA cases.[19] We also collect data on SCAs from the Stanford Securities Class Action Clearinghouse for the period January 1, 1996, to December 31, 2020. Of the 2,594 settled SCAs, we exclude those not involving a Rule 10b-5 allegation, non-GAAP violations, or those unrelated to financial statement misreporting, resulting in 545 unique firms after excluding overlaps with AAERs.[20]

Finally, we identify restatements from Audit Analytics (AA) between January 1, 1995, and December 31, 2020. Of the 18,870 restatements, we focus on the first restatement related to fraud or accounting issues, excluding overlaps with AAERs and SCAs, resulting in 6,421 unique firms.

We exclude firms with less than 22 data items in their quarterly reports, as this is the minimum required to compute BL deviations (Horton et al. 2020). We also require a minimum of two years of consecutive quarterly data immediately preceding the misstatement, based on findings from Dechow et al. (2011) and Amiram et al. (2015). This results in a final

[18]For example, in 0.0232 the first non-zero digit is 2, these non-zero numbers are observed as COMPUSTAT scales all financial statement to millions.

[19]Several of the firms received both an AAER and SCA relating to the same reporting period and therefore consistent with Donelson et al. (2021) we only keep the one that comes chronologically first.

[20]We find during Donelson et al. (2021) sample period - 1998 to 2014 - a similar magnitude of SCAs although not identical. The slight difference we believe is down to the classification of a GAAP violation which can only be determined by reading the SCA document and involves reader judgement in some cases. Given we find only a small difference this provides us with some confidence that we have successfully replicated as far as possible Donelson et al's (2021) sample identification strategy.

misreporting sample of 469 AAER firms, 424 SCA firms, and 4,783 restating firms (of which 121 related to fraud).

Our control sample consists of all US public firms in COMPUSTAT not identified as misreporting during the period 1961 to 2020. Excluding misreporting firms and those with less than 22 digits in their quarterly reports results in a control sample of 24,075 firms.

IV. RESULTS

Descriptive Statistics

Table 2, Panel A reports descriptive statistics of BL conformity measures. Specifically, we report the distributions of the *Adj_FSD*, *Abn_FSD, HI* and *String* for all firms. To gain some understanding of our measures we split the sample into those firms who we now know committed misreporting and those that did not. To emphasize that our focus is only on the period before the first misreporting takes place, we label our samples either as *Pre_Misreport* or *No_Misreport*. For example, *Pre_Misreport* represents those firms who we know, due to perfect hindsight, will misreport in the future. We find *Adj_FSD* and *Abn_FSD* scores, for *Pre_Misreport* firms have significantly higher averages (column 6) relative to the *No_Misreport* firms (column 9) (t-test and Wilcoxon test reported in columns 11 and 12).[21] Specifically, on average 58.9 percent of *Pre_Misreport* firms have an increase in human intervention *(HI)* compared to 52.4 percent of *No_Misreport* firms. We find the average *String* length over the whole sample period is 1.64 quarters. *Pre_Misreport* firms have a significantly higher average *String* lengths of 2.12 quarters relative to the *No_Misreport* firms who have an average length of 1.43 quarters.

[21]In untabulated results we find our basic *FSD* (prior to multiplying by pool of digits) has an average (median) score for all firms of 0.026 (0.023) which is in line with Amiram et al. (2015) who documents an average (median) *FSD* score of 0.029 (0.029). We observe *Pre_Misreport* firms have significantly lower *FSD* scores ($p<0.001$) than the *No_Misreport* firms which is inconsistent with the expectations of human intervention but consistent with the prior misreporting literature (Amiram et al. 2015; Chakrabarty et al. 2024). This inconsistency is due to the relatively higher number of digits reported by *Pre_Misreport* firms (average 123 digits) compared to the *No_Misreport* firms (average of 98 digits).

As expected, on average all firms have some level of human intervention in their accounts irrespective of whether they go on to misreport or not. However, we would expect, if there is a potential slippery slope that there will be a unique time-based distribution of *HI*s between the *Pre_Misreport* and *No_Misreport* groups. Specifically, that the human intervention in the *No_Misreport* group will be randomly spread across all quarters, as opposed to the *Pre_Misreport* group where they should appear to be accumulated and sequential. Table 2 Panel B details the number of quarters in a specific year when firms have an escalation in human intervention (i.e., when *HI* equals 1). If the *Pre_Misreport* firms are indeed on a slippery slope, we would anticipate longer strings of human intervention across the quarters in the years preceding the misreporting period, such as in the first year before *(t-1)*, second year before *(t-2)*, and so on, compared to *No_Misreport* firms.

In Panel B, we compare the distribution of these *HI*s across the *Pre_Misreport* and *No_Misreport* samples.[22] Overall, the results reveal that *Pre_Misreport* firms (column 1) exhibit a lower frequency of *HI* =1 in two quarters or less (combinations 1 and 2) during a year in comparison to the *No_Misreport* firms (column 4). However, this pattern shifts when examining three or all four quarters (combination 3), with *Pre_Misreport* firms showing a higher frequency of *HI*=1 compared to the *No_Misreport* firms. Specifically, 50.74 percent of *Pre_Misreport* firms in the four quarters just before the misreporting period have an increasing frequency of *HI*=1 in at least three or all four quarters (combination 3) compared to only 37.00 percent of *No_Misreport* firms. The overall distribution of *HI=1* for the *Pre_Misreport* sample is significantly different from that of the *No_Misreport* sample (see Chi-square tests).

[22]To manage the complexity of the analysis, we focus solely on four-quarter *HI*s, as the number of possible permutations becomes exponentially vast for *HI*s longer than four quarters (Chu et al. 2019). For a four-quarter *HI*s, there are 16 possible permutations.

Cross-sectional Analysis: Construct Validity

Before implementing our forecasting model, we validate our measures of human intervention in the accounting process by examining their correlation with established indicators of earnings management in the literature (Richardson, Sloan, Soliman, and Tuna 2005; Kothari, Leone, and Wasley 2005). The results are presented in Table 3 and all variable definitions are in Appendix 1. Consistent with our expectations, both *HI* (columns 1 and 2) and *String* (columns 5 and 6) show positive and statistically significant associations with signed discretionary accruals (*DA*) and the rate of change in accruals ($\Delta RSST$), with significance at the 1 percent level.

Additionally, we find that *HI* and *String* are significantly correlated with motivations underlying escalating human intervention, particularly the achievement of benchmarks (Chu et al. 2019). Specifically, we observe a significant increase in the likelihood of firms reporting positive earnings strings (*E_String*; columns 3 and 7) and meeting or beating analyst forecasts (*M/B*; columns 4 and 8) as human intervention escalates.[23]

These findings support the validity of our metrics for escalating human intervention, demonstrating their interrelation with recognized indicators of earnings management. This validation reinforces the potential efficacy of our model as an analytical tool for forecasting misreporting.

Forecasting Misreporting for our Hold-Out-Sample

Table 4 reports the estimates from the Cox model and the forecast model's accuracy. Panel A provides the estimates for the Cox model's annual iterations run between 2006 (using data from 1961 to 2006) and 2017 (using data from 1961 to 2017). We find the coefficients are consistent with our expectations of the slippery slope phenomenon. We find both the lagged *String* and

[23] When the dependent variable is *M/B* we also include analyst specific controls consistent with prior literature: average number of companies (*F_Follow)* and industries (*I_Follow*) analyst follows and average number of forecasting experience of the analyst: in totality (*G_Exp)*; at firm level (*F_Exp)* and at industry level (*I_Exp*). See Appendix 1 which provides details on the measurement of these variables.

HI, are positive and statistically significant (p<0.001). The coefficients on lagged *String* ranges from 0.019 to 0.022 indicating that for each additional consecutive quarter of increasing human intervention in time *t-1*, the risk of a firm misreporting in the future increases by approximately 1.02 times. Furthermore, if a firm also increases the level of human intervention in the current quarter the firm is approximately 2.599 to 3.636 times more likely to misreport in the future compared to firms with no *HI;* the coefficients on *HI* range between 0.955 to 1.291. We also find negative coefficients on *Abn_FSD* indicating that for each unit change in the abnormal deviation from BL, the likelihood the firm will misreport in the future decreases by approximately 4.1 percent - 11.4 percent, all else being equal.[24] The Harrell's C-index (equivalent to the AUC for a Cox model) of between 0.646 – 0.668 indicates the model has discriminative power.

Notably the magnitude of the coefficients remains relatively unchanged regardless of the time period analyzed, suggesting that the effects of the slippery slope are not artifacts of specific time periods or transient conditions but rather reflect persistent patterns. These results underscore the model's generalizability and suggests the model is a highly credible tool for forecasting future misreporting regardless of the time horizon under consideration.

We initiated our forecasting model in 2006, utilizing data from 1961 up to 2006 for the initial training phase. The year 2006 was chosen as the model had access to a substantial number of misreporting events, enabling it to discern meaningful patterns. Specifically, by the end of 2006, the model could learn from 64.31 percent of all revealed misreporting events, totaling 3,650 incidents out of 5,676. Each subsequent year the model incorporates the latest data on revealed misreporting events. For instance, by 2010, the model included an additional

[24]It is worth pointing out that the total effect of the slippery slope over the hazard rate of misreporting is given by $\widehat{\beta_1}+\widehat{\beta_2}+\widehat{\beta_3}$ and that a negative $\widehat{\beta_3}$ would not imply that increasing (decreasing) human intervention decreases (increases) the hazard rate of future misreporting, because the interpretation should be made in conjunction with $\widehat{\beta_1}$ and $\widehat{\beta_2}$.

806 newly revealed misreporting cases, reflecting 78.51 percent of the total misreporting cases, compared to the model in 2006. For years post-2018, we found no revelations related to firms misreporting for the first time during our sample period ending in 2020, consistent with the inherent time lag associated with such revelations.[25]

For each yearly run of the model, we derive a misreporting hazard score for each firm and assign a red flag if their hazard score exceeds the likelihood ratio for at least two consecutive quarters.[26] Using the hold-out sample, Table 4 Panel B reports the model's forecasting ability to correctly identify firms that will misreport in the future over differing temporal spans: one, two, and three or more years before the misreporting event. The initial run of the model in 2006 correctly forecasts one year ahead for 51.71 percent of firms that misreported for the first time in 2007 (column 4), two years ahead for 46.80 percent of firms that misreported for the first time in 2008 (column 7), and three or more years ahead for 49.95 percent of firms that misreported in 2009 or later (column 10). Specifically, out of the total sample of 5,710 firms in 2006, 1,368 went on to misreport post 2006, of which we correctly flagged a total of 681 whilst also correctly identifying 3,780 firms who did not misreport in the future. For the 2006 run, the model achieved an overall accuracy of 78.13 percent, with a precision of 54.79 percent and an overall sensitivity rate of 49.78 percent (see columns 13 to 15 respectively).

Through our iterative process, we observed the model's forecasting ability significantly improved over time with annual updates of newly revealed misreporting events. For instance, in 2011, it accurately forecasted 81.77 percent of misreporting events that took place one year ahead in 2012. By the end of 2016, the model could accurately forecast 96.00 percent of

---

[25]To reiterate we are only interested in a firm's first misreporting event. So, although during the period 2012 to 2018 for example there were a total of 2,111 misreporting years only 796 firms were misreporting for the first time. For the years post 2018 we find there are no revelations relating to firms who misreported for the first time during our sample period ending in 2020.
[26]This requirement of two consecutive quarters is to avoid spurious 'red flags'.

misreporting firms one year ahead of the 2017 misreporting events. Overall, from 2011 to 2017, the model accurately forecasted, on average, 90.58 percent of misreporting events one year ahead, 83.61 percent two years ahead, and 74.14 percent three or more years ahead, with an overall average accuracy of 73.24 percent and a precision of 36.29 percent.

Our findings indicate that once our model flags a firm as being at a high-risk for future misreporting, this warning remains consistent until the misreporting occurs. This sustained alert highlights the model's effective detection of increasing risk factors, providing continuous and dependable warnings and significantly enhancing stakeholders' ability to pre-empt potential financial misreporting.

Notably, as our observational window narrows toward more recent years the accuracy and precision decreases. This trend may indicate a truncation bias, reflecting the considerable time lag between the actual occurrence of a misreporting event and its subsequent disclosure. Chakrabarty et al. (2024) suggests this latency could extend up to seven years. Hence, many red flags raised by the model, for instance at the end of 2016, 2017, might indeed be accurate, but the respective misreporting events have yet to be revealed by the end of our sample period 2020, thereby underestimating our true level of accuracy. Preliminary data from the post-2020 period reveals that of the 1,267 red flags identified in 2017, approximately 313 can be attributed to misreporting incidents disclosed only between 2020 and June 2024. This suggests a recalibrated 2017 precision rate of approximately 24.78 percent, a significant increase from the initially observed 0.08 percent, and an accuracy rate that has improved to 76.65 percent from 68.985 percent.

Alternative Model Specifications

In our investigation of potential enhancements to the predictive capacity of our model, we explored the application of alternative measures within the slippery slope framework. Historically, literature has identified elevated reported accruals preceding manipulation periods, with noted observations during the years leading up to such financial discrepancies (Beneish, 1997; Dechow et al., 2011; Chu et al., 2019). These elevated accruals often correspond with periods characterized by frequent benchmark surpassing (Chu et al., 2019). Additionally, the prevalence of increased *F-Score* metrics has been documented consistently up to three years before a misreporting period. Based on these precedents, our analysis incorporates a detailed examination of the components of the *F-Score*, discretionary accruals and earnings strings. This investigation aims to evaluate both their standalone predictive efficacy and to determine the impact of integrating these components into our model as alternative measures. For example, we revise our model to incorporate the *F-Score* variables as outlined by Dechow et al. (2011) and evaluate the performance differences between our original standalone model and the modified versions including these variables. To address the variability in sample sizes resulting from different data constraints across model specifications, we methodically recalibrate our primary model to maintain consistency and ensure comparability among all variations.

Table 5 reports the comparative efficacy of different model configurations on our hold-out sample spanning from 2011 to 2017. When employing the *F-Score* components exclusively, we noted a predictive accuracy of 93.48 percent in forecasting misreporting firms a year ahead, which compares favorably to the 88.85 percent accuracy of our model. However, the increase in sensitivity observed with the *F-Score* components was offset by a marked reduction in specificity. Specifically, only 37.79 percent of firms not misreporting were correctly identified using the *F-Score* components alone, compared to 78.84 percent by our model. Thus, while integrating the *F-Score* may enhance sensitivity, it also results in a significant increase in false positives, substantially lowering both the overall accuracy and precision. The accuracy deteriorated by 32.61 percent, and precision fell by 29.53 percent, relative to our model, underscoring a significant reduction in the area under the ROC curve when restricted to comparisons across all models.

In a similar vein, the inclusion of performance-matched signed discretionary accruals (*DA*) measured using the methodology of Kothari et al. (2005), or earning strings (*E_String)*, yielded an average forecasting accuracy of 93.62 percent and 93.84 percent respectively, for misreporting firms a year ahead, outperforming our model's 88.34 percent. The model combining both earning strings (*E_String*) and discretionary accruals (*DA*) marginally improves the model's sensitivity by 1.34 percent. However, these measures, similar to the alternative *F-Score* implementation, introduce substantial noise, significantly diminishing the overall model's accuracy and precision. For instance, the model based on discretionary accruals alone, reduces the overall accuracy to 37.02 percent and precision to 18.29 percent. These metrics represent reductions of 38.68 percent and 30.47 percent in accuracy and precision, respectively, compared to our model. These declines are also reflected in a markedly lower AUC.

The aggregate findings from these examinations underscore the significant benefits of employing our BL measure over time, in contrast to the periodic measures employed to capture slippery slope attributes such as accruals or the *F-Score* components. The empirical measures such as accruals or the *F-Score* are influenced by varying firm-specific characteristics over time (Owens et al., 2017) and therefore have more difficulty differentiating between firms, unlike the deviations from BL measures which are less likely to bear any ex-ante relationship with the underlying firm characteristics or its economic performance across different periods (Amiram et al., 2015). Consequently, the reliance on these alternative empirical measures generally results in lower accuracy due to the increment in false positives. Figure 2 provides an overview of the different specifications and their accuracy levels.

We also examine alternative specifications to our model to assess its sensitivity. Specifically, we examine different combinations of our predictive variables and examine new predictive variables: contemporaneous *String*$_t$, the sum of *HI* (*ΣHI*) over all available quarters, the percentage of *HI* over all available quarters *(%HI*), and Amiran's *FSD* score. In addition, we also examine the forecasting ability of the model when we only include the industry fixed effects. Lastly, we re-examine the construction of *HI* by using the abnormal changes in the *F-Score* as an alternative to the abnormal changes in BL deviations.

The specification that utilizes only industry fixed effects as a predictive variable exhibits the lowest accuracy level among all the models evaluated. The remaining specifications have accuracy ranging between 68.73 percent to 94.87 percent with correct *No_Misreport* identification ranging between 24.37 percent to 81.00 percent. It is important for stakeholders, when deciding which specification to use, to consider the trade-off between improved accuracy in anticipating misreporting events and the increased potential for misclassifying non-misreporting firms given the cost of this misclassification (Beneish and Vorst 2021).

Fraud Only Sample

We also examine a subset of our misreporting sample involving fraud cases only, given that fraud constitutes the gravest form of misreporting. To this end, a refined sample was constructed, encompassing only instances of fraud and non-misreporting firms. This sample thus deliberately excluded firms involved in restatements that were publicly revealed at the time of running the model consistent with Chakrabarty et al. (2024). Furthermore, a new likelihood ratio specifically tailored to assess the likelihood of fraud was constructed, paralleling the methodology used for the misreporting likelihood ratio (Dechow et al, 2011). In untabulated results the model's predictive efficacy for fraud cases using our hold-out sample 2011-2017 was found to be 89.36 percent one year prior to the fraud event, 85.71 percent two years in advance, and 79.31 percent three or more years before the fraud occurrence. The model also correctly identifies 69.99 percent of *No_Misreport* firms, resulting in an overall accuracy of 70.20 percent with a precision of 3.21 percent.

In-Sample Comparison

Consistent with the existing literature, the in-sample forecasting capability of our model was also examined (e.g., Chakrabarty et al., 2024; Dechow et al., 2011). Table 6 presents the comparative efficacy of our model against alternative specifications. Notably, our model demonstrates significantly higher accuracy and precision, effectively forecasting 99.86 percent of all future misreporting firms, whereas the *F-Score* component model and the discretionary accruals plus string model forecasts 97.97 percent and 99.19 percent of such firms, respectively.[27] These results highlight the superior predictive capabilities of our model across both sensitivity and specificity metrics compared to alternative models. Specifically, our model achieves an accuracy of 87.68 percent and a precision of 60.79 percent, outperforming the *F-Score* components model, which has an accuracy of 69.07 percent and a precision of 39.71 percent. This heightened accuracy is further evidenced in the AUC results reported in column (7), confirming the robustness of our model in varying analytical scenarios.

V. CONCLUSION

Overall, our study provides an innovative analytical tool that to date has not been available but has been much needed to further address the concerning issues of misreporting. Extending upon prior research that recognized the slippery slope as a precursor to misreporting, our work offers a new instrument that significantly enhances the comprehension of the processes leading to financial misreporting. Notably, our model forecasts future misreporting with impressive accuracy: 90.58 percent for the following year, 83.61 percent for two years ahead, and 74.14

[27]The in-sample forecast rates for our model are very similar to the out-of-sample forecast rates presented in Table 5, which suggests our model does not suffer from overfitting (Chakrabarty et al. 2024).

percent for three years ahead. By systematically tracking deviations from BL to identify patterns of increasing human intervention in financial reporting, our approach provides stakeholders with the necessary insights to help them mitigate potential adverse outcomes. This model not only forecasts misreporting but also enriches our understanding of its early signs. Its straightforward design and use of public data ensures its adaptability in a wide range of settings, transcending the constraints of accounting practices and economic sectors. By highlighting emerging risks, the model empowers decision-makers to pre-emptively address factors that might lead to future misreporting, thus offering both a preventive and diagnostic function that is vital in the dynamic landscape of financial oversight.

REFERENCES

Amiram D, Bozanic Z and Rouen E. 2015. Financial statement errors: Evidence from the distributional properties of financial statement numbers. *Review of Accounting Studies* 20: 1540-1593. doi: https://doi.org/10.1007/s11142-015-9333-z

Amiram D, Huang, S and Rajgopal S. 2020. Does financial reporting misconduct pay off even when discovered? *Review of Accounting Studies* 25: 811-854. doi: https://doi.org/10.1007/s11142-020-09548-7

Barney B J and Schulzke K S. 2016. Moderating "cry wolf" events with excess MAD in Benford's Law research and practice. *Journal of Forensic Accounting Research* 1(1): A66-A90. doi: https://doi.org/10.2308/jfar-51622

Beneish M D. 1997. Detecting GAAP violation: implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16(3): 271-309. doi: https://doi.org/10.1016/S0278-4254(97)00023-9

Beneish M D. The detection of earnings manipulation. 1999. *Financial Analysts Journal* 55(5): 24-36. doi: https://doi.org/10.2469/faj.v55.n5.2296

Beneish M D, Farber D B, Glendening M and Shaw K W. 2023. Aggregate Financial Misreporting and the Predictability of U.S. Recessions and GDP Growth. *The Accounting Review* 98(5): 1-31. doi: https://doi.org/10.2308/TAR-2021-0160

Beneish M D and Vorst P. 2022. The cost of fraud prediction errors. *The Accounting Review* 97(6): 91-121. doi: https://doi.org/10.2308/TAR-2020-0068

Brazel J F, Jones K L, Thayer J and Warne R C. 2015. Understanding investor perceptions of financial statement fraud and their use of red flags: Evidence from the field. *Review of Accounting Studies* 20: 1373-1406. doi: https://doi.org/10.1007/s11142-015-9326-y

Brown T J. 2014. Advantageous comparison and rationalization of earnings management. *Journal of Accounting Research* 52(4), 849-876. doi: https://doi.org/10.1111/1475-679X.12054

Cecchini M, Aytug H, Koehler G J and Pathak P. 2010. Detecting management fraud in public companies. *Management Science* 56(7): 1146- 1160. doi: https://doi.org/10.1287/mnsc.1100.1174

Chakrabarty B, Moulton P C, Pugachev L and Wang X F. 2024. Catch Me If You Can: In Search of Accuracy, Scope, and Ease of Fraud Prediction. *Review of Accounting Studies* (forthcoming). Available at *SSRN*: https://ssrn.com/abstract=3352667.

Chava S and Jarrow R A. 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8(4): 537-569. doi: https://doi.org/10.1287/mnsc.2017.2739

Chen A, Lee S and Zapatero F. 2022. Clinging Onto the Cliff: A Model of Financial Misconduct. Available at *SSRN*: https://ssrn.com/abstract=3964115.

Chu J, Dechow P M, Hui K W and Wang A Y. 2019. Maintaining a reputation for consistently beating earnings expectations and the slippery slope to earnings manipulation. *Contemporary Accounting Research* 36(4): 1966-1998. doi: https://doi.org/10.1111/1911-3846.12492

Cox D R and Oakes D. 1984. *Analysis of Survival Data*. Volume 21: CRC press.

Dechow P M, Ge W and Schrand C. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics* 50(2-3): 344-401. doi: https://doi.org/10.1016/j.jacceco.2010.09.001

Dechow P. M, Ge W, Larson C R and Sloan R G. 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28(1): 17-82. doi: https://doi.org/10.1111/j.1911-3846.2010.01041.x

Dechow P M, Sloan R G and Sweeney A P. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC.

Contemporary Accounting Research 13(1): 1-36. doi: https://doi.org/10.1111/j.1911-3846.1996.tb00489.x
Dichev I D, Graham J R, Harvey C R and Rajgopal S. 2013. Earnings quality: Evidence from the field. *Journal of Accounting and Economics* 56(2-3): 1-33. doi: https://doi.org/10.1016/j.jacceco.2013.05.004
Donelson D C, Kartapanis A, Mcinnis J and Yust C G. 2021. Measuring accounting fraud and irregularities using public and private enforcement. *The Accounting Review* 96(6): 183-213. doi: https://doi.org/10.2308/TAR-2018-0592
Dyck A, Morse A and Zingales L. 2024. How pervasive is corporate fraud? *Review of Accounting Studies* 29: 736-769. doi: https://doi.org/10.1007/s11142-022-09738-5
Gino F and Bazerman M H. 2009. When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. *Journal of Experimental Social Psychology* 45(4): 708-719. doi: https://doi.org/10.1016/j.jesp.2009.03.013
Ham, C., Lang, M., Seybert, N., & Wang, S. (2017). CFO narcissism and financial reporting quality. *Journal of Accounting Research*, *55*(5), 1089-1135. doi: https://doi.org/10.1111/1475-679X.12176
Hill T P. 1998. The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist* 86(4): 358-363. doi: https://www.jstor.org/stable/27857060
Horton J, Krishna Kumar D and Wood A. 2020. Detecting academic fraud using Benford law: The case of Professor James Hunton. *Research Policy* 49(8). doi: https://doi.org/10.1016/j.respol.2020.104084
Hsü E. An experimental study on "mental numbers" and a new application. 1984. *The Journal of General Psychology* 38(1): 57-67. doi: https://doi.org/10.1080/00221309.1948.9711768
International Auditing and Assurance Standards Board (IAASB) 2022. *IAASB Handbook*.
Johnson G G and Weggenmann J. 2013. Exploratory research applying Benford's law to selected balances in the financial statements of state governments. *Academy of Accounting and Financial Studies Journal*.
Karpoff J M, Koester A, Lee D S and Martin G S. 2017. Proxies and databases in financial misconduct research. *The Accounting Review* 92(6): 129-163. doi: https://doi.org/10.2308/accr-51766
Kothari S P, Leone A J and Wasley C E. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics*. 2005; 39(1): 163-197. doi: https://doi.org/10.1111/1475-679X.00069
Kubovy M. Response availability and the apparent spontaneity of numerical choices. 1977. *Journal of Experimental Psychology: Human Perception and Performance* 3(2): 359-364. doi: https://psycnet.apa.org/doi/10.1037/0096-1523.3.2.359
Malenko N, Grundfest J A. and Shen Y. Quadrophobia: Strategic Rounding of EPS Data. 2022. *Journal of Financial and Quantitative Analysis* 1-43. doi: https://doi.org/10.1017/S0022109022001375
Nigrini M J. A taxpayer compliance application of Benford's Law. 1996. *The Journal of the American Taxation Association* 18(1): 72-91.
Owens E L, Wu J S and Zimmerman J. Idiosyncratic Shocks to Firm Underlying Economics and Abnormal Accruals. 2017. *The Accounting Review*. 92(2): 183-219. doi: https://doi.org/10.2308/accr-51523
Public Company Accounting Oversight Board (PCAOB), 2022. Chair Williams challenges the profession to uphold the highest standards of quality, Speech, December 12, 2022.,

available at https://pcaobus.org/news-events/speeches/speech-detail/pcaob-chair-williams-delivers-remarks-at-international-institute-on-audit-regulation

Pinkham R S. On the distribution of first significant digits. 1961. *The Annals of Mathematical Statistics* 32(4): 1223-1230. doi: https://www.jstor.org/stable/2237922

Pulliam, S. 2005. At Center of Fraud, WorldCom Official Sees Life Unravel. *The Wall Street Journal.* Available at https://www.wsj.com/articles/SB111162088282688144.

Richardson S, Tuna I and Wu M. 2002. Predicting earnings management : The case of earnings management Available at https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=41dff156a33b6809aad23e70919cf8a747b9cd9d

Richardson S, Sloan R G, Soliman M T and Tuna I. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39(3): 437-485. doi: https://doi.org/10.1016/j.jacceco.2005.04.005

Schrand C M and Zechman S L C. 2012. Executive overconfidence and the slippery slope to financial misreporting. *Journal of Accounting and Economics* 53(1-2): 311-329. doi: https://doi.org/10.1016/j.jacceco.2011.09.001

Securities Exchange Commission (SEC). 2022. SEC Announces Enforcement Results for FY22. Available at https://www.sec.gov/news/press-release/2022-206

Shumway T. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business* 74(1): 101-124. doi: https://www.jstor.org/stable/10.1086/209665

APPENDIX 1: Variable Definitions

Variable	Definition
Abn_FSD_{it}	Difference between current quarter *Adj_FSD* and average *Adj_FSD* for preceding quarters for the same firm $Abn\_FSD_{it}=Adj\_FSD_{it} - Avg\_Adj\_FSD_{i(t-1)}$ subscripts *i* and *t* stand for firm and quarter respectively
Adj_FSD_{it}	*FSD* adjusted for the variability and number of line items (N_{it}) and is calculated as follows: $Adj\_FSD_{it}=Mod\_FSD_{it}*N_{it}$ where, Mod_FSD_{it} is the Johnson and Weggnmann (2013)'s adjusted FSD_{it} and is calculated as follows. $Mod\_FSD_{it}=\frac{\sum_{j=1}^{9}\lvert(X_{jit})- FSD_{it}\rvert}{9}$ and X_{jt} is the difference between the actual frequency of each first digit *j* for firm *i* quarter *t* financial statement items (*AF*), and the expected frequency determined by the BL (*EF*) distribution. FSD_{it} is calculated as follows: $FSD_{it}=\frac{\sum_{j=1}^{9}\lvert AF_{jit} - EF_{jit}\rvert}{9}$ subscripts *j, i* and *t* stand for digit, firm and quarter respectively.
N_{it}	Total number of items reported in the quarterly financial statements (i.e., balance sheet, income statement, cash flow statement and notes to accounts). Subscripts *i* and *t* stand for firm and quarter respectively.
HI_{it}	Is an indicator variable that takes the value 1 if Abn_FSD_{it} is greater than zero; and 0 otherwise. Subscripts *i* and *t* stand for firm and quarter respectively.
$String_{it}$	The sum of consecutive *HI* i.e., it is the uninterrupted sequence of human intervention. Subscripts *i* and *t* stand for firm and quarter respectively.
$Pre\_Misreport_{i}$	Is an indicator variable that takes the value 1 if a firm will receive during our sample period either an SEC Accounting and Auditing Enforcement Releases (AAER) due to an accounting manipulation; is subject to a Security Class Action (SCA) for an accounting manipulation; or announces a restatement (Audit Analytics); and 0 otherwise. Subscripts *i* stand for firm.
$No\_Misreport_{i}$	Is an indicator variable that takes value 1 if a firm will not receive/announce during our out-of-sample period an Accounting and Auditing Enforcement Releases (AAER); is not subject to a Security Class Action (SCA); and has not announced any restatement of their financials (Audit Analytics); and 0 otherwise. Subscripts *i* stand for firm.
$RSST_{it}$	$(\Delta WC_{it} + \Delta NCO_{it} + \Delta FIN_{it})/$ *Average total assets*$_{it}$, where WC_{it} = [$\underline{Current\ Assets_{it} - Cash_{it}\ and\ Short\text{-}term\ Investments_{it}}$] – [$Current\ Liabilities_{it} - Debt\ in\ Current\ Liabilities_{it}$]; NCO_{it} = ($Total\ Assets_{it} - Current\ Assets_{it}$) – $Investments\ and\ Advances_{it}$] – $Total\ Liabilities_{it} - Current\ Liabilities_{it} - Long\text{-}term\ Debt_{it}$]; FIN_{it} = $Short\text{-}term\ Investments_{it} + Long\text{-}term\ Investments_{it}$] – $Long\text{-}term\ Debt_{it} + Debt\ in\ Current\ Liabilities_{it} + Preferred\ Stock_{it}$]; following Richardson et al. (2005) and Dechow et al. (2011). Subscripts *i* and *t* stand for firm and year respectively.
$\Delta RSST_{it}$	Is *RSST* of firm *i* in year *t* divided by *RSST* of firm *i* in year *t-1*.
DA_{it}	Residual from the Performance based Discretionary Accruals model explained by Kothari et al. (2005) and is calculated as follows: $Total\ Accruals_{it}=\alpha+\beta_1\frac{1}{Assets_{it-1}} + \beta_2\Delta Sales_{it}+\beta_3 PPE_{it}+\beta_4 ROA_{it} + \varepsilon$ Where *Total Accruals* is change in current asset excluding cash minus change in current liabilities excluding current debt minus depreciation, *Assets* is the total

	asset, *Δ sales* is the simple change in sales between years, *PPE* is net property plant and equipment, and *ROA* is return on assets calculated as Net Income over lagged total assets. All the variables are scaled by lagged total assets. Subscripts *i* and *t* stand for firm and year respectively.
E_String_{it}	Is an indictor variable that takes value 1 if the change in net income is positive and 0 otherwise. Subscripts *i* and *t* stand for firm and year respectively.
M/B_{it}	Is an indictor variable that takes value 1 if the actual EPS for a quarter is equal to or greater than the analyst consensus and 0 otherwise. Subscripts *i* and *t* stand for firm and quarter respectively.
$Sales\_G_{it}$	The natural logarithm of change in net sales of firm *i* from year *t* to *t-1*, divided by net sales in for firm *i* in year *t-1*.
$Leverage_{it}$	Leverage of client firm *i* in year *t* defined as long-term debt plus short-term debt scaled by total assets.
$Size_{it}$	The natural logarithm of total assets of firm *i* in year *t*.
ROA_{it}	The natural logarithm of return on Assets of client firm *i* in year *t* measured as net income before extraordinary items divided by total assets.
$Loss_{it}$	Is an indicator variable which equals 1 if the net income of firm *i* in year *t* is negative, and 0 otherwise.
C_Ratio_{it}	Current Asset divided by Current Liabilities. Subscripts *i* and *t* stand for firm and year respectively.
F_Follow_{it}	The average number of companies all analysts' following firm *i* follows in quarter *t*
G_Exp_{it}	The average number of years of forecasting experience for all analysts' following firm *i* in quarter *t*.
F_Exp_{it}	The average number of years of forecasting firm *i* by all analysts' following firm *i* in quarter *t*.
I_Exp_{ijt}	The average number of years of forecasting industry *j* by all analysts' following firm *i* in quarter *t*.
I_Follow_{it}	The average number of industries all analysts' following firm *i* follows in quarter *t*
$Likelihood\ Ratio_{it}$	Consistent with Dechow et al. (2011) methodology the probability is calculated by dividing the number of misreporting firms up to and including quarter *t* that is revealed till the end of our sample period by the total number of firms up to and including quarter *t*. $Likelihood\ Ratio_t = \frac{\#Revealed\ misreporting\ firms\ up\ to\ and\ including\ quarter\ t}{\#Total\ firms\ up\ to\ and\ including\ quarter\ t}$
Red_Flag_{it}	Is an indictor variable that takes value 1 if our hazard score is greater than likelihood ratio for two consecutive quarters and 0 otherwise Subscripts *i* and *t* stand for firm and year respectively.

TABLE 1: Sample Selection

This table presents the sample selection description of SEC enforcement (AAERs), security class actions (SCA) and Audit Analytics restatements. Panel A reports the number of AAERs issued and the number of unique firms during the sample period 1 January 1982 to 1 May 2021. Panel B reports SCAs filed and the unique number of firms that received a SCA during 1 January 1996 to 1 May 2021. Panel C reports the Audi Analytics restatement sample during 1 January 1995 to 1 May 2021. Panel D describes the final sample of *Misreport* and *No_Misreport* firms.

Panel A: AAERs Sample

	Total
AAER issued	**4,219**
Less:	
Missing AAERs[28]	(87)
Against individuals for insider trading	(20)
Against the auditors or audit firms or investment funds	(95)
AAER with no specific company information	(85)
AAER related to review requirements	(23)
Report of Investigation	(1)
Document defining Covered Argument	(1)
	3,907
Less: AAER's not related to the first chronological fraud	(1,852)
	2,055
Less: Manipulation unrelated to Financial Statements Digits	
Audit Issues	(156)
Disclosure Issues	(108)
Bribe	(94)
Other	(56)
Less: No Specific Reporting Dates available	(167)
	1,474
Less: No COMPUSTAT accounting data available	(766)
Overlaps	(66)
Final Sample of AAERs	**642**

Panel B: Security Class Actions (SCA) Sample

	Total
SCAs Issued	**5,933**
Less: Pending and Dismissed cases	(2,979)
Settled SCAs	2,594
Less: No Section 10(b)/Rule 10b-5	(315)
Fraudulent settled SCAs	2,279
Less: No GAAP Violation	(1,037)
	1,242
Less: SCA's not related to the first chronological fraud	(53)
	1,189
Less: Manipulation unrelated to Financial Statements	
Disclosure	(326)
Shares	(37)
Other Issues	(1)
	825
Less: No COMPUSTAT accounting data available	(255)
Overlaps	(25)
Final Sample of SCAs	**545**

[28] Out of 131, 15 were intentionally omitted by the SEC. The remaining AAERs could still be under litigation and will be made publicly available only after the completion of enforcement/litigation.

Panel C: Audit Analytics (AA) Sample

	Total
AA Restatement	
Fraud	312
Other Restatements	18,565
	18,877
Less: AA's not related to the first chronological fraud	(8,252)
	10,625
Less: No COMPUSTAT accounting data available	(3,252)
Overlaps - Fraud	(60)
Overlaps - Restatement	(892)
Final Sample of AA Restatements	**6,421**
Of which:	
Fraud	**137**
Restate	**6,284**

Panel D: Final Treated (*Misreport*) and Control Sample (*No_Misreport*)

		Misreport						*No Misreport*
	Total	*Fraud*				*Restate*		
		AAER	**SCA**	**AA**	**Total**	**AA**	**Total**	**Total**
Number of Misreporting	39,220	642	545	137	1,324	6,284	7,608	**31,612**
Less: Firms with < 22 digits per quarter	8,552	(0)	(0)	(3)	(3)	(1,012)	(1,015)	(7,537)
Unrestricted Full Sample	**30,668**	**642**	**545**	**134**	**1,321**	**5,272**	**6,593**	**24,075**
Less: Firms with < 2 years of consecutive quarters immediately preceding misstatement quarter	(917)	(173)	(12)	(13)	(307)	(610)	(917)	(0)
2 Year Restricted Sample	**29,751**	**469**	**424**	**121**	**1,014**	**4,662**	**5,676**	**24,075**

TABLE 2: Descriptive Statistics for All, *Misreport* and *No_Misreport* Firms

Panel A reports descriptive statistics used to identify human intervention for the period 1961 to 2020. Specifically, it provides the descriptive statistics for the *Adj_FSD*, *Magnitude*, *HI*, and *String* for *Pre_Misreport* firms and *No_Misreport* firms. For the *Misreport* firms we only use the quarters prior to the misreporting period for all measures. *Adj_FSD* score is the *FSD* adjusted for smaller sample size using the Johnson and Weggenmann (2013) methodology and then multiplied by the pool of digits to adjust for the changes in pool of digit over time. *Abn_FSD* is the difference between *Adj_FSD* score and the average of *Adj_FSD* for all the preceding quarters. Our variable of interest *String* is the aggregate number of consecutive *HI's;* where *HI* takes the value of one if *Abn_FSD* >0 and zero otherwise. Panel B provides a comparison of permutations of a four-quarter string in each year preceding the misreporting period across the *Pre_Misreport* and *No_Misreport* samples. All variables are defined in Appendix 1.

Panel A: Descriptive Statistics

	All Firms				*Pre_Misreport*			*No_Misreport*			*Pre_Misreport Vs No_Misreport*	
	Obs.	Mean	Median	SD	Obs.	Mean	Median	Obs.	Mean	Median	t-test	Wilcoxon
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11) [6-9]	(12) [7-10]
Adj_FSD	1,429,493	2.537	2.261	1.307	459,754	2.744	2.488	969,739	2.428	2.148	0.316***	0.340***
Abn_FSD	1,429,493	0.289	0.091	1.063	459,754	0.428	0.208	969,739	0.227	0.045	0.201***	0.163***
HI	1,429,493	0.545	1.000	0.498	459,754	0.589	1.000	969,739	0.524	1.000	0.065***	0.000***
String	1,429,493	1.644	1.000	1.644	459,754	2.117	1.000	969,739	1.434	1.000	0.683***	0.000***

Panel B: Frequency and Persistency when firms have a *HI* = 1

	Pre_Misreport	*No_Misreport*
Combination	**(1)**	**(2)**
(1) *HI*=1 in no quarter (0,0,0,0)		
1st before misreporting event (FY-1)	4.26%	6.05%
2nd year before (FY -2)	5.27%	9.10%
3rd year before (FY -3)	4.64%	6.34%
4th year before (FY -4)	4.33%	6.37%
5th year before (FY -5)	3.94%	6.06%
6h year before (FY -6)	3.97%	5.54%
(2) *HI*=1 in two or one quarter		
FY -1	45.00%	56.95%
FY -2	47.30%	55.24%
FY -3	45.20%	57.17%
FY -4	46.44%	57.80%
FY -5	47.28%	57.18%
FY -6	47.29%	57.48%
(3) *HI*=1 in three or all quarters		
FY -1	50.74%	37.00%
FY -2	47.43%	35.66%
FY -3	50.17%	36.49%
FY -4	49.23%	35.83%
FY -5	48.78%	36.76%
FY -6	48.72%	36.98%

Chi-Square Tests	
Pre_Misreport vs. No_Misreport	
For All Permutations	
FY -1	500.09***
FY -2	378.02***
FY -3	525.39***
FY -4	417.30***
FY -5	285.48***
FY -6	246.89***

TABLE 3: Cross-sectional Analysis: Slippery Slope Validation and Association with Positive Outcomes

This table presents results of the cross-sectional analysis of the slippery slope characteristics, specifically, the indicator variable for human intervention (*HI*) and the cumulative string of HI's (*String*) on accruals and various outcomes using a panel regression in columns (1), (2), (5) and (6) and a panel logistic regression in columns (3), (4), (7) and (8) for our sample period 1961-2020. Our variable of interest *HI* takes the value of one *Abn_FSD* >0 and zero otherwise. The dependent value in column (1) and (5) is the change in yearly accruals (*ΔRSST*) which is determined using the Richardson et al. (2005) model. In Column (2) and (6) is discretionary accruals (*DA*) from the performance matched discretionary accrual model by Kothari et al. (2005). In Column (3) and (7) is an indicator variable if whether the firm has an earning string with the dependent variable (*E_String*) taking the value of one if the firm has reported higher earnings relative to the prior year and zero otherwise. The dependent variable in column (4) and (8) is an indicator variable equal to one if the firm meets or surpasses the analyst consensus based on Chu et al. (2019) (*M/B*) and zero otherwise. All continuous variables are winsorized at the 1st and 99th percentile. Standard errors are adjusted for clustering at the firm level. Variables are defined in Appendix 1. ***, ** and * indicate two-tailed significance at 0.01, 0.05 and 0.1 levels.

	ΔRSST	***DA***	***E_String***	***M/B***	***ΔRSST***	***DA***	***E_String***	***M/B***
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HI	**0.029*** **	**0.016*** **	**0.046*** **	**0.066*** **				
	(0.001)	**(0.002)**	**(0.004)**	**(0.014)**				
String					**0.001*** **	**0.002*** **	**0.003** **	**0.353*** **
					(0.000)	**(0.000)**	**(0.001)**	**(0.075)**
Sales_G	0.032***	-0.002	0.198***	-0.099***	0.029***	-0.005	0.189***	-0.178
	(0.003)	(0.003)	(0.006)	(0.029)	(0.003)	(0.003)	(0.006)	(0.136)
Leverage	-0.019**	0.021**	0.363***	-0.997***	-0.014	0.025**	0.361***	-1.775***
	(0.009)	(0.010)	(0.015)	(0.073)	(0.009)	(0.010)	(0.015)	(0.238)
Size	-0.033***	0.004*	-0.112***	0.273***	-0.019***	0.012***	-0.104***	0.270***
	(0.002)	(0.002)	(0.002)	(0.009)	(0.001)	(0.002)	(0.002)	(0.029)
ROA	0.318***	0.156***	0.522***	2.468***	0.311***	0.152***	0.511***	6.949*
	(0.013)	(0.012)	(0.017)	(0.330)	(0.013)	(0.012)	(0.017)	(3.749)
Loss	-0.056***	-0.019***	-1.653***	-1.812***	-0.054***	-0.019***	-1.649***	-1.548***
	(0.003)	(0.004)	(0.011)	(0.032)	(0.003)	(0.004)	(0.011)	(0.178)
C_Ratio	-0.002***	0.004***	-0.022***	-0.001	-0.003***	0.003***	-0.022***	0.010
	(0.001)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)	(0.012)
Constant	0.121***	-0.022**	1.252***	-0.594***	0.130***	-0.022**	1.328***	-0.626***
	(0.007)	(0.009)	(0.018)	(0.059)	(0.007)	(0.009)	(0.016)	(0.189)
Analyst Controls	No	No	No	Yes	No	No	No	Yes
Obs.	222,596	264,017	267,338	88,812	266,427	266,427	266,427	88,812
Adj R2	0.060	0.008			0.058	0.008		

TABLE 4: Forecast Model for Hold-Out Sample

This table reports the Cox Model estimates, the years the first misreporting was revealed and the forecast accuracy of the model. Panel A presents the estimated coefficients with the standard errors in parentheses from the Cox model run on data from 1961 to years between 2006 and 2017. For the Cox model the hazard is the first quarter of the revealed misreporting which takes the value of 1, and zero for all prior quarters. *String* is the aggregate number of consecutive *HI* where *HI* is equal to one if in a quarter if *Abn_FSD* (the difference between *Adj_FSD* and *Avg_Adj_FSD*) for all the preceding quarters is positive. Panel B presents the forecast accuracy of the model to correctly identify firms who will misreport in the future over differing temporal spans, one year before the misreporting event, two years before and three years or more before. We forecast a firm as highly likely to misreport if the absolute risk to hazard is greater than likelihood ratio. Sensitivity is the true positive rate computed as the ratio of correct forecast *Misreport* to total future *Misreport*. Accuracy is the total correct forecast rate computed as ratio of correct forecasted *Misreport* and *No_Misreport* to Total firms. Precision is the total correct forecast rate computed as the ratio of correct forecast *Misreport* to all red flags. Specificity is the true negative rate computed as the ratio of correct forecast *No_Misreport* to total *No_Misreport*. ** and * indicate two-tailed significance at 0.01, 0.05 and 0.1 levels. Robust standard errors are clustered at firm level and reported in parentheses.

Panel A: Cox Model Linear Predictors

	Sample Data from 1961 to:											
	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$String_{t-1}$	0.019***	0.021***	0.020***	0.021***	0.022***	0.017***	0.018***	0.018***	0.020***	0.021***	0.021***	0.021***
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
HI_t	0.982***	0.955***	1.018***	1.052***	1.117***	1.090***	1.215***	1.259***	1.274***	1.291***	1.287***	1.271***
	(0.055)	(0.051)	(0.049)	(0.047)	(0.046)	(0.043)	(0.042)	(0.041)	(0.040)	(0.039)	(0.039)	(0.038)
Abn_FSD_t	-0.041*	-0.042*	-0.059***	-0.062***	-0.047**	-0.078***	-0.084***	-0.091***	-0.093***	-0.101***	-0.109***	-0.114***
	(0.024)	(0.022)	(0.021)	(0.020)	(0.019)	(0.020)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	821,967	842,672	862,913	882,797	874,674	890,339	906,328	922,254	937,495	952,014	965,357	1,003,245
Unique firms:												
Total	24,573	24,961	25,278	25,711	26,225	26,592	27,375	27,810	28,109	28,436	28,579	29,579
Misreport	5,098	5,180	5,245	5,334	5,412	5,467	5,543	5,586	5,602	5,676	5,676	5,676
R^2	0.018	0.018	0.018	0.019	0.021	0.020	0.024	0.025	0.026	0.027	0.026	0.026
Wald Chi	864.53	967.36	1,093.96	1,229.59	1,424.11	1,477.53	1,896.63	2089.55	2,220.59	2,905.59	2,407.63	2,416.18
Harrel's C	0.668	0.665	0.662	0.662	0.661	0.646	0.647	0.646	0.646	0.646	0.645	0.645

Panel B: Forecasting Accuracy of the Model

	Misreport: Forecast Period									*No_Misreport*		Overall Model Efficacy		
	One year ahead			Two years ahead			Three or more years							
Model Run in Year:	Year	Red Flags	Correct[1]	Year	Red Flags	Correct[1]	Year onwards	Red Flags	Correct[1]	No Flag	Specificity[2]	Accuracy[3]	Precision[4]	Sensitivity[5]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
2006	2007	121	51.71%	2008	95	46.80%	2009	465	49.95%	3780	87.06%	78.13%	54.79%	49.78%
2007	2008	116	57.14%	2009	102	48.11%	2010	83	58.05%	3473	83.71%	77.51%	50.00%	56.15%
2008	2009	119	56.13%	2010	100	59.17%	2011	104	63.37%	2832	79.95%	75.62%	47.99%	61.27%
2009	2010	114	67.46%	2011	129	69.35%	2012	133	71.75%	2935	77.71%	76.27%	44.02%	70.50%
2010	2011	146	78.49%	2012	155	80.73%	2013	137	75.62%	3055	75.84%	76.12%	39.53%	77.47%
2011	2012	157	81.77%	2013	144	75.00%	2014	111	72.00%	2886	76.15%	76.06%	36.38%	75.58%
2012	2013	171	89.06%	2014	151	84.36%	2015	79	70.98%	3145	73.84%	74.73%	29.18%	81.38%
2013	2014	168	93.85%	2015	119	88.81%	2016	57	84.27%	3042	69.04%	70.80%	20.97%	90.05%
2014	2015	130	97.01%	2016	66	90.41%	2017	20	87.50%	2995	67.33%	68.65%	12.99%	93.94%
2015	2016	70	95.89%	2017	24	96.00%	2018	1	100.00%	3088	67.53%	68.13%	6.01%	95.96%
2016	2017	24	96.00%	2018	1	100.00%	2019			2872	68.02%	68.20%	1.82%	96.15%
2017	2018	1	100.00%	2019			2020			2814	68.97%	68.98%	0.08%	100.00%
Averages Over:												**Average based on one-year ahead**		
2006-17			74.28%			69.35%			63.41%		74.41%	74.37%	51.36%	74.28%
2011-17			90.58%			83.61%			74.14%		69.99%	73.24%	36.29%	90.58%

In 2006, Total Firms = 5,710; Total *Misreport* firms = 1,368; Total *No_Misreport* firms = 4,342; Total Red Flags in 2006 = 121+95+465+562 = 1,243.

[1]Correct (columns 4,7,10) is calculated as follows: *Misreport* Flagged / Total *Misreport*. For one year ahead in 2006 = 121/234 =51.71%; two years ahead=95/203=46.80%; three or more years ahead = 465/931 =49.95%.

[2]*Specificity* is calculated as follows: Correct *No_Misreport* / Total *No_Misreport*. For 2006 = 3,780/4,342 = 87.06%

[3]*Accuracy* is calculated as follows: (Correct *No_Misreport* + Correct *Misreport*)/Total Firms. For 2006 = (3,780+681)/5,710 = 78.13%

[4]*Precision* is calculated as follows: Correct *Misreport* / Total Red Flags. For 2006 = 681/1,243=54.79%

[5]*Sensitivity* is calculated as follows: Correct *Misreport* / Total *Misreport*. For 2006 = 681/1,368 = 49.78%

TABLE 5: Utilizing Alternative Misreporting Detection Variables

This table reports the average forecast hold-out accuracy for alternative model specifications based on one year ahead forecasts for 2011 to 2017. In rows 3,6 and 9 the *HI model* is run on the different samples sizes to aid comparability. *F-Score Comp* is the *F-Score* components as described by Dechow et al. (2011). Row 1 includes only *F-Score* components in the model, and row 2 includes the *F-Score* components in addition to our predictive variables. *DA* is performance matched discretionary accruals consistent with Korthari et al. (2005); row 4 includes only *DA* in the model, and row 5 included *DA* in addition to our predictive variables. *E_String* is a dummy variable which equals one if the change in net income is positive and zero otherwise. Row 7 includes only *E_String* in the model, and row 8 includes *E_String* in addition to our predictive variables. Sensitivity is the true positive rate computed as the ratio of correct forecast *Misreport* to total future *Misreport*. Accuracy is the total correct forecast rate computed as ratio of correct forecasted *Misreport* and *No_Misreport* to Total firms. Precision is computed as the ratio of correct forecast *Misreport* to sum of true and false positives. Specificity is the true negative rate computed as the ratio of correct forecast *No_Misreport* to total *No_Misreport*. Area under the Receiver Operating Characteristics curve (AUC) is the area under the curve that plots the true positive and false positive rates using every observation as a possible cut-off.

		Sample	*Misreport*		*No_Misreport*				
	Model Specification	**#Firms**	**Red Flags**	**Sensitivity**	**No Flag**	**Specificity**	**Accuracy**	**Precision**	**AUC**
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
1.	*F_*Score *Comp*	3,899	545	93.48%	1,253.14	37.79%	46.12%	17.31%	0.7764
2.	*F_Score Comp + HI Model*	3,899	553	94.85%	1,247.29	37.61%	46.17%	17.53%	0.7769
3.	***HI Model***	**3,899**	**518**	**88.85%**	**2,556.14**	**78.84%**	**78.84%**	**46.84%**	**0.8200**
4.	*DA*	3,325	602	93.62%	731.29	24.86%	37.19%	18.11%	0.6776
5.	*DA + HI Model*	3,325	609	94.71%	725.43	24.67%	37.24%	18.28%	0.6775
6.	***HI Model***	**3,325**	**568**	**88.34%**	**2,145.57**	**72.93%**	**75.70%**	**48.76%**	**0.8246**
7.	*E_String*	5,050	695	87.31%	901.75	21.20%	31.62%	19.22%	0.6384
8	*E_String + HI Model*	5,050	747	93.84%	921.43	21.66%	33.04%	17.56%	0.6389
9.	***HI Model***	**5,050**	**721**	**90.58%**	**2,977.43**	**69.99%**	**73.24%**	**36.29%**	**0.8139**
10.	*DA + E_String*	3,325	610	94.87%	716.00	24.37%	37.02%	18.29%	0.6375
11.	*DA + E_String + HI Model*	3,325	611	95.02%	716.43	24.39%	37.07%	18.32%	0.6377

TABLE 6: Forecast Model: In-Sample for All Specifications

This table reports the average forecast in-sample accuracy for all specifications. In rows 3,6 and 9 the *HI model* is run on the different samples to aid comparability. *F-Score* Comp are the *F-Score* components as described by Dechow et al. (2011). Row 1 includes only *F-Score* components in the model, and row 2 includes the *F-Score* components in addition to our predictive variables. *DA* is performance matched discretionary accruals consistent with Korthari et al. (2005); row 4 includes only *DA* in the model, and row 5 included *DA* in addition to our predictive variables. *E_String* is a dummy variable which equals one if the change in net income is positive and zero otherwise. Row 7 includes only *E_String* in the model, and row 8 includes *E_String* in addition to our predictive variables. Sensitivity is the true positive rate computed as the ratio of correct forecast *Misreport* to total future *Misreport*. Accuracy is the total correct forecast rate computed as ratio of correct forecasted *Misreport* and *No_Misreport* to Total firms. Precision is computed as the ratio of correct forecast *Misreport* to sum of true and false positives. Specificity is the true negative rate computed as the ratio of correct forecast *No_Misreport* to total *No_Misreport*. Area under the Receiver Operating Characteristics curve (AUC) is the area under the curve that plots the true positive and false positive rates using every observation as a possible cut-off.

		Sample	***Misreport***		***No_Misreport***				
	Model Specification	**#Firms**	**Red Flags**	**Sensitivity**	**No Flag**	**Specificity**	**Accuracy**	**Precision**	**AUC**
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
1.	*F-Score Comp*	23,735	4,771	97.97%	11,622	61.61%	69.07%	39.71%	0.7979
2.	*F-Score Comp + HI Model*	23,735	4,678	95.85%	12,041	63.83%	70.40%	40.62%	0.8041
3.	***HI Model***	**23,735**	**4,864**	**99.88%**	**14,926**	**79.12%**	**83.38%**	**55.25%**	**0.8185**
4.	*DA*	23,317	4,649	99.19%	10,486	56.29%	64.91%	36.34%	0.7896
5.	*DA + HI Model*	23,317	4,649	99.19%	11,461	61.53%	69.10%	39.34%	0.8001
6.	***HI Model***	**23,317**	**4,631**	**98.81%**	**13,093**	**70.28%**	**76.01%**	**45.54%**	**0.8133**
7.	*E_String*	29,751	5,671	99.91%	14,568	60.51%	68.03%	37.36%	0.7979
8	*E_String + HI Model*	29,751	5,670	99.89%	15,622	64.89%	71.57%	40.15%	0.7982
9.	***HI Model***	**29,751**	**5,668**	**99.86%**	**20,419**	**84.81%**	**87.68%**	**60.79%**	**0.8161**
10.	*DA + E_String*	23,317	4,649	99.19%	11,364	61.00%	68.68%	39.02%	0.7973
11.	*DA + E_String + HI Model*	23,317	4,649	99.19%	11,365	61.01%	68.69%	30.03%	0.7974

Figure

FIGURE 1: Differing Objectives of Prediction Models versus Forecasting Model

We provide an illustration to highlight the objectives of the prediction models versus our forecast model. Assume period *t* marks the year when the firm for the first time misreported their published financial statements. Period *t-n* to *t-1* represents the period before misreporting takes place, *t+1* to *t+n* encompasses the post-misreporting period up to the time it is sanctioned in our illustration when an AAER for instance is issued by the SEC. The information available to the misreporting prediction model at time *t* includes the misreported accounts published in time *t*. If the prediction model is accurate then it will predict in time *t* that the financial statements published in time *t* are misreported. The information available to the forecasting model, for example in time *t-3* is all quarterly reports between and including the periods *t-n* to *t-3*. At time *t-3* the forecasting model will provide a probability of future misreporting, if for example it forecasts a high probability, we then identify this firm as likely to misreport in the future, i.e., *t-3* onwards. This provides stakeholders with a three-year prevention window (*t-3* to *t-1*) to act and potentially avert the misreporting from taking place in time *t*. On the contrary, with a prediction model, intervention is restricted to the cure period (*t+1* to *t+n*) since the misreporting has already occurred.

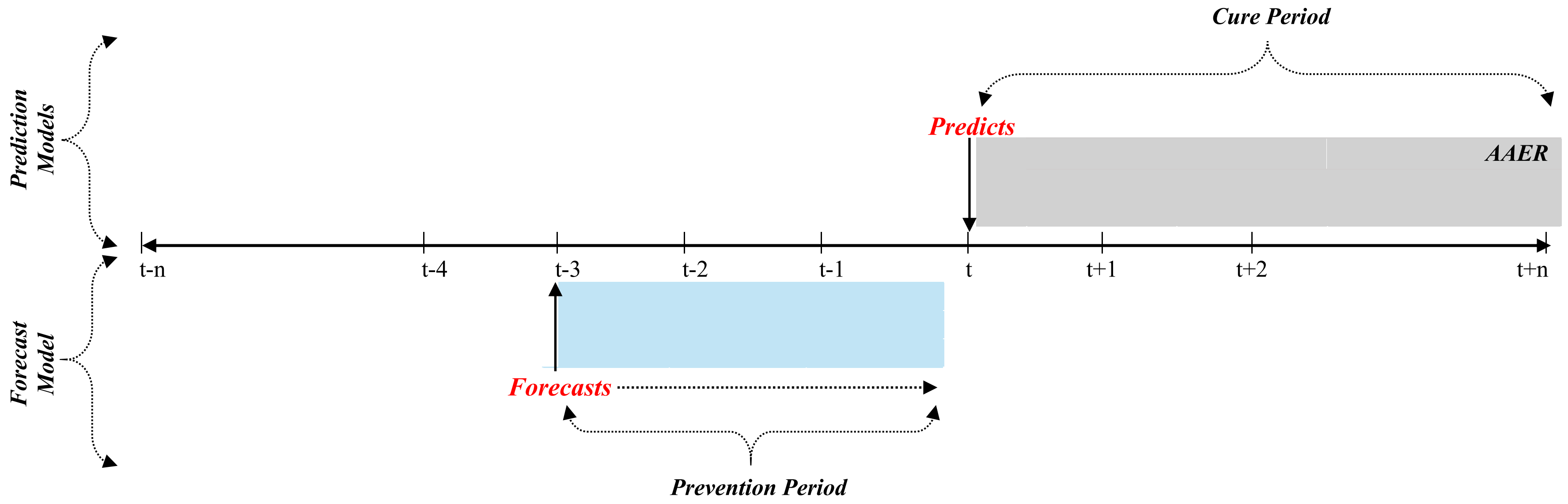

FIGURE 2: Accuracy of Alternative Specifications of the Forecast Model

The figure reports the accuracy of using alternative predictive variables to capture the slippery slope in the Cox Model (in all specifications we continue to include industry fixed effects and cluster standard errors at firm level). The alternative predictive variables considered are detailed in the data label and the blue dot represents the accuracy of each of these specifications. Our main model is highlighted in blue. The forecast model is run yearly starting in 2011. The horizontal axis reports the percentage of firms that the model correctly forecasted as likely to misreport in the future. The vertical axis reports the percentage of firms that the model correctly forecasted as not likely to misreport in future. *ΣHI* equals the cumulation *HI* over all available quarters. *%HI* is *ΣHI* divided by the number of available quarters. *Mag* equals *Abn_FSD$_t$*. All other variables are as previous reported.