Product Market Similarity, Benchmarking, and Corporate Fraud *

Audra Boone† William Grieser† Rachel Li‡ Parth Venkat‡

December 3, 2022

Abstract

We document a strong negative relation between firms' product market similarity to their rivals and their observed fraud rate. This disciplining effect is larger than most documented fraud predictors and remains virtually unchanged after controlling for those predictors and traditional measures of competition. By exploiting variation from rivals' newly available public disclosures, as well as cross-sectional variation in firm complexity, institutional ownership, and analyst coverage, we show that the effect is most likely driven by firms' external information environment. Our analyses suggest that greater product similarity enhances the external information environment, which improves external monitoring and disciplines manager reporting behavior.

Keywords: Corporate Fraud, Product Market Competition, Product Similarity, Benchmarking, Corporate Governance

*We thank Anne Albrecht, Khrystyna Bochkay, Donald Bowen, Jeff Chen, Joey Engelberg, Will Gerken, Umit Gurun, Gerard Hoberg, Andy Imdieke, Simi Kedia, Zack Liu, Michelle Lowry, Tanakorn Makaew, Gonzalo Maturana, Karen Nelson, Jordan Neyland, Joerg Picard, Vesa Pursiainen, Nathan Swem, Xinxin Wang, Jared Wilson, participants at the 2019 Conference on Financial Market Regulation, the 2018 FMA Asia Annual Conference, the 2018 FMA Annual Meeting, the 2018 Australasian Banking and Finance Conference, the 2018 New Zealand Finance Conference, and the 2019 Midwest Finance Association Conference, as well as seminar participants at Drexel University, Clemson University, Southern Methodist University, Texas Christian University, Universidad de los Andes, University of Nevada Las Vegas, the U.S. Securities and Exchange Commission, and Villanova University for helpful comments.

†Department of Finance, Neeley School of Business, Texas Christian University, e-mail: audra.boone@tcu.edu, 817-257-4267; w.grieser@tcu.edu, 817-257-4910.

‡Economics, Finance, and Legal Studies, Culverhouse College of Business, University of Alabama, e-mail: qli@cba.ua.edu, 205-348-7858; prvenkat@ua.edu, 205-348-1383.

1 Introduction

Financial reporting facilitates efficient resource allocation by providing information about a firm's business activities, financial performance, and relevant circumstances to stakeholders. Financial statement fraud undermines the trust essential to a well-functioning financial system (Greenspan, 2008), which can erode firm value (Karpoff et al., 2008b; Dyck et al., 2010), impose negative externalities (Kedia and Philippon, 2009), discourage stock market participation (Guiso et al., 2008), and distort investors' allocation decisions (Giannetti and Wang, 2016; Gurun et al., 2018). It is, therefore, important to understand how stakeholders' ability to detect fraud affects managers' incentives to commit fraud. In this paper, we explore the role that product market similarity plays in these forces.

Product market similarity to peers can affect corporate financial fraud through several channels. One possible channel is that more closely related rivals provide stakeholders with better benchmarks, which enhances that firms information environment (Badertscher et al., 2013; Kim et al., 2013; Hsu et al., 2017; Engelberg et al., 2018). In turn, a stronger information environment should facilitate better monitoring and evaluation of firm performance and increase the probability of detecting fraud and other financial reporting abnormalities, such as earnings management (Dichev et al., 2013). Given that getting caught committing financial fraud results in severe pecuniary and reputational costs (Karpoff et al., 2008c; Armour et al., 2017), product similarity can discipline managers and reduce the ex-ante likelihood of committing fraud.

Another possible channel is through more intense competition, which affects managerial behavior. On the one hand, competition can reduce agency problems as a more competitive product market can reduce managerial rent extraction (Nickell, 1996; Ades and Di Tella, 1999), reduce the benefits of shirking (Hart, 1983), or force managers to focus on maximizing profits to avoid bankruptcy (Schmidt, 1997). Thus, it would reduce the likelihood that managers would commit fraud. On the other hand, competition can pressure managers to employ accounting manipulations in a race for survival (Shleifer, 2004). In this scenario, competition would lead to a higher incidence of fraud.

Our analyses suggest that higher product market similarity is associated with a reduced incidence of fraud. That is, a one standard deviation increase in product similarity is associated with around a 20% decrease in the rate of settled SEC enforcement actions and fraud-driven class action lawsuits. Our main proxy of product market similarity is the average pairwise product similarity score between the firm and all of its rivals developed by Hoberg and Phillips (2016) based on the business descriptions in firms 10-Ks. Higher average similarity scores indicate that rivals operate in similar product markets, share similar economic fundamentals, and are exposed to similar economic shocks. Importantly, the magnitude of the discipline effects using our measure of product market similarity is significantly larger than that of many other predictors of fraud documented in previous studies.

We next take several approaches to explore whether the disciplining effect of product similarity stems from the benchmarking or competition channels. The first approach uses variation in the size of shocks to a given firms information environment while controlling for the level of competition. Specifically, we exploit time-series changes in peer-firm similarity and conduct a two-stage least squares analysis only using the pair-wise similarities of rivals with newly available public disclosures (i.e., rivals undergoing an initial public offering). This framework mitigates concerns that our results are driven by firms' product market choices or pervasive differences in the rate of observed fraud across industries (e.g., Povel et al., 2007; Wang et al., 2010). First, rival IPOs present a discrete change in publicly-available information that is outside the underlying firm's sway. Second, we control for amount of capital raised and the number of IPOs, both of which could affect the intensity of competition. Thus, the remaining variation in competition should be limited as the rival firms already operate in the same product market prior to IPO. Indeed, we find strong corroborating evidence that the disciplining effect of product similarity operates through superior benchmarks that provide a better information environment.

The second approach uses a shock to competition while controlling for the information environment. We use large tariff reductions as a shock because these increase the intensity of foreign competition (Fresard, 2010) without affecting the quantity and quality of available public information. Our results show a persistent negative relation between product market similarity and fraud, which suggests the disciplining effect of product similarity operates through an information channel and affects the outcome of corporate fraud outside of changes to competition. Whereas we observe a small marginal pressure effect from tariff reductions, which is inconsistent with managerial rent extraction channel.

Next, we run a series of cross-sectional tests to provide more direct evidence that the observed results are driven by firms' information environment. The effect of benchmarking in disciplining managers should be stronger if the marginal value of public information increases, which in turn enhances the probability of getting caught committing fraud. On the margin, it is likely more challenging to detect abnormal behavior for complex firms operating across many segments than for firms with a simple organizational structure (e.g., Cohen and Lou, 2012). Consistent this notion, we find that the disciplining effect of product similarity increases monotonically in firm complexity. The economic magnitude is more than three times as large for the most complex firms when compared to the least complex firms.

The effect of benchmarking in disciplining firms also depends on the degree to which stakeholder monitoring relies on public disclosures. Product similarity is expected to have a stronger disciplining effect when monitors rely more on public information (e.g., financial statements), and vice versa. To explore this possibility, we consider two important corporate governance actors: institutional investors and sell-side analysts. Institutional investors often engage with management directly (McCahery et al., 2016) and have access to internal information to detect reporting anomalies which reduces the dependence on public signals; whereas analysts rely more on financial reports and industry trends to uncover misrepresentations (Shleifer and Vishny, 1986; Dyck et al., 2010). Our findings show a muted disciplining effect of product market similarity on fraud for firms with high institutional ownership and an enhanced effect for firms with greater analyst coverage. Furthermore, the effect of analyst coverage is driven by the post-Reg FD period after which analysts' reliance on publicly available information increased substantially. These findings further imply that benchmarking plays a greater role for monitors that rely more heavily on public information.

To more explicitly explore the competition channel, we control for a variety of widely-used measures of competition, such as the Herfindahl index. Prior literature has found mixed evidence on how the competitive landscape affects financial misconduct with some work documenting a positive association (i.e., pressure, see Karuna et al., 2012; Datta et al., 2013; Markarian and Santalo, 2014; Shi et al., 2018; Wang and Winton, 2021) and other work showing a negative association (i.e., discipline, see, Laksmana and Yang, 2014; Balakrishnan and Cohen, 2014). We find virtually no effect of traditional competition measures on our primary estimates on the relation between fraud and product market similarity and these competition measures do not have a significant impact on observing fraud in our sample. One explanation is that product market similarity provides information that is not only firm-specific but also from time-varying competitors and networks that are most relevant in assessing the landscape of product market competition. The ability to capture the intensity of the competitive relationships within-industry helps us to better control for unobserved heterogeneity and mitigate measurement error compared to traditional measures of competition that are coarse and at the industry level.

An empirical challenge in identifying predictors of fraud is that we only observe detected fraud rather than all committed fraud (see Dimmock and Gerken, 2012; Dyck et al., 2013). Thus, the negative relation between fraud and product market similarity can either be driven by reduced commission or diminished detection. Theory indicates that informational context provided by peers is sharpened by benchmarking (Tirole, 2010), which is consistent with empirical evidence that peer disclosures enhance board decisions (Murphy, 1986), improve analyst forecasts (De Franco et al., 2011), reduce audit opinion errors (Zhang, 2018) and earnings management behavior (Sohn, 2016). This evidence suggests that higher product similarity is likely to enhance fraud detection. Alternatively, it is less obvious how new information (i.e., the similarity score with IPO rivals) would make it more difficult on regulators and other stakeholders to detect fraud. The estimates from a bivariate probit model following Wang et al. (2010) corroborate the interpretation that our results are more consistent with reduced commission than with reduced detection.

Overall, our results support the notion that product market similarity provides external monitors with better benchmarks and an enhanced information environment, which then disciplines managers. Our findings add to the literature on external forms of corporate governance (e.g., Beasley, 1996; Farber, 2005; Khanna et al., 2015). Corporate governance mechanisms are endogenous responses to the costs and benefits of different internal governance mechanisms, as well as external monitoring from entities such as sell-side analysts, banks, or institutions (Gillan et al., 2011). Our paper offers evidence that product market competition is an important external governance force (Smith, 1776; Shleifer and Vishny, 1997) and can substitute for other formal corporate governance mechanisms (Giroud and Mueller, 2010; Chhaochharia et al., 2016). Our findings also complement those of Wang and Winton (2021), who show that industry-level information affects fraud detection. Our results indicate that information contained in firm-specific product market networks further enhances fraud detection beyond industry-level information. Our findings also reaffirm that peers are an avenue through which external parties can obtain information useful in the detection of fraud.

2 Data and Empirical Measures

2.1 Product Market Similarity

We use a measure of product market similarity developed by Hoberg and Phillips (2010, 2016) that captures the pairwise similarity of a firm's product market with each other firm that files a 10-K in a given year. Specifically, the authors calculate the cosine similarity, which measures the angle between two word vectors on a unit sphere, between two firms' business descriptions in their 10-K annual filings. The pairwise similarity is higher when the product market descriptions between the two firms exhibit greater overlap. The measure ranges from 0 (no similarity) to 1 (perfect similarity).

Peer firms are determined using the text-based network industry classification (TNIC) (Hoberg and Phillips, 2016), which includes all firms with a pairwise similarity above a given threshold.[1] The TNIC contains granular information that results in an industry classification scheme that is less constrained and more accurate than SIC or NAICS codes. One reason for the improved accuracy is that TNIC is updated annually. For example, when Exxon sold its retail gas stations in 2008, this event was reflected by the change in its competitor set (TNIC) and average pairwise similarity with all the competitors in its TNIC (from 0.035 to 0.012). However, the divestment resulted in no change to Exxon's SIC code.

In addition, most industry classifications are transitive and have no intensive margin. In contrast, TNIC is intransitive which allows for situations like Nike, Titlist and Addidas. Nike competes with Titlist in golf and Addidas in shoes but Titlist and Addidas do not compete with each other. This is useful in panel regressions because we can include industry fixed effects in firm-year level regressions. With respect to the intensive margin, the Coca-Cola Company and PepsiCo are not considered competitors according to their four-digit SIC code because PepsiCo has historically owned more restaurants. In contrast, they are TNIC competitors and have a high pairwise similarity score (80th percentile) as their main product, soda, is very similar.

We create our main variable of interest; *Product Market Similarity*, as the average pairwise similarity of all competitors within a firm's TNIC-3 classification in each year. As shown in

[1] Hoberg and Phillips (2016) only provide pairwise similarities above the threshold, c, where c is set such that the TNIC set of firms matches the same degree of coarseness as the three-digit SIC code classification scheme. In other words, a firm's peer group is of similar size when using either TNIC-3 or three-digit SIC code. We use TNIC-3 in this paper which Hoberg and Phillips (2016) define as firms with an *pairwise similarity* of above 0.03.

Table 1, firms in our sample have an average of 74 competitors.

One potential issue with this measure is that the number of competitors and the degree of similarity across competitors varies substantially. This issue could be a concern if one firm has several moderately close rivals while a different firm has some nearly identical rivals and some loosely related rivals that clear the threshold. While both firms could have the same average product similarity score, we expect near identical rivals to provide more precise signals regarding a firm's information environment. To address these concerns, we create alternative measures outlined in Section 3.3.4. All our constructs of product similarity are highly correlated (See Table A.1) and our results are robust with any of these constructions. For this reason, we use the simplest, the mean, in our main tests.

2.2 Financial Fraud

Consistent with prior work, we define corporate financial fraud as "the intentional, material misstatement of financial statements that causes damages to investors" and use Accounting and Auditing Enforcement Releases (AAER) misstatements and securities class action lawsuits as our main proxy(Donelson et al., 2021).[2] We focus on accounting misstatements because we are interested in whether product market similarity disciplines managerial actions that are intentionally misleading and damaging to investors.[3]

We obtain AAER data for the sample period 1996-2010 from the Center for Financial Reporting and Management. The U.S. Securities and Exchange Commission (SEC) issues

[2]Though Karpoff et al. (2017) note that each potential database on misconduct only encapsulates a portion of such events, we use a combination of public and private enforcement actions through AAER and class action lawsuits to increase the likelihood of capturing the most egregious and intentional misreporting. We also note the importance of additional checks of the sources such as whether or not the SCAC was dismissed to more accurately capture instances of fraud.

[3]Although some forms of financial fraud can be viewed as an extreme form of earnings management, Dechow and Skinner (2000) argue that fraud is more well defined and that earnings management is harder to measure empirically than fraud. There is a conceptual distinction between fraudulent accounting (that clearly demonstrates an intent to deceive) and judgments or estimates that fall within GAAP and which may even represent benign or justified forms of earnings smoothing that empirical measures of earnings management often cannot discern from more maligned motives.

AAERs during, or at the conclusion of, an investigation against a company, an auditor, or an officer for alleged accounting or auditing misconduct. The misstatement investigations in our sample occur because of bribery, fraud, inflated assets, financial reporting related enforcement actions concerning civil lawsuits brought in federal court, and orders concerning the institution and/or settlement of administrative proceedings.[4] We construct our sample of class action lawsuits following the work of Thompson and Sale (2003), Griffin et al. (2004), Choi et al. (2009) and Jayaraman and Milbourn (2009). We download all class action lawsuits from the Securities Class Action Clearing house (SCAC) for 1996 through 2010 and only include 10-b5 class action lawsuits, which eliminates lawsuits that occur for non-financial reasons.

Our primary dependent variable, *fraud*, is a binary variable equal to one for all firm years in which fraud occurred that was later disclosed via a AAER or SCAC. Because it takes time for lawsuits and investigations to occur and our fraud years are the years where the accounting fraud occurred, not where it was disclosed. We exclude financial firms and utilities, firms headquartered outside the United States, ADRs, REITs, firms with assets less than $1M, penny stocks and unit offerings, and firms with missing assets or sales items in Compustat. Our final sample of corporate fraud events contains 498 firms and 1,217 firm-years flagged as fraudulent from 1996 to 2010. These figures are closely in line with those of Dyck et al. (2010). As shown in Table 1, the incidence of fraud in our sample is 1.9%. 935 firm-years are affected by AAER misstatements from 322 unique firms and there are 311 class action lawsuits affecting 299 firms.[5]

[4]The AAER dataset provides information on the nature of the misconduct, the named individuals, and the entities involved, as well as their effect on the financial statements.

[5]All our results hold when we only include AAERs and ignore SCAC only years.

2.3 Additional Data and Control Variables

Our financial data comes from the Center for Research on Securities Prices (CRSP) and Compustat. In addition, we obtain data on 4488 IPOs for the period 1996 through 2010 from Thomson Reuters' SDC platinum financial securities database.

To construct our control variables, we follow recent literature related to fraud.[6] We include Richardson et al. (2005) (RSST) accruals, which measures the change in non-cash net operating assets, including both working capital accruals and long-term operating capital. Bergstresser and Philippon (2006) show that changes in accounts receivable (ΔAR) and inventory (ΔInventory) are associated with incentives to improve sales growth and gross profit margin. A firm's soft assets as a percentage of total assets (% soft assets), change in cash-based sales (ΔCash Sales) and nd change in return on assets (ΔROA) are associated with discretion for earnings management. ΔEmployees captures the labor costs and must be expensed and can be used to boost a firm's short-term financial performance. Finally, we include a dummy variable (Dummy Security Issue) equal to one for firm years in which a firm issues debt or equity, which can incentivize earnings management (Rangan, 1998). We refer to specifications including only the controls from Dechow et al. (2011) as the Dechow set of controls.

In addition, in some specifications we include institutional ownership, analyst coverage, research and development expenses (R&D),[7] industry stock return r-squared (Ind R2) following Wang and Winton (2021), firm size (Ln Assets), book leverage, and Whited and Wu (2006) Index for financial constraints.

We also exploit large industry-level tariff reductions and calculate industry tariff rates as duties collected by U.S. Customs divided by the value of U.S. imports. The duties and

[6] Detailed variable definitions are reported in Appendix 1, and all control variables are at year t.

[7] To handle observations with missing R&D, we follow the method outlined in Koh and Reeb (2015) and replace each missing observation with the industry year average and include a dummy variable for whether the firm has missing R&D (R&D dummy).

customs values are collected from the U.S. International Trade Commission. We aggregate the values from ten-digit U.S. Harmonized System codes to each three-digit SIC, using the concordance table provided by Pierce and Schott (2012), and indentify *Tariff shocks*.[8]

Relatedly, managerial incentives to commit fraud has been linked to managerial wealth (Armstrong et al., 2013), which is strongly linked to relative performance evaluation (RPE). Firms with higher similarity scores may be more likely to use RPE which could pressure managers to misstate earnings to outperform benchmarks (Cheng, 2011) and work against our hypothesis. Therefore, we include a measure of RPE in our specifications following Wang and Winton (2021), specifically the difference in performance (*return on assets*) between firm i and the weighted average across firm i's three-digit SIC code.[9]

Product market similarity is also related to accounting comparability, which is the degree to which companies' information is reported in a similar manner according to choices in accounting treatments (e.g., De Franco et al., 2011; Sohn, 2016; Zhang, 2018). We employ a commonly used output-based measure of accounting comparability developed by De Franco et al. (2011), which is based on the degree to which common economic events translate into similar financial statements.

Table 1 provides summary statistics for all variables and we present estimates obtained from the winsorized data (1% in each tail).

[8]We thank Chotibhak Jotikasthira for kindly sharing the methodology to calculate tariff shocks.

[9]We estimate the following equation separately for each three-digit SIC: $pr(\text{CEO Turnover}_{i,t-1}) = \gamma_1 RP^+_{i,t} + \gamma_2 RP^-_{i,t} + \epsilon_{i,t}$, where $RP^+_{i,t}$ ($RP^-_{i,t}$) is equal to relative performance when relative performance is above (below) 0, and zero otherwise. We define RPE equal to one for industries where $\hat{\gamma}_2 < 0$.

3 Empirical Results

3.1 Product Market Similarity and Financial Fraud

Our primary goal is to explore the relation between product market similarity and the incidence of corporate fraud. High product market similarity indicates that those firms operate in more homogeneous product markets in which rivals are subject to similar demand and cost conditions (Tirole, 1988). Consequently, the rivals of these firms likely provide sharper context to assess its economic circumstances. Furthermore, the pecuniary and reputational costs of being caught for such divergent behavior can be substantial for both managers and firms (Karpoff et al., 2008a, 2014; Armour et al., 2017). As such, we posit that managers exhibit more discipline in their reporting practices, i.e. commit less financial fraud, in response to stakeholders' and auditors' ability to benchmark financial disclosures against those of comparable firms with greater product market similarity. Hypothesis 1 is:

Hypothesis 1 *Ceteris paribus, product market similarity is negatively associated with financial fraud.*

The null hypothesis is that there is no effect or a positive effect. If product market similarity enhances the ability to detect fraud without an offsetting effect on managerial behavior, we would observe a positive association between product similarity and fraud. In addition, higher product similarity could be associated with higher incidences of fraud as it intensifies competition, which in turn can pressure managers to employ financial statements manipulations in a race for survival (see Shleifer, 2004) or in raising capital Hoberg and Lewis (2017) and result in firms exhibiting fraud risk factors Graham and Bedard (2003). If rival information has no systematic effect on managerial discipline or any discipline effect is offset by a positive effect, then we would expect to not see an effect on fraud.

Table 2 reports OLS results where we formally test Hypothesis 1. The unit of observation is firm-year, and the t-statistics are calculated from standard errors clustered by three-digit

SIC code.[10] Across all columns, *Product Market Similarity* exhibits a consistent, economically meaningful, and statistically significant mitigating effect on fraud. For example, the standardized regression in column 6 indicates that a one standard deviation increase in *Product Market Similarity* (0.023) is associated with a roughly 0.4 percentage point decline in the rate of fraud from 1.8% to 1.4%, a substantial economic effect.

Our results also suggest that product market similarity is an important predictor of fraud, even when controlling for predictors that can impact managerial incentives (see Dechow et al., 2011), as well as firm size, innovation, and corporate governance.[11] These results are robust to several different sample periods (i.e., before and after Sarbanes Oxley).[12]

By design, most variation is driven by differences in *Product Market Similarity* across firms within the same broadly defined industry, rather than through time. However, to ensure that our results are not driven by innovation or enforcement waves that influence both product similarity and fraud (Hoberg and Lewis, 2017; Wang et al., 2010), we further isolate the cross-sectional variation. First, we include industry×year fixed effects, which is a substantial improvement over prior studies in controlling for unobserved heterogeneity and pervasive industry-level differences in fraudulent activity. Second, we estimate separate regressions for each year of our sample using our main specification (column 4 of Table 2). We repeat this exercise using a 5-year rolling average *Product Market Similarity* as the independent variable. The coefficient estimates for each year, plotted in Figure 1, are all substantially negative[13] and are strongly consistent over time. This test shows that at any

[10] Our results are robust to probit and logit specifications (Table A.2), to clustering at various industry levels (e.g., two-digit SIC) and at the firm level, to controlling for F-score from Dechow et al. (2011) (Table A.3), to controlling for auditor quality (fees as in Markelevich and Rosner (2013) and a Big 4 Indicator, not reported) and size quartile splits (not reported).

[11] These control variables are defined in section 2.3 and Table 1. Including Institutional Ownership in our specifications induces a large drop in the number of observations (39,519 to 28,912) and does not appear to have a meaningful effect on fraud detection. Considering these issues, we refer to Column 4 as our primary specification in our remaining analyses, and referred as "Full" controls in the tables.

[12] For robustness, we also show that Product Market Similarity exhibits a strong discipline effect on fraud-driven restatements, but no effect when including all restatements, such as restatements driven by clerical errors or unintentional misapplications of Accounting Standards Updates. These results are reported in Table A.7 of the Appendix.

[13] Except for 2010 which is an abbreviated sample

point in our sample, firms with more similar rivals exhibit significantly lower fraud rates in the cross section, which strongly imply that managers' incentives to not materially mislead investors are systematically different in the cross-section.

3.2 Benchmarking Channel

3.2.1 Shocks to information

We propose that the negative relationship between product market similarity and financial fraud is driven by a benchmarking channel, that more closely related rivals enhance a firms information environment that facilitates more effective stakeholder monitoring. To attempt to isolate this benchmarking channel, we exploit a plausibly exogenous shock to firm's information environment. For each year, we focus on the product market similarity of only rivals issuing an IPO as opposed to the product market similarity all rivals in an two stage least squares (2SLS) framework. When a rival's disclosures become public for the first time, this provides a discrete change in publicly available information provided by rival firms. Specifically, these disclosures convey information that can bolster stakeholders' ability to assess a firm's economic circumstances (Holmstrom, 1982; Nalebuff and Stiglitz, 1983). Intuitively, this effect should be greater for two firms that serve markets that have greater overlap. If rivals that undergo an IPO on average have product similarities to a firm that are greater (lesser) than the focal firm's average score, rival IPOs provide a positive (negative) shock to overall product market similarity, which should increase (decrease) discipline on financial reporting.

Hypothesis 2 *Ceteris paribus, firm's product market similarity with new rival is negatively associated with financial fraud.*

This test enables us to exploit time-series within-industry heterogeneity that is likely outside of a firm's control, which mitigates the concerns that our baseline results are driven by pervasive differences in the rate of observed fraud across industries (see Povel et al., 2007; Wang et al., 2010).[14] For instance, CEOs that are more likely to commit fraud might select into product markets that are innovative and risky (i.e. more likely to be highly dissimilar). Alternatively, in mature industries, firms could have settled into well-differentiated positions and simultaneously commit less fraud for fear of damaging valuable brands or reputations.

We implement two-stage least squares regressions using the product market similarity between firm i and its rivals that issue an IPO in year t as a large time series change to firm i's overall product market similarity in year *t+1*. For each pairwise observation of competitors, i and j, we flag whether firm j issued an IPO in year t. We then create our instrumental variable, *Rival-IPO Similarity*, equal to firm i's product market similarity with all rival firms issuing an IPO in year t. Rivals already public in year t are excluded from this calculation. We also create an indicator variable ($CompetitorIPO$) that is equal to 1 if any of a firm's rivals underwent an IPO in year t, and 0 otherwise. On average, there are 3.2 rival IPOs per firm in our sample, with a median of 0 (43% of firms have at least one rival IPO). Contingent on having at least one IPO rival, each firm has an average of 7.6 rivals launching IPOs.

For this instrument to be valid, it must meet the relevance criterion, which requires that our instrumental variable exhibits a strong relation to a firm's change in ex post average score. Unless the product market similarity between firm i and its rivals conducting an IPO is exactly equal to the average similarity between firm i and its pre-existing public rivals, this activity will change firm i's average product market similarity. Thus, the relevance condition should hold by construction which it does as shown in in Column 1-3 of Table 3. These columns report a strong positive correlation between *Rival-IPO Similarity* and *Overall Product Market Similarity*. The positive sign indicates that rivals conducting IPOs that are more similar to firm i, increase firm i's ex post *Product Market Similarity* while rivals

[14] While our inclusion of industry and industry-year fixed effects should partially mitigate this concern, we cannot perfectly control for differences in industry characteristics related to fraud.

conducting IPOs that are less similar to firm i, decrease firm i's ex post *Product Market Similarity*. The smallest F-statistic in the first stage is 70.89 (8.42^2) and *Rival-IPO Similarity* does not appear to be a weak instrument.

The exclusion restriction requires that product similarity with rivals undergoing an IPO is unrelated to the focal firm's ex ante propensity to commit fraud. One potential concern is that IPOs have been shown to occur in waves (e.g. Lowry and Schwert, 2002), which could be endogenously related to industry opportunities and fraudulent behavior. However, we argue that a firm has little, if any, control regarding which rivals go public, the timing of that decision, and most importantly the incremental information produced. The variation we exploit is not whether there is an IPO (which is not random), but how similar the IPO rivals are to the focal firm. While all rival IPOs provides additional information, some increase information significantly more (increase average product market similarity post IPO) than others (decrease product market similarity). Importantly, prior to the IPO, rivals operate in the product market and impact the competitive landscape of the focal firm, however, provide significantly less information. To control for the change in the intensity of competition as a result of rival IPOs, we also include the number of rival IPOs, the total capital raised and industry fixed effects.

In Column 4-6 of Table 3, we report the second stage results of two-stage least squares regressions using *Rival-IPO Similarity* as an instrument for *Product Market Similarity*. We find strong corroborating evidence that the incidence of fraud is significantly lower for firms with greater product market similarity. In particular, the coefficient estimates range from 0.495-0.618 across all specifications, suggesting a consistent and economically meaningful effect. Importantly, these findings persist with the inclusion of industry fixed-effects in Column 6, which further helps to mitigate endogeneity concerns. To confirm our results are driven by information contained in the instrument, we repeat our analysis on a sample including only firm-years with at least one rival IPO and we find similar results reported in columns 4 and 8.

The coefficient estimates in the two-stage analysis are roughly twice as large as the OLS coefficients. The larger estimates could imply that omitted variables work against the effect documented in our initial analyses, and that the actual impact of information provided by close rivals on fraud is indeed larger than our initial estimates suggest. Alternatively, the larger coefficient estimates could be capturing a local average treatment effect where the partial effects are concentrated in firms with more rival IPO activity.

While we cannot entirely rule out the potential for omitted variables to jointly determine a firm's fraudulent reporting and the product market similarity with rivals issuing IPOs, the two-stage results are suggestive of a causal relationship between product market similarity and fraud. In addition, we believe this is a useful setting in suggesting our results operate through an information channel more than a competition one.

3.2.2 Shocks to competition

One alternative explanation for our negative relationship between product market similarity and fraud is that the disciplining effect works by facilitating investors control of the agency problem by reducing managerial slack (Machlup, 1967), rather than through the benchmarking channel. For example, competition can cause managers to exert more effort by diminishing the benefits of shirking Hart (1983) and reduce resources available for rent extraction (e.g., Nickell, 1996; Ades and Di Tella, 1999). Extending this concept to corporate fraud, competition potentially reduces the economic profits that may be extracted through reporting manipulations.[15] Furthermore, the availability of product market substitutes offered by rivals may exacerbate lost market share due to the reputational costs of fraud. We refer to this channel of product market discipline as the managerial slack effect.

To explore whether the disciplining effect of product market similarity is driven by variation in managerial slack, we exploit large tariff reductions at the industry level. Tariff re-

[15] For instance, competition can mitigate the benefits of earnings manipulations in order to maintain higher valuations during acquisition activity or capital raising (Shleifer, 2004).

ductions have been shown to increase the intensity of foreign competition (Fresard (2010)), which can ultimately decrease managerial slack (Hart, 1983). While firms could respond to changes in foreign competition in the long run by adjusting their product mix, in the short-term, changes in tariff rates directly affect foreign rivals' ability of to offer competitive prices without directly affect the quality or quantity of readily available information through financial disclosures in 10-Ks. Thus, tariffs provide a good setting to analyze the intermediate effects of changes in competition that are likely independent of the benchmarking channels. Our third hypothesis is

Hypothesis 3 *Ceteris paribus, the negative relationship between firm's product market similarity and financial fraud is persistent in spite of changes in competitive environment without accompanied changes in information available for benchmarking.*

Following the literature, industry tariff rates are calculated as duties collected by U.S. Customs divided by the value of U.S. imports for consumption. *Tariff shock* is an indicator variable that takes value of 1 if the 4-year percentage change in tariff rate is the bottom (tercile/quartile/quintile), and 0 otherwise. Table 4 presents the tariff results for each specification with and without *Product Market Similarity* included as a control. Although the results is suggestive of a mild pressure effect from the increase in foreign competition, the coefficient is only statistically significantly different than zero in Column 3.

More importantly, the coefficient estimates for *Product Market Similarity* are highly consistent with the magnitude found throughout our other analysis. This finding suggests that product market similarity operates through a different channel than tariff reductions and works in the opposite direction. Thus, it does not appear that our results are driven by changes to managerial slack.

3.2.3 Marginal Value of Public Information: Complexity

If product market similarity disciplines managers by enhancing informational context, this effect should also depend on other characteristics of a firm's information environment. As such, our remaining hypotheses concern cross-sectional predictions related to information regarding the effect of product market similarity on financial fraud. Specifically, if product market similarity is operating through a information channel, we should see the impact of product market similarity be greater if a firm operates in a more opaque environment as the marginal value of public product market information is higher.

We first examine the relation between firm complexity and the disciplining effects of product market similarity. Bushman et al. (2004) show that firm complexity, measured via product line diversification, is associated with more costly governance, suggesting that complex firms are more difficult to monitor. Similarly, Cohen and Lou (2012) argue that a firm with multiple operating segments requires more complicated analysis to impound information into its share price. Therefore, one expects that managers of complex firms can more easily obfuscate information, imposing challenges to evaluating financial reporting behavior and have greater ability to conceal financial information. Peterson (2012) and Hoitash and Hoitash (2017) find consistent empirical evidence that firm complexity is associated with a greater likelihood of accounting misstatements. However, if product market similarity operates through the benchmarking channel, it can substantially reduce the cost of assimilating information for complex firms. Therefore, all else equal, the disciplining effect of product market similarity should be stronger for complex firms where information should have a larger marginal effect.[16] Hypothesis 4 is:

Hypothesis 4 *Ceteris paribus, product market similarity has a stronger negative association with financial fraud for complex firms.*

[16] Note that, if complex firms have fewer readily available benchmarks or their scope for manipulating financial reporting could be sufficient to evade the effects of comparable peers without necessarily producing red flags, we might not observe significant marginal discipline effect.

We define complexity as the number of unique industries (three-digit SIC codes) in which a firm operates each year. To calculate this value, we sum the number of distinct industries spanned by a firms TNIC-based competitor set. For example, if a firm has three rivals that each operate in a different three-digit SIC code, then we consider that firm to be operating in three distinct product markets.[17] We split our sample into quartiles according to complexity, and then estimate our main specification for the relation between fraud and *Product Market Similarity* separately for each quartile in Table 5.

In Panel A, we report the average number of unique SIC codes and the number of competitors in each firms TNIC. Each specification is estimated using our main specification. We estimate regressions separately for each complexity quartile in Panel B. The disciplining effect of *Product Market Similarity* increases monotonically across complexity quartiles for Panel B. The partial effect for the top quartile is more than four times as large as that for the lowest quartile. To put this finding into perspective, a one standard deviation increase in *Product Market Similarity* for the least complex firms leads to a decrease in propensity of fraud from 1.6% to approximately 1.3%, or a 19% percent decline. By comparison, a one standard deviation change in *Product Market Similarity* for the most complex firms decreases the propensity of fraud from 2.1% to 0.85%, or a 60% percent decline. While large firms are more complex, on average, complexity appears to be distinct from firm size.

Running regressions separately for each quartile does not constrain coefficients on the control variables to be the same across quartiles, ensuring the greatest degree of flexibility. We also estimate specifications with an interaction term and find corroborating results in Column 5 of Panel B. Overall, the monotonic relationship between product market similarity and fraud as firm complexity increases is strongly supportive of Hypothesis 4. Specifically, while simple firms with close rivals are less likely to commit fraud, the effect on fraud rates of having a close rival is over three times larger for complex firms.

[17]Our measure of complexity builds on the intuition provided by Cohen and Lou (2012) and Bushman et al. (2004) who also define firm complexity using the number of segments in which a firm operates.

3.2.4 Marginal Value of Public Information: External Monitors

Next we consider the marginal impact of product market similarity on the monitoring abilities of two particularly important corporate governance actors that have varying degrees of reliance on public financial statements: institutional investors and sell-side analysts.

The negative effects of corporate fraud can be particularly costly to investors with large ownership stakes (i.e., institutional investors), thus creating strong monitoring incentives (Shleifer and Vishny, 1986). Furthermore, higher ownership stakes by institutional investors can facilitate the access to managers and internal information relevant to fraud (Piotroski and Roulstone, 2004; Bushee and Goodman, 2007; McCahery et al., 2016). This access to private information reduces their reliance on public signals of firm performance; therefore, the disciplining effect of product market similarity based on publicly available financial statements should have a muted effect for firms with highly concentrated institutional ownership which we capture in our next hypothesis.

Hypothesis 5 *Ceteris paribus, product market similarity has a weaker negative association with financial fraud for firms with a greater concentration of institutional ownership.*

Sell-side analysts also play an important governance role in monitoring firms and detecting fraud. For instance, Dyck et al. (2010) find that analysts and auditors uncover roughly 24% of fraud and Chen et al. (2015) find that analysts play a meaningful role in scrutinizing management behavior. While institutional investors potentially have access to fraud-relevant inside information, "analysts ... [have] to delve through details of financial reports and industry trends to uncover misrepresentations" (Dyck et al., 2010). Rivals with high product market similarity can provide context through which analysts benchmark financial disclosures to better detect abnormal reporting behavior and increase the marginal disciplining effect for firms with more analyst coverage. Moreover, Regulation Fair Disclosure (Reg FD) increased analysts' reliance on public information by reducing access to private information.[18] As such, the enhanced disciplining effect of product market similarity for firms with greater analyst coverage should be concentrated in the post-Reg FD period. We capture these potential effects in the following hypotheses.

Hypothesis 6a *Ceteris paribus, product market similarity has a stronger negative association with financial fraud for firms with greater analyst coverage.*

Hypothesis 6b *Ceteris paribus, the association between product market similarity, financial fraud, and analyst coverage is stronger post Reg FD.*

Following Hartzell and Starks (2003), we proxy for institutional ownership influence and monitoring intensity using a Herfindahl Index (HHI) of institutional ownership as they show the concentration of institutional ownership, rather than the level, determines the intensity of institutional monitoring. To explore the differential effect of product market similarity, we estimate our primary specifications for subsets of our sample split into quartiles based on Institutional Ownership HHI. As shown in Panel A of Table 6, the partial effect of *Product Market Similarity* on fraud is much higher for firms with dispersed ownership than it is for firms with concentrated ownership. For instance, for firms with the least concentrated ownership structure (i.e., bottom quartile in IO-HHI in column 1), one standard deviation increase in product market similarity is associated with a 0.97% decrease in fraud, which is approximately 51.4% drop compared to the unconditional mean. In contrast, column 4 reports insignificant discipline effects for concentrated firms from product market similarity. In Column 7 of Panel A, we report estimates of a specification that includes an interaction term, which reaffirms the quartile-split results. These findings suggest that the disciplining effect of product market similarity is less significant for firms with highly concentrated institutional owners, who have greater access to fraud-relevant information and can discipline firms without relying on benchmarks.

[18] According to Investor.gov from the SEC, "Regulation FD addresses the selective disclosure of information by publicly traded companies and other issuers. Regulation FD provides that when an issuer discloses material nonpublic information to certain individuals or entitiesgenerally, securities market professionals, such as stock analysts, or holders of the issuer's securities who may well trade on the basis of the informationthe issuer must also make public disclosure of that information".

In Panel B of Table 6, we present results for samples split by quartiles based on the number of analysts covering the firm, and by whether a firm is covered by at least one star analyst. Higher analyst coverage or the presence of star analysts should be associated with an increased degree of scrutiny over a firm's financial statements which increases the value of readily available benchmarks.[19] Indeed, we find that the partial effect of *Product Market Similarity* on fraud is significantly larger for firms with greater analyst coverage (or at least one star analyst). For the firms with the most analyst coverage (i.e., top quartile in column 4), a one standard deviation increase in product market similarity is associated with a 1.1% drop in fraud rate, or equivalent to 58.9% decrease compared to the unconditional mean; and we do not find any effect for the firms with lowest analyst coverage (i.e. bottom quartile in column 1).

In Panels C and D of Table 6, we split the samples before and after Regulation Fair Disclosure (Reg FD). Prior to Reg FD, analysts likely had better access to non-public information, but since its implementation have to rely more on public information. Consistent with this notion, the results show that there is a stronger differential impact across different degrees of analyst scrutiny after Reg FD.[20] Finally, in Column 7 of each panel, we estimate specifications that include interaction terms. The coefficient estimates on the interaction terms reaffirm the quartile-split results and they follows a similar pattern across panels. Overall, these findings offer support of our conjectures and increase our confidence that the effect is driven by reliance on public signals provided by stronger benchmarks.

We interpret the disparate impact of institutional investors and analysts on the partial effect of product market similarity on fraud as being driven by differential access to information. To reinforce this assertion, we use data from (Dyck et al., 2010) regarding who blows the whistle on corporate fraud to discern whether each case was caught internally (e.g. firm

[19]Star analysts are defined by Institutional Investor All-America Research Team designations.

[20]We test whether the coefficients are statistically different across quartiles in Table 6 and report the results in Table A.8. With the exception of the Star/Non-Star split in Panel C, all coefficient estimates are statistically different across sub-samples.

employees or auditors) or externally (e.g. regulators, analysts, or the media). In untabulated results, we find that firms that are caught by external whistle blowers have a significantly higher *Product Market Similarity* than firms caught by internal whistle blowers. This finding is also consistent with external monitors relying more on public information.

3.3 Robustness and Additional Analyses

3.3.1 Financial Statement Comparability

Product market similarity is closely related to accounting comparability, which is the degree to which companies' information is reported in a similar manner according to choices in accounting treatments. Accounting comparability lowers the cost of acquiring information and also increases the ability to benchmark reported financial information and should enhance the benchmarking effect of product market similarity. Accordingly, we estimate the partial effect of *Product Market Similarity* for firms while controlling for the degrees of accounting comparability.

Specifically, we use an output-based measure of accounting comparability based on the degree to which common economic events translate into similar financial statements following De Franco et al. (2011). Results are reported in Table 7 for *Product Market Similarity* and fraud for firms with different levels of *Accounting Comparability*. Panel A indicates that the correlation between *Accounting Comparability* and *Product Market Similarity* is low (-1.5%). Panel B reports regression results using our main specification. Columns 1-4 include Year and industry fixed effects while Columns 5-7 include industry×year fixed effects. Column 1 contains observations where *Accounting Comparability* is missing. In Columns 2-3 (5-6) we split the data between low (below median) and high (above median) *Accounting Comparability*. In column 4 and 7 we include an interaction term between *Product Market Similarity* and *Accounting Comparability*.

The results suggest a strong reinforcement effect between Product Market Similarity and accounting comparability. In particular, the disciplining effect imposed by Product Market Similarity is isolated to firms with more comparable accounting statements. The interaction effect reaffirms this association. These findings support the notion that accounting comparability and product market similarity collectively enhance external monitoring through peer firm comparisons.

3.3.2 Controlling for the Confounding Effects of Competition

Product market similarity can also operate through concomitant effects of competition as classic models of competition suggest that competition becomes more intense as more firms offer marketable substitutes (Tirole, 1988). Competition can pressure managers to commit more fraud in order to survive (which is inconsistent with our main results) or it could reduce profits that can be extracted through reporting manipulations mitigating the benefits of earnings manipulations to aid acquisition activity or capital raising (Shleifer, 2004).

To account for potential confounding effects of competition we control for a wide array of competition measures widely utilized in prior literature, such as HHI (Hirschman, 1945; Herfindahl, 1950), profit margin (Bain, 1951), the sales concentration ratio of the largest four firms in an industry (Heflebower, 1957), and the number of competitors in a firm's TNIC or three-digit SIC code. To exploit more accurately defined product markets, we also create a TNIC-based HHI by weighting the sales of each rival by the firm's product market similarity with that rival. We also include newer measures of competition such as product market fluidity, which captures competitive pressure from potential entrants that captures each firms ex ante competitive threats (Hoberg et al., 2014). Finally, we use a firm-level measure of the frequency that firms discuss competition in their 10-Ks, which captures the degree of competitive threats perceived by the focal firm (see Li et al., 2013).[21]

[21] Table A.1 reveals a weak correlation between proxies for competition and *Product Market Similarity*.

Each specification reported in Table 8 uses our main specification along with several combinations of proxies for competition. Notably, the relation between *Product Market Similarity* and fraud remains consistent in both significance and magnitude across all columns.[22] Note that our inclusion of industry×year fixed effects allows controlling for a high level of unobserved heterogeneity is not possible under common definitions of competition. In unreported results we estimate the relation between fraud and proxies for competition, excluding *Product Market Similarity* and fail to find a robust link between competition proxies and fraud. The apparent lack of power exhibited by these measures of competition highlight one potential reason for a lack of consistent evidence between product market characteristics and corporate fraud in the literature.

We also exploit vertically related firms that can provide economic context for upstream and downstream firms, but are less confounded by other competitive forces. For instance, auto-manufacturing is heavily dependent on upstream steel and aluminum industries, and it could send up a red flag if an auto-manufacturer's major steel or aluminum supplier experiences a substantial surge in the cost of goods sold, but this change does not translate into an increase in the firm's own reported cost of production. We use the Vertical Text-based Network Industry Classifications (VTNIC) developed by Frésard et al. (2020), who combine vocabulary in 10-Ks and vocabulary from the Bureau of Economic Analysis's (BEA) input/output tables to identify "the extent to which two firms operate in vertically-related product markets.".

We present results with *V-Similarity* as the primary independent variable in Table A.5. Overall, the results, while weaker in magnitude, reaffirm the findings that information from related firms can discipline managers. For example, a one standard deviation increase in *V-Similarity* leads to a 7.9-18.8% decrease in accounting fraud, relative to the unconditional mean.

[22]We explore several alternative combinations of covariates and different sample periods and find that these results are not sensitive to model specification.

3.3.3 Fraud Detection vs. Commission

One concern is that empirical measures of fraud capture the joint outcome of a firm committing fraud and being caught, making it difficult to interpret whether estimated effects are due to changes in detection or commission rates (Dimmock and Gerken, 2012; Dyck et al., 2013). Our findings suggest that product market similarity either reduces commission or lowers detection. However, it is difficult to ascertain a plausible explanation for why product market similarity decreases outsiders' ability to detect reporting manipulations. Indeed, similar rivals should inform about economic circumstances, but, at worst, would provide no incremental information (e.g. Tirole, 2010).

To further substantiate our interpretation, we estimate a bivariate probit model employed by Wang (2011). The bivariate probit model is a latent variable model that exploits timing differences between detection and commission, with commission necessarily occurring prior to detection. The model solves two simultaneous probit specifications (one for commission and one for detection) and achieves identification through exclusion restrictions: namely, that some variables are only associated with detection, while others are only associated with commission. Following Wang (2011) we include RPE, ROA, Equity Finance Needed, Book Leverage, and Institutional Ownership, only in the commission regression and Abnormal Industry Litigation, Abnormal Stock Return Volatility, Abnormal Turnover, and a Disastrous Return indicator, only in the detection regression. All other controls are included in both regressions. In Table A.6 of the Internet Appendix, we report estimates from the bivariate probit model that suggest product market similarity is strongly associated with a decline in fraud commission and weakly related to enhanced detection.

3.3.4 Alternate Constructions of Product Market Similarity

Lastly, our primary construction of *Product Market Similarity* is an equally weighted average that aggregates pairwise similarity across each of a firm's rivals. We implement alternative constructions to ensure that our results are not driven by any particular features of our main construct. First, we average pairwise similarity across only top 5, 10 or 15 closest competitors, which focuses on each firm's closest rivals that are most likely to provide the greatest information externalities. Second, we also calculate a sum of pairwise comparability scores, which is higher when a firm has more rivals and rivals that are more similar. Finally, we develop a measure, *Precision*. If the noise from the signal generated by a rival benchmark is normally distributed, then there is an inverse squared relationship between product market similarity and the quality of the signal. We define *Precision* as:

$$precision_{i,t} = \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \frac{1}{(1 - score_{i,j,t})^2} \right)^{0.5} \tag{1}$$

where N_i is the number of competitors in firm i's TNIC, and $score_{i,j,t}$ is the product similarity score between firm i and competitor j in year t.[23]

In Panel A of Table 9, each alternative construction yields similar results to our main analysis, which mitigates concerns that aggregation obscures the association between fraud and product market similarity.[24] To facilitate comparison of economic magnitudes, we report results from standardized regressions in Panel B. The coefficient estimates exhibit some monotonicity with the closest (i.e., highest similarity) rivals yielding the greatest effect.

4 Conclusion

When firms manipulate financial reporting, they undermine the trust that is essential to a well-functioning financial system (Greenspan, 2008). We show that by enhancing the external information environment and facilitating stakeholder monitoring, product market similarity disciplines managers, i.e., reduces the incidence of financial fraud. Our findings

[23] We thank Gerard Hoberg for suggesting this measure.

[24] Correlations between these alternative constructions are presented in Table A.1 of the Internet Appendix.

help shed light on circumstances that effect managers' incentives to commit fraud and reveal that market-based mechanisms can minimize those incentives and improve the ability of various stakeholders to detect such behavior.

We also show the discipline effect from product market similarity is economically larger than that of many other documented predictors of fraud and significant in all time periods in our sample. This finding contributes to the extensive literature on financial fraud by proposing another important predictor to include in fraud regressions in order to reduce omitted variable concerns and increase power.

Finally, because regulators must allocate limited resources available for fraud detection, our findings offer them some real world suggestions. Consider focusing less on firms that likely face more market discipline, i.e, firms that have similar product market rivals.

Appendix: Variable Descriptions

Variable	Definitions
AAER Misstatement	Equal to 1 for firm-years for which firms have settled with the SEC for corporate Fraud. Note: This is not the actual settlement year, which is usually several years after the alleged fraud, but the year in which the fraud allegedly occurred.
SCAC	Securities and Class Action Equal to 1 for all firm-years for which firms settle a securities class action lawsuit for an alleged 10B-5 fraud allegation.
Fraud	Equal to 1 for all firm-years with an AAER or SCA.
SIC3 HHI	Herfindahl-Hirschman index based on firm sales and three-digit SIC code industry classifications
TNIC HHI	Herfindahl-Hirschman index based on firm sales Text-based Network Industry classifications (TNIC) from Hoberg and Phillips.
Product Market Similarity (PMS)	Mean Hoberg and Phillips Similarity Score for all rivals within each firm-years TNIC
Precision	Defined as $(\frac{1}{NCOMP_{TNIC}} \times \sum \frac{1}{(1-score)^2})^{\frac{1}{2}}$
Profit Margin	Average EBITDA/sales ratio for firms within each three-digit SIC code
Top 4 Concentration	Proportion of sales within a three-digit SIC code attributable to the four largest firms within an industry
Age	Number of years the firm has been in Compustat
Analyst Num	Number of analysts covering the firm in each year from IBES (0 if missing).
Inst Ownership	Percentage of shares outstanding held by 13-F institutions
Assets	Total Assets
Capex	Capital Expenditures / log Assets
Book Leverage	(Total Long-Term Debt +Debt in Current Liabilities)/ log Assets
ROA	Net Income / Assets
EFN	Equity Finance Needed defined as ROA/(1 - ROA).
RSST Accruals	$(NOA_t - NOA_t - 1)/NOA_t - 1$. NOA (Net Operating Assets) = OA- OL where OA (Operating Assets) = sum of COA (current operating assets) and NCOA (noncurrent operating assets) and OL = sum of COL (current operating liabilities) and NCOL (noncurrent liabilities). COA = Current Assets - Cash and Short-Term Equivalents. NCOA = Total Assets - Current Assets - Investments and Advances. COL = Current Liabilities - Debt in Current Liabilities. NCOL = Total Liabilities - Current Liabilities - Long-Term Debt
Security Issuance	An indicator variable equal to 1 if the firm issued securities during year
Change AR	Change in Accounts Receivable/Total Assets
Change Inventory	Change in Inventory/Total Assets
Pct Soft Assets	(Total Assets - PP&E- Cash and Cash Equivalent)/Total Assets
Change in Cash Sales	Percentage change in Cash Sales - Change in Accounts Receivable
Change in ROA	Change in Return on Assets
Change in Employee	Percentage change in the number of employees - percentage change in assets
R&D	Research and Development scaled by assets. Missing observations are filled with either the firm average, if a time series exists, or the industry average if not.
R&D (indicator)	Equal to 1 if R&D is missing and 0 otherwise. We follow Koh and Reeb (2015) when using R&D.
NCOM SIC3	Number of competitors within the three-digit SIC Code.
Ind R2	Following Wang and Winton (2021), we first regress each firms daily stock returns on the weighted-average daily market return and the weighted-average daily industry return. Then, we take the average r-squared for each firm in each three-digit SIC code.
RPE Flag	See Page 16
NCOMP TNIC	Number of competitors according to Text-based Network Industry classifications (TNIC) from Hoberg and Phillips.

References

Ades, A., and R. Di Tella. 1999. Rents, Competition, and Corruption. *The American Economic Review* 89:982–993.

Armour, J., C. Mayer, and A. Polo. 2017. Regulatory sanctions and reputational damage in financial markets. *Journal of Financial and Quantitative Analysis* 52:1429–1448.

Armstrong, C. S., D. F. Larcker, G. Ormazabal, and D. J. Taylor. 2013. The relation between equity incentives and misreporting: The role of risk-taking incentives. *Journal of Financial Economics* 109:327 – 350.

Badertscher, B., N. Shroff, and H. D. White. 2013. Externalities of Public Firm Presence: Evidence from Private Firms Investment Decisions. *The Journal of Financial Economics* 109:682–706.

Bain, J. S. 1951. Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940. *The Quarterly Journal of Economics* 65:293–324.

Balakrishnan, K., and D. A. Cohen. 2014. Competition and Financial Accounting Misreporting. *Working Paper* .

Beasley, M. S. 1996. An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud. *The Accounting Review* 71:443–465.

Bergstresser, D., and T. Philippon. 2006. CEO incentives and earnings management. *Journal of Financial Economics* 80:511–529.

Bushee, B. J., and T. H. Goodman. 2007. Which institutional investors trade based on private information about earnings and returns? *Journal of Accounting Research* 45:289–321.

Bushman, R., Q. Chen, E. Engel, and A. Smith. 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics* 37:167–201.

Chen, T., J. Harford, and C. Lin. 2015. Do analysts matter for governance? Evidence from natural experiments. *Journal of Financial Economics* 115:383 – 410.

Cheng, I.-H. 2011. Corporate Governance Spillovers. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Chhaochharia, V., Y. Grinstein, G. Grullon, and R. Michaely. 2016. Product Market Competition and Internal Governance: Evidence from the Sarbanes – Oxley Act. *Management Science* 63:1405–1424.

Choi, S. J., K. K. Nelson, and A. C. Pritchard. 2009. The Screening Effect of the Private Securities Litigation Reform Act. *Journal of Empirical Legal Studies* 6:35–68.

Cohen, L., and D. Lou. 2012. Complicated firms. *Journal of Financial Economics* 104:383–400.

Datta, S., M. Iskandar-Datta, and V. Singh. 2013. Product market power, industry structure, and corporate earnings management. *Journal of Banking and Finance* 37:3273 – 3285.

De Franco, G., S. P. Kothari, and R. S. Verdi. 2011. The Benefits of Financial Statement Comparability. *Journal of Accounting Research* 49:895–931.

Dechow, P. M., W. Ge, C. R. Larson, and R. G. Sloan. 2011. Predicting Material Accounting Misstatements. *Contemporary Accounting Research* 28:17–82.

Dechow, P. M., and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting horizons* 14:235–250.

Dichev, I. D., J. R. Graham, C. R. Harvey, and S. Rajgopal. 2013. Earnings quality: Evidence from the field. *Journal of Accounting and Economics* 56:1–33.

Dimmock, S. G., and W. C. Gerken. 2012. Predicting fraud by investment managers. *Journal of Financial Economics* 105:153–173.

Donelson, D., A. Kartapanis, J. McInnis, and C. Yust. 2021. Measuring Accounting Fraud and Irregularities Using Public and Private Enforcement. *The Accounting Review* .

Dyck, A., A. Morse, and L. Zingales. 2010. Who Blows the Whistle on Corporate Fraud? *The Journal of Finance* 65:2213–2253.

Dyck, I. J. A., A. Morse, and L. Zingales. 2013. How Pervasive is Corporate Fraud? *Working Paper* .

Engelberg, J., R. D. McLean, and J. Pontiff. 2018. Analysts and Anomalies. SSRN Scholarly Paper ID 2939174, Social Science Research Network, Rochester, NY.

Farber, D. B. 2005. Restoring Trust after Fraud: Does Corporate Governance Matter? *The Accounting Review* 80:539–561.

Fresard, L. 2010. Financial Strength and Product Market Behavior: The Real Effects of Corporate Cash Holdings. *The Journal of Finance* 65:1097–1122.

Frésard, L., G. Hoberg, and G. M. Phillips. 2020. Innovation Activities and Integration through Vertical Acquisitions. *Review of Financial Studies* .

Giannetti, M., and T. Y. Wang. 2016. Corporate Scandals and Household Stock Market Participation. *The Journal of Finance* 71:2591–2636.

Gillan, S. L., J. C. Hartzell, and L. T. Starks. 2011. Tradeoffs in Corporate Governance: Evidence From Board Structures and Charter Provisions. *Quarterly Journal of Finance* 01:667–705.

Giroud, X., and H. M. Mueller. 2010. Does corporate governance matter in competitive industries? *Journal of Financial Economics* 95:312–331.

Graham, L., and J. C. Bedard. 2003. Fraud Risk and Audit Planning. *International Journal of Auditing* 7:55–70.

Greenspan, A. 2008. *The Age of Turbulence: Adventures in a New World*. Penguin.

Griffin, P. A., J. A. Grundfest, and M. A. Perino. 2004. Stock Price Response to News of Securities Fraud Litigation: An Analysis of Sequential and Conditional Information. *Abacus* 40:21–48.

Gurun, U. G., N. Stoffman, and S. E. Yonker. 2018. Trust Busting: The Effect of Fraud on Investor Behavior. *The Review of Financial Studies* 31:1341–1376.

Hart, O. D. 1983. The Market Mechanism as an Incentive Scheme. *The Bell Journal of Economics* 14:366–382.

Hartzell, J. C., and L. T. Starks. 2003. Institutional Investors and Executive Compensation. *The Journal of Finance* 58:2351–2374.

Heflebower, R. B. 1957. Barriers to New Competition. *The American Economic Review* 47:363–371.

Herfindahl, O. 1950. *Concentration in the steel industry*. Columbia Univ., Ph. D. thesis. Columbia University.

Hirschman, A. O. 1945. *National Power and the Structure of Foreign Trade*. University of California Press. Google-Books-ID: BezqxPq50dwC.

Hoberg, G., and C. Lewis. 2017. Do fraudulent firms produce abnormal disclosure? *Journal of Corporate Finance* 43:58–85.

Hoberg, G., and G. Phillips. 2010. Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis. *The Review of Financial Studies* 23:3773–3811.

Hoberg, G., and G. Phillips. 2016. Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy* 124:1423–1465.

Hoberg, G., G. Phillips, and N. Prabhala. 2014. Product Market Threats, Payouts, and Financial Flexibility. *The Journal of Finance* 69:293–324.

Hoitash, R., and U. Hoitash. 2017. Measuring accounting reporting complexity with XBRL. *The Accounting Review* 93:259–287.

Holmstrom, B. 1982. Moral Hazard in Teams. *The Bell Journal of Economics* 13:324–340.

Hsu, C., X. Li, Z. Ma, and G. M. Phillips. 2017. Does Product Market Competition Influence Analyst Coverage and Analyst Career Success? *Working Paper* .

Jayaraman, S., and T. Milbourn. 2009. Does equity-based CEO compensation really increase litigation risk. *Working Paper* .

Karpoff, J. M., A. Koester, D. S. Lee, and G. S. Martin. 2017. Proxies and databases in financial misconduct research. *The Accounting Review* 92:129–163.

Karpoff, J. M., D. S. Lee, and G. S. Martin. 2008a. The consequences to managers for cooking the books. *Journal of Financial Economics* 88:193–215.

Karpoff, J. M., D. S. Lee, and G. S. Martin. 2008b. The Cost to Firms of Cooking the Books. *The Journal of Financial and Quantitative Analysis* 43:581–611.

Karpoff, J. M., D. S. Lee, and G. S. Martin. 2008c. The cost to firms of cooking the books. *Journal of financial and quantitative analysis* 43:581–611.

Karpoff, J. M., D. S. Lee, and G. S. Martin. 2014. The consequences to managers for financial misrepresentation. In *Accounting and Regulation*, pp. 339–375. Springer.

Karuna, C., K. Subramanyam, and F. Tian. 2012. Industry product market competition and earnings management. In *American Accounting Association Financial Accounting and Reporting Section Mid-Year Conference*.

Kedia, S., and T. Philippon. 2009. The Economics of Fraudulent Accounting. *The Review of Financial Studies* 22:2169–2199.

Khanna, V., E. H. Kim, and Y. Lu. 2015. CEO Connectedness and Corporate Fraud. *The Journal of Finance* 70:1203–1252.

Kim, S., P. Kraft, and S. G. Ryan. 2013. Financial statement comparability and credit risk. *Review of Accounting Studies* 18:783–823.

Koh, P.-S., and D. M. Reeb. 2015. Missing R&D. *Journal of Accounting and Economics* 60:73–94.

Laksmana, I., and Y.-W. Yang. 2014. Product market competition and earnings management: Evidence from discretionary accruals and real activity manipulation. *Advances in Accounting* 30:263 – 275.

Li, F., R. Lundholm, and M. Minnis. 2013. A Measure of Competition Based on 10-K Filings. *Journal of Accounting Research* 51:399–436.

Lowry, M., and G. W. Schwert. 2002. IPO Market Cycles: Bubbles or Sequential Learning? *The Journal of Finance* 57:1171–1200.

Machlup, F. 1967. Theories of the Firm: Marginalist, Behavioral, Managerial. *The American Economic Review* 57:1–33.

Markarian, G., and J. Santalo. 2014. Product Market Competition, Information and Earnings Management. *Journal of Business Finance & Accounting* 41:572–599.

Markelevich, A., and R. L. Rosner. 2013. Auditor Fees and Fraud Firms. *Contemporary Accounting Research* 30:1590–1625.

McCahery, J. A., Z. Sautner, and L. T. Starks. 2016. Behind the Scenes: The Corporate Governance Preferences of Institutional Investors. *The Journal of Finance* 71:2905–2932.

Murphy, K. J. 1986. Incentives, Learning, and Compensation: A Theoretical and Empirical Investigation of Managerial Labor Contracts. *The RAND Journal of Economics* 17:59–76.

Nalebuff, B. J., and J. E. Stiglitz. 1983. Prizes and Incentives: Towards a General Theory of Compensation and Competition. *The Bell Journal of Economics* 14:21–43.

Nickell, S. J. 1996. Competition and Corporate Performance. *Journal of Political Economy* 104:724–746.

Peterson, K. 2012. Accounting complexity, misreporting, and the consequences of misreporting. *Review of accounting studies* 17:72–95.

Pierce, J. R., and P. K. Schott. 2012. A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries. *Journal of Economic and Social Measurement* 37:61–96.

Piotroski, J. D., and D. T. Roulstone. 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The accounting review* 79:1119–1151.

Povel, P., R. Singh, and A. Winton. 2007. Booms, Busts, and Fraud. *The Review of Financial Studies* 20:1219–1254.

Rangan, S. 1998. Earnings management and the performance of seasoned equity offerings. *Journal of Financial Economics* 50:101–122.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and . Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39:437–485.

Schmidt, K. M. 1997. Managerial Incentives and Product Market Competition. *The Review of Economic Studies* 64:191–213.

Shi, G., J. Sun, and L. Zhang. 2018. Product market competition and earnings management: A firm-level analysis. *Journal of Business Finance & Accounting* 45:604–624.

Shleifer, A. 2004. Does Competition Destroy Ethical Behavior? *Working Paper* .

Shleifer, A., and R. W. Vishny. 1986. Large Shareholders and Corporate Control. *Journal of Political Economy* 94:461–488.

Shleifer, A., and R. W. Vishny. 1997. A survey of Corporate Governance. *The Journal of Finance* LII:737–783.

Smith, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. William Strahan, Thomas Cadell.

Sohn, B. C. 2016. The effect of accounting comparability on the accrual-based and real earnings management. *Journal of Accounting and Public Policy* 35:513–539.

Thompson, R. B., and H. A. Sale. 2003. Securities Fraud as Corporate Governance: Reflections upon Federalism. *Vanderbilt Law Review* 56:859–910.

Tirole, J. 1988. *The Theory of Industrial Organization*. MIT Press.

Tirole, J. 2010. *The Theory of Corporate Finance*. Princeton University Press.

Wang, T. Y. 2011. Corporate Securities Fraud: Insights from a New Empirical Framework. *The Journal of Law, Economics, and Organization* 29:535–568.

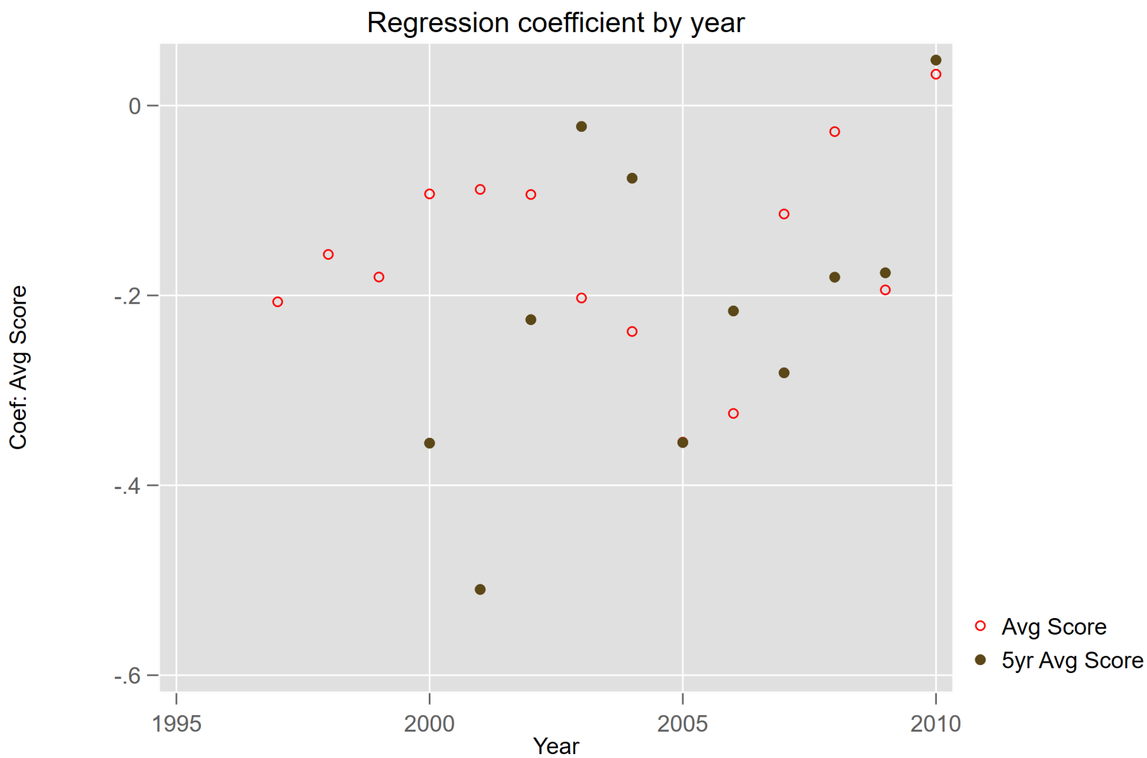
Wang, T. Y., and A. Winton. 2021. Industry informational interactions and corporate fraud. *Journal of Corporate Finance* 69:102024.

Wang, T. Y., A. Winton, and X. Yu. 2010. Corporate Fraud and Business Conditions: Evidence from IPOs. *The Journal of Finance* 65:2255–2292.

Whited, T. M., and G. Wu. 2006. Financial Constraints Risk. *The Review of Financial Studies* 19:531–559.

Zhang, J. H. 2018. Accounting comparability, audit effort, and audit outcomes. *Contemporary Accounting Research* 35:245–276.

Figure 1: Product Market Similarity and Fraud (Independent Cross-Sectional Regressions)



This figure plots coefficient estimates for regressions of fraud on *Product Market Similarity* run independently for each year. The coefficient estimates for each year are obtained using the "Full" set of control variables from our main model specification (Column 4) in Table 2. The red (hollow) dots indicate the 14 estimated coefficients for each firm's most recent (one year lagged) *Product Market Similarity*. We repeat this exercise using a 5 year rolling average of *Product Market Similarity* as our primary independent variable. The 11 coefficients for each year obtained using the 5-year rolling average are indicated by black (solid) dots.

Table 1: Summary Statistics

This table reports summary statistics of firm characteristics at the firm-year level. Variable definitions are provided in the Appendix. Our sample spans 1996 through 2010.

Panel A:

Variable	No. Obs	Mean	Std. Dev.	10th Percentile	90th Percentile
AAER Misstatement	55,381	0.014	0.119	0.000	0.000
SCAC	55,381	0.006	0.074	0.000	0.000
Fraud	55,381	0.019	0.135	0.000	0.000
Product Market Similarity	55,381	0.030	0.023	0.012	0.055
Top5 Similarity	55,381	0.080	0.058	0.017	0.156
Top10 Similarity	55,381	0.066	0.050	0.014	0.135
Top15 Similarity	55,381	0.059	0.047	0.013	0.123
Similarity Precision	55,381	1.002	0.103	0.924	1.053
Sum Similarity	55,381	2.847	4.999	0.074	7.659
Product Market Fluidity	50,402	7.182	3.292	3.292	11.685
SIC3 HHI	55,381	0.176	0.145	0.062	0.332
SIC3 Profit Margin	55,381	-0.039	0.272	-0.346	0.156
TNIC HHI	55,381	0.235	0.197	0.064	0.518
NCOMP TNIC	55,381	74.204	90.520	5.000	204.000
NCOMP SIC3	55,381	121.607	170.694	6.000	351.000
RSST accruals	51,487	0.024	0.240	-0.182	0.220
Change AR	55,381	0.010	0.065	-0.045	0.070
Change Inventory	55,060	0.006	0.049	-0.028	0.050
Pct Soft Assets	55,377	0.541	0.245	0.175	0.852
Change in Cash Sales	51,888	0.195	0.710	-0.214	0.574
ROA	51,497	-0.005	0.195	-0.205	0.141
Change in ROA	54,671	-0.007	0.175	-0.149	0.120
Change in employee	54,053	-0.080	0.469	-0.365	0.241
Dummy Security Issue	55,381	0.920	0.272	1.000	1.000
Whited-Wu Index	54,954	-0.196	0.198	-0.389	0.012
Book Leverage	55,237	0.299	0.294	0.000	0.733
Capex	55,381	0.060	0.093	0.000	0.140
R&D	55,381	0.069	0.117	0.000	0.184
R&D dummy	55,381	0.627	0.484	0.000	1.000
Age	53,295	15.353	11.825	4.0000	35.000
Inst Ownership	43,018	0.516	0.315	0.068	0.922
Number of Analysts	55,381	5.837	7.008	0.000	15.000
Stock Industry Return R2	53,238	0.342	0.173	0.121	0.580
Relative Perf Eval Flag	55,179	0.677	0.467	0.000	1.000
Ln Asset	55,381	5.618	1.937	3.155	8.181

Panel B:

Unique Fraud	516	Unique Fraud Company	498
Unique AAER	332	Unique AAER Company	320
Unique SCAC	206	Unique SCAC Company	201
Fraud Duration	2.2	Fraud Firms	5.5%

Table 2: Product Market Similarity and Fraud

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity*. Our proxy for financial fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. The specification in Column 1 does not include control variables. The specification in Column 2 includes controls used in Dechow et al. (2011). In Columns 3-5 we include our "Full" set of controls as described in Section III and Column 3 also includes Institutional Ownership. In Column 6 we report the standardized regression. All specifications are run at the firm-year level, include year fixed effects, and include explanatory variables lagged by one year. Column 4 also includes three-digit SIC code (SIC3) fixed effects, Column 5 adds year $\times$ SIC3 fixed effects. In Column 6, we run the specification from Column 4 but with standardized regressors. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud	(6) Fraud (Std Reg)
Product Market Similarity	-0.113**	-0.180***	-0.220***	-0.171***	-0.163***	-0.004***
	(-2.089)	(-3.517)	(-3.915)	(-3.946)	(-3.243)	(-3.946)
Ln Asset		0.006***	0.005***	0.006***	0.007***	0.011***
		(4.617)	(3.245)	(2.913)	(3.546)	(2.913)
R&D			0.009	-0.011	-0.012	-0.001
			(0.740)	(-1.008)	(-1.143)	(-1.008)
R&D dummy			-0.000	-0.003	-0.004	-0.003
			(-0.011)	(-0.914)	(-1.135)	(-0.914)
Ln number analysts			0.001	0.001	0.001	0.001
			(0.227)	(0.267)	(0.392)	(0.267)
Inst Ownership			0.007			
			(0.792)			
Whited-Wu Index			0.005	-0.000	0.033*	-0.000
			(0.859)	(-0.057)	(1.892)	(-0.057)
RSST accruals		0.002	-0.003	0.001	0.002	0.000
		(0.538)	(-0.636)	(0.342)	(0.572)	(0.342)
Change AR		0.022*	0.023	0.016	0.023*	0.001
		(1.791)	(1.460)	(1.296)	(1.808)	(1.296)
Change Inventory		0.016	0.026	0.020	0.027	0.001
		(0.756)	(1.258)	(0.937)	(1.225)	(0.937)
Pct Soft Assets		0.019***	0.022***	0.019***	0.019***	0.005***
		(4.241)	(4.186)	(3.825)	(3.733)	(3.825)
Change in Cash Sales		0.005**	0.004**	0.005**	0.005**	0.003**
		(2.250)	(2.057)	(2.227)	(2.367)	(2.227)
Change in ROA		-0.023***	-0.017***	-0.021***	-0.018***	-0.004***
		(-6.132)	(-3.336)	(-5.887)	(-5.212)	(-5.887)
Change in employee		-0.004**	-0.004*	-0.003	-0.003	-0.001
		(-2.101)	(-1.868)	(-1.456)	(-1.414)	(-1.456)
Ln Age		-0.010***	-0.009***	-0.008***	-0.007***	-0.006***
		(-3.413)	(-3.040)	(-3.513)	(-2.974)	(-3.513)
Dummy Security Issue		0.003	-0.001	0.002	0.001	0.002
		(1.327)	(-0.258)	(0.737)	(0.275)	(0.737)
Stock Industry Return R2			-0.009	0.017		0.003
			(-0.822)	(1.515)		(1.515)
Relative Perf Eval Flag			0.007**			
			(2.124)			
TNIC NCOMP			0.001	0.002	0.001	0.002
			(1.205)	(0.945)	(0.756)	(0.945)
Constant	0.015***	-0.004	-0.001	-0.009	-0.007	0.020***
	(6.859)	(-0.708)	(-0.129)	(-0.707)	(-0.635)	(7.128)
Observations	50,526	39,519	28,912	37,144	38,916	37,144
R-squared	0.005	0.015	0.014	0.034	0.079	0.034
FE	Year	Year	Year	Year+Sic3	Year×Sic3	Year+Sic3

Table 3: Shock to Information Using Rival IPOs: 2SLS

This table reports 2SLS estimates for the relation between *Product Market Similarity* and fraud. In columns 1-4, we report the first stage result for 2SLS regression, namely the relationship between a firm's IPO-rival's lagged *Product Market Similarity* on the firm's overall *Product Market Similarity.* In columns 5-8, we use *Product Market Similarity* with competitors undergoing an IPO as an instrument for the firms *Product Market Similarity* on fraud. Column 4 and 8 only include the subsample where the number of competitor IPO > 0. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. The specifications include our "Full" set of control variables described in Section III. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Product Market Similarity (1st Stage)				Fraud (2nd Stage)			
Rival-IPO Product Market Similarity	0.342***	0.291***	0.216***	0.258***				
	(8.424)	(10.554)	(9.382)	(8.859)				
Product Market Similarity (Inst)					-0.618***	-0.530***	-0.495**	-0.314*
					(-4.498)	(-3.181)	(-2.361)	(-1.673)
Ln Num Competitor IPO	-0.003***	-0.004***	-0.002***	-0.000	0.006***	0.005**	0.004**	0.003
	(-3.002)	(-6.682)	(-4.426)	(-1.133)	(3.106)	(2.477)	(2.122)	(1.092)
IPO Size ($)	0.001**	-0.001*	-0.001***	0.016***	0.000	-0.000	-0.000	0.127***
	(2.304)	(-1.741)	(-2.654)	(3.055)	(0.037)	(-0.403)	(-0.722)	(3.025)
Observations	37,335	37,335	37,335	14,823	37,335	37,335	37,335	14,823
R-squared	0.168	0.229	0.352	0.679	0.010	0.014	0.033	0.046
Kleibergen-Paap F-statistics					70.963	111.377	88.762	69.201
Controls	No	Full	Full	Full	No	Full	Full	Full
FE	Year	Year	Year+Sic3	Year+Sic3	Year	Year	Year+Sic3	Year+Sic3

Table 4: Shocks to Competition Using Tariff Reductions

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity* and large industry-level tariff reductions. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. We aggregate ten-digit U.S. Harmonized System codes to year-SIC3 level tariff levels, and identify large year over year tariff reductions. Columns 1 and 2 use the top tercile of reductions, 3 and 4 the top quartile and 5 and 6 the top quintile. Regressions 1,3, and 5 are run without *Product Market Similarity* . All specifications include our "Full" set of controls as described in Section III. Regressions include Year and SIC3 fixed effects. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud	(6) Fraud
Product Market Similarity		-0.172**		-0.173**		-0.172**
		(-2.453)		(-2.462)		(-2.446)
Big Tariff Reduction (Tercile)	0.008	0.008				
	(1.582)	(1.612)				
Big Tariff Reduction (Quartile)			0.009**	0.009**		
			(2.072)	(2.117)		
Big Tariff Reduction (Quintile)					0.006	0.006
					(1.410)	(1.418)
Observations	22,636	22,636	22,636	22,636	22,636	22,636
R-squared	0.036	0.037	0.036	0.037	0.036	0.036
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3

Table 5: Marginal Value of Public Information: Firm Complexity

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity* with the sample split into firm complexity quartiles. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. We define complexity as the number of unique SIC codes spanned by a firms set of competitors according to the TNIC developed by Hoberg and Phillips, 2016. Panel A reports competitor and fraud classifications for each quartile. Panel B reports OLS estimates for each quartile. The specification in Column 7 includes an interaction term, rather than quartile splits. All specifications include the "Full" set of control variables described in Section III. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

Complexity	Low			High	
	Q1	Q2	Q3	Q4	
Panel A: Summary Statistics					
Unique SICs in TNIC	3.3	8.2	13	22.5	
Competitors in TNIC	13	49	117	150	
% Fraud	1.6	1.7	1.9	2.1	
Product Market Similarity	2.7	2.8	3.3	3.5	
Panel B: Regression Results					
Product Market Similarity	-0.168***	-0.197**	-0.201	-0.683***	0.027
	(-3.418)	(-2.029)	(-1.508)	(-5.053)	(0.460)
Complexity					-0.003
					(-0.873)
Product Market Similarity × Complexity					-0.145***
					(-4.054)
Observations	9,995	9,628	9,018	8,503	37,144
R-squared	0.016	0.017	0.024	0.026	0.017
FE	Year	Year	Year	Year	Year

Table 6: Marginal Value of Public Information: External Monitors

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity* with the sample split into quartiles based on Institutional Ownership HHI and Analyst Coverage. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Panel A reports OLS estimates for each Institutional ownership HHI quartile. Panel B contains quartile splits based on the number of sell-side analysts covering a firm in a given year and whether that firm is covered by at least one star analyst or not. Panel's C and D repeat the analysis in Panel B but splits the sample bye pre- and post- REG FD (i.e., pre- and post- 2000). Column 7 includes specifications with interaction terms, rather than quartile-splits. All specifications include the "Full" set of control variables described in Section III. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

Panel A Institutional Ownership HHI							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Low			High			
	Q1	Q2	Q3	Q4			
Institutional Ownership							-0.021**
							(-2.284)
Product Market Similarity (PMS)	-0.425***	-0.298***	-0.107**	-0.066			-0.282***
	(-3.373)	(-3.145)	(-2.175)	(-1.217)			(-3.829)
PMS × Institutional Ownership							0.318*
							(1.844)
Observations	7,943	8,225	8,085	7,153			31,415
R-squared	0.023	0.022	0.012	0.007			0.016
	Low			High	Non Star	Star	
	Q1	Q2	Q3	Q4			
Panel B Analyst Coverage							
Analyst Coverage							0.018***
							(4.063)
Product Market Similarity (PMS)	-0.001	-0.204**	-0.456***	-0.487***	-0.133*	-0.388***	0.193**
	(-0.014)	(-2.417)	(-3.940)	(-4.378)	(-1.737)	(-3.985)	(2.170)
PMS × Analyst Coverage							-0.255***
							(-3.986)
Observations	7,742	7,179	6,420	6,450	12,944	7,221	27,791
R-squared	0.006	0.014	0.038	0.028	0.011	0.028	0.019
Panel C Analyst Coverage - Pre 2000							
Analyst Coverage							0.018**
							(1.978)
Product Market Similarity (PMS)	0.058	-0.267**	-0.786***	-0.581**	-0.336**	-0.508*	0.275**
	(1.037)	(-2.170)	(-3.296)	(-2.487)	(-2.442)	(-1.774)	(1.996)
PMS × Analyst Coverage							-0.381***
							(-2.817)
Observations	2,704	2,580	2,263	2,201	3,275	2,167	9,748
R-squared	0.012	0.020	0.072	0.039	0.024	0.057	0.029
Panel D Analyst Coverage - Post 2000							
Analyst Coverage							0.020***
							(4.541)
Product Market Similarity (PMS)	-0.041	-0.164	-0.218*	-0.426***	-0.068	-0.312***	0.221*
	(-0.730)	(-1.533)	(-1.778)	(-2.939)	(-0.830)	(-2.880)	(1.681)
PMS × Analyst Coverage							-0.224***
							(-2.789)
Observations	5,038	4,599	4,157	4,249	9,669	5,054	18,043
R-squared	0.008	0.013	0.025	0.029	0.011	0.022	0.016

Table 7: Robustness: Financial Statement Comparability

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity* and an output-based measure of *Accounting Comparability* from De Franco et al. (2011). Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Panel A presents estimates for the relation between fraud and *Product Market Similarity*, controlling for *Accounting Comparability* (Firm and Industry Level). Panel B presents the mean *Product Market Similarity* and % Fraud for firm-years with below and above median *Accounting Comparability* and the correlation between the two measures. Panel C includes regressions for sub samples split on high and low *accounting comparability*. The specification in Column 1-2, 4-5 we split the data between low (below median) and high (above median) *Accounting Comparability*. In column 3 and 6 we include the interaction term of *EProduct Market Similarity* and *High AcctComp*, which is an indicator variable that equals one if a firm's *Accounting Comparability* is above the median. All specifications are run at the firm-year level, columns 1-3 include Year + Sic3 fixed effects, columns 4-6 include Year×Sic3 fixed effect, and explanatory variables are lagged by one year. All regression specifications also include the "Full" set of controls from Table 2, including Year and SIC3 fixed effect transformations. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

Panel A: Controlling for Accounting Comparability (Dependent Variable = Fraud)						
Product Market Similarity	-0.194***	-0.153***	-0.155**	-0.193***	-0.153***	-0.154**
	(-3.299)	(-2.712)	(-2.437)	(-3.260)	(-2.708)	(-2.422)
Accounting Comparability	-0.000	-0.001	-0.001			
	(-0.252)	(-0.782)	(-1.139)			
Accounting Comparability (Industry)				-0.000	-0.000	-0.001
				(-0.597)	(-0.908)	(-0.799)
Observations	20,197	20,192	19,886	20,197	20,192	19,886
R-squared	0.017	0.038	0.091	0.017	0.038	0.091
FE	Year	Year+sic3	Year×Sic3	Year	Year+sic3	Year×Sic3
Panel B: Accounting Comparability: Summary Statistics						
Accounting Comparability				All	Low	High
Product Market Similarity				3.1	3.1	2.9
% Fraud				2.0	1.3	2.0
Observations				33,400	16,898	16,558
Corr(Econ Com, Acc Comp)						-1.5%
Panel C: Dependent Variable = Fraud						
	(1)	(2)	(3)	(4)	(5)	(6)
Accounting Comparability	Low	High	All	Low	High	All
Product Market Similarity	-0.040	-0.248***	-0.073	0.004	-0.269***	-0.061
	(-0.485)	(-4.258)	(-1.071)	(0.038)	(-4.128)	(-0.800)
High AcctComp			0.002			0.002
			(0.432)			(0.383)
PMS # High AcctComp			-0.158*			-0.179*
			(-1.843)			(-1.956)
Observations	9,848	10,330	20,192	9,367	9,890	19,886
R-squared	0.057	0.056	0.038	0.175	0.119	0.091
FE	Year+Sic3	Year+Sic3	Year+Sic3	Yea×Sic3	Year×Sic3	Year×Sic3

Table 8: Robustness: Controlling for Competition

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity*, controlling for measures of competition. Our proxy for corporate fraud includes a combination of misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Column 1 includes sales based Herfindahl-Hirschman Index (HHI) according to three digit SIC code (SIC3). Column 2 also includes the number of competitors (logged) in the same SIC3. Column 3 also includes the profit margin and an industry concentration measure. In Column 4 we include the sales based HHI according to the firm's TNIC. Column 5 also includes the number of competitors within a firm's TNIC. Column 6 also includes the product market fluidity. In Column 7 we included a 10-k based competition measure from Li et al. (2013) which is only available for a subset of our sample. All specifications include the "Full" set of controls as described in Section III. All specifications are run at the firm-year level, include year and SIC3 fixed effects, and explanatory variables lagged by one year. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud	(6) Fraud	(7) Fraud
Product Market Similarity	-0.159***	-0.161***	-0.159***	-0.161***	-0.171***	-0.183***	-0.172**
	(-3.317)	(-3.305)	(-3.248)	(-3.518)	(-3.936)	(-4.209)	(-2.264)
SIC3 HHI	0.018	0.035	0.009				
	(0.842)	(1.435)	(0.383)				
SIC3 NCOMP		0.016**	0.016***				
		(2.155)	(2.680)				
SIC3 PM sale			-0.007				
			(-0.700)				
SIC3 Top 4 Concentration			0.042**				
			(2.121)				
TNIC HHI				-0.002	0.006	0.006	
				(-0.286)	(0.968)	(0.944)	
TNIC NCOMP					0.002	0.002	
					(1.107)	(0.912)	
Product Market Fluidity						0.001	
						(1.373)	
Competition 10K							0.004
							(1.074)
Observations	37,144	37,144	37,144	37,144	37,144	35,999	18,696
R-squared	0.034	0.034	0.034	0.034	0.034	0.035	0.048
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3

Table 9: Robustness: Alternate Independent Variable Constructions

This table reports OLS estimates for the incidence of fraud on alternative constructions of our primary independent variable. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Panel A reports OLS regressions and Panel B reports the regressions with standardized regressors. Column 1 presents results for the main dependent variable used throughout our analysis. Columns 2-4 replace *Product Market Similarity* with a firms similarity averaged across its closest 15, 10, and 5 competitors, respectively. In Column 5, we replace *Product Market Similarity* with the *Similarity Precision* measure outlined in section III. The unit of observation in this analysis is the firm-year. All specifications include the "Full" set of controls as described in Section III. They include year and SIC3 fixed effects, and the explanatory variables lagged by one year. The t-statistics, calculated from standard errors clustered at three digit SIC code (SIC3) level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud
Panel A: OLS Regressions					
Product Market Similarity	-0.171***				
	(-3.946)				
Top 15 Similarity		-0.089**			
		(-2.020)			
Top 10 Similarity			-0.091**		
			(-2.244)		
Top 5 Similarity				-0.086***	
				(-2.730)	
Similarity Precision					-0.037***
					(-3.893)
Observations	37,144	37,144	37,144	37,144	37,144
R-squared	0.034	0.034	0.034	0.034	0.034
Panel B: Standardized Regressions					
Product Market Similarity	-0.004***				
	(-3.946)				
Top 15 Similarity		-0.004**			
		(-2.020)			
Top 10 Similarity			-0.005**		
			(-2.244)		
Top 5 Similarity				-0.005***	
				(-2.730)	
Similarity Precision					-0.004***
					(-3.893)
Observations	37,144	37,144	37,144	37,144	37,144
R-squared	0.034	0.034	0.034	0.034	0.034
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3

Internet Appendix for: Product Market Similarity, Benchmarking, and Corporate Fraud

Table A.1: Correlations

Correlation coefficients are reported for various measures of *Product Market Similarity* and competition. Our sample covers 1996 through 2011.

	Prod Sim.	Top 15 Sim.	Top 10 Sim.	Top 5 Sim.	Sim. Precision	Sum Sim.	SIC3 HHI	SIC3 PM	SIC3 NCOMP	TNIC HHI	TNIC NCOMP	Acct Comp.	10-K Competition
Product Market Similarity	1												
Top 15 Similarity	0.781	1											
Top 10 Similarity	0.76	0.993	1										
Top 5 Similarity	0.725	0.949	0.975	1									
Similarity Precision	0.656	0.593	0.611	0.656	1								
Sum Similarity	0.331	0.718	0.693	0.628	0.267	1							
SIC3 HHI	-0.059	-0.171	-0.172	-0.161	-0.106	-0.189	1						
SIC3 Profit Margin	0.073	-0.001	0.009	0.024	-0.018	-0.244	0.226	1					
SIC3 NCOMP	-0.056	0.093	0.082	0.056	0.078	0.284	-0.38	-0.521	1				
TNIC HHI	-0.154	-0.435	-0.461	-0.466	-0.4	-0.385	0.142	0.025	-0.092	1			
TNIC NCOMP	0.161	0.606	0.586	0.525	0.245	0.897	-0.227	-0.315	0.419	-0.443	1		
Acct Comparability	-0.057	-0.013	-0.014	-0.016	0.004	0.045	-0.104	-0.085	0.113	-0.043	0.078	1	
10-K Competition	-0.049	0.039	0.036	0.027	0.017	0.096	-0.126	-0.156	0.202	-0.059	0.197	0.118	1

Table A.2: Product Market Similarity and Fraud (Non-Linear Specifications)

This table reports non-linear regression estimates for the incidence of fraud on *Product Market Similarity.* Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Columns 1, 2, 4 and 5 are logit specifications and 3 and 6 are probit specifications. Columns 1-3 only include size as a control variable, while Columns 4-6 include the "Full" set of controls as described in Section II. In specifications 2-3, 5-6, we standardize the variable of interest for the purposes of understanding economic magnitude. The unit of observation in this analysis is the firm-year. All specifications include year fixed effects and the "Full" set of controls as described in Section III. Control variables are lagged by one year. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud	(6) Fraud
Product Market Similarity	-9.709*			-21.955***		
	(-1.926)			(-3.847)		
Standardized Product Market Similarity		-0.236*	-0.098**		-0.533***	-0.213***
		(-1.926)	(-2.124)		(-3.847)	(-3.901)
Ln Asset	0.269***			0.287***	0.287***	
	(11.378)			(3.887)	(3.887)	
Standardized Ln Asset		0.534***	0.223***			0.2338***
		(11.378)	(9.400)			(3.602)
Constant	-5.627***	-4.409***	-2.246***	-6.458***	-7.129***	-3.041***
	(-24.338)	(-26.371)	(-34.920)	(-9.594)	(-11.118)	(-11.273)
Observations	54,852	54,852	54,852	37,144	37,144	37,144
Specification	Logit	Logit	Probit	Logit	Logit	Probit
Controls	None	None	None	Full	Full	Full
FE	Year	Year	Year	Year	Year	Year

Table A.3: Product Market Similarity: Controlling for Accounting Comparability and F-score

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity.* Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. We control for *Accounting Comparability* based on De Franco et al. (2011), and F-score based on Dechow et al. (2011). *AcctCompInd* is the median *Accounting Comparability* across all rivals in firm *i*'s industry. *F-score* is the scaled predicted probability of misstatements based on financial characteristics of misstating firms, often used as a red flag of the likelihood of earnings management or misstatement. Columns 1-3 include controls listed and Columns 4-6 includes the "Full" set of controls as described in Section III. The firm-year is the unit of observation in this analysis. All control variables, lagged by one year, and the level of fixed effects are indicated in the table. The t-statistics, calculated from standard errors clustered at the SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1) Fraud	(2) Fraud	(3) Fraud	(4) Fraud	(5) Fraud	(6) Fraud
Product Market Similarity	-0.224***	-0.161***	-0.161**	-0.197***	-0.154***	-0.148**
	(-3.604)	(-2.786)	(-2.456)	(-3.329)	(-2.730)	(-2.337)
Accounting Comparability				-0.000	-0.001	-0.001
				(-0.103)	(-0.575)	(-1.046)
AcctCompInd	-0.000	-0.000	-0.001			
	(-0.597)	(-0.907)	(-0.758)			
F-score	0.016***	0.014***	0.014***	0.018**	0.020**	0.017**
	(4.265)	(3.713)	(3.605)	(2.180)	(2.589)	(2.016)
R&D	0.010	-0.004	-0.007	0.013	-0.001	-0.003
	(0.761)	(-0.407)	(-0.628)	(1.025)	(-0.111)	(-0.236)
R&D dummy	-0.001	-0.005	-0.005	-0.001	-0.004	-0.005
	(-0.244)	(-0.844)	(-0.928)	(-0.330)	(-0.813)	(-0.841)
Ln number analysts	-0.001	-0.001	-0.002	0.000	-0.000	-0.001
	(-0.189)	(-0.336)	(-0.618)	(0.026)	(-0.145)	(-0.276)
Whited-Wu Index	0.001	0.003	0.010	0.002	0.002	0.005
	(0.182)	(0.553)	(0.431)	(0.337)	(0.368)	(0.172)
Ln Age	-0.011**	-0.011**	-0.011**	-0.012***	-0.011**	-0.011**
	(-2.482)	(-2.384)	(-2.406)	(-2.622)	(-2.467)	(-2.383)
Stock Industry Return R2	0.001	0.028***		0.005	0.028***	
	(0.109)	(2.670)		(0.423)	(2.676)	
Relative Perf Eval Flag	0.007*			0.007**		
	(1.956)			(2.032)		
TNIC NCOMP	0.003*	0.002	0.002	0.003**	0.002	0.002
	(1.699)	(0.729)	(0.695)	(2.167)	(0.906)	(0.782)
Ln Asset	0.005***	0.007***	0.008***	0.005***	0.007***	0.007***
	(2.971)	(3.323)	(3.560)	(2.952)	(3.218)	(2.836)
Constant	-0.008	-0.013	-0.004	-0.007	-0.009	-0.000
	(-0.614)	(-0.942)	(-0.245)	(-0.528)	(-0.631)	(-0.032)
Observations	20,197	20,192	19,886	20,197	20,192	20,429
R-squared	0.016	0.038	0.090	0.017	0.039	0.094
Additional Controls	Dechow	Dechow	Dechow	Full	Full	Full
FE	Year	Year sic3	Year#sic3	Year	Year sic3	Year#sic3
Cluster	sic3	sic3	sic3	sic3	sic3	sic3

Table A.4: Product Market Similarity and Fraud (Other Fixed Effects)

This table reports OLS estimates for the incidence of fraud on *Product Market Similarity.* Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. All specifications include the "Full" set of controls as described in Section III. All specifications are run at the firm-year level, include year and TNIC200 (TNIC300/firm) fixed effects, and explanatory variables are lagged by one year. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)
	Fraud	Fraud	Fraud	Fraud	Fraud
Product Market Similarity	-0.139***	-0.114**	-0.155***	-0.129**	-0.093*
	(-3.171)	(-2.387)	(-3.337)	(-2.496)	(-1.691)
Observations	37,143	39,042	37,136	38,771	36,380
R-squared	0.029	0.066	0.032	0.081	0.437
Controls	Full	Full	Full	Full	Full
FE	Year+TNIC200	Year # TNIC200	Year+TNIC300	Year # TNIC300	Year+Firm

Table A.5: V-Score Similarity of Vertical Relationships and Fraud

This table reports OLS estimates of fraud on *V-Similarity*. *V-Similarity* is calculated from vertical similarity scores obtained from Frésard et al. (2020), which measure the extent to which a firm operates in a common product market with its vertically-related peers. Our proxy for fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Column 2 includes controls used in Dechow et al. (2011). In Columns 3-7 we include our "Full" set of controls as described in Section III. All specifications include year fixed effects and explanatory variables lagged by one year. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Fraud	Fraud	Fraud	Fraud	Fraud	Fraud	Fraud
V-Score Similarity	-1.150*	-1.490**	-1.128*	-0.623	-1.445**	-0.909	-1.461**
	(-1.827)	(-2.332)	(-1.719)	(-0.818)	(-2.350)	(-1.145)	(-2.186)
R&D			0.015	-0.011	-0.010	-0.013	-0.015
			(0.893)	(-0.739)	(-0.903)	(-0.858)	(-1.460)
R&D dummy			0.000	-0.005	-0.007	-0.005	-0.008
			(0.143)	(-1.254)	(-1.350)	(-1.440)	(-1.607)
Ln number analysts			0.001	-0.001	-0.001	-0.001	-0.000
			(0.345)	(-0.359)	(-0.402)	(-0.355)	(-0.227)
Whited-Wu Index			-0.003	-0.004	-0.002	0.039*	0.003
			(-0.374)	(-0.650)	(-0.219)	(1.939)	(0.290)
RSST accruals		0.001	0.001	-0.000	0.000	0.002	-0.001
		(0.245)	(0.299)	(-0.024)	(0.009)	(0.395)	(-0.278)
Change AR		0.018	0.014	0.012	0.017	0.018	0.020
		(1.195)	(0.889)	(0.796)	(1.091)	(1.128)	(1.228)
Change Inventory		0.027	0.027	0.026	0.022	0.039	0.026
		(1.169)	(1.147)	(1.118)	(0.924)	(1.537)	(1.015)
Pct Soft Assets		0.025***	0.030***	0.021***	0.025***	0.020***	0.026***
		(4.648)	(5.302)	(3.750)	(4.193)	(3.505)	(4.066)
Change in Cash Sales		0.005**	0.005**	0.005**	0.006***	0.006**	0.005**
		(2.267)	(2.393)	(2.347)	(2.618)	(2.524)	(2.527)
Change in ROA		-0.022***	-0.021***	-0.020***	-0.021***	-0.017***	-0.021***
		(-4.705)	(-4.301)	(-4.585)	(-4.837)	(-3.853)	(-4.127)
Change in employee		-0.004	-0.004	-0.003	-0.003	-0.003	-0.002
		(-1.256)	(-1.211)	(-0.795)	(-0.788)	(-0.903)	(-0.471)
Ln Age		-0.009***	-0.009***	-0.009***	-0.007***	-0.008***	-0.007***
		(-3.498)	(-3.304)	(-3.465)	(-3.236)	(-3.009)	(-3.338)
Dummy Security Issue		0.003	0.001	0.002	0.002	0.001	0.002
		(1.105)	(0.485)	(0.739)	(0.699)	(0.435)	(0.589)
Ind R2			-0.004	0.015	-0.007		
			(-0.324)	(1.435)	(-0.475)		
RPE Flag			0.008**		0.005		
			(2.334)		(1.283)		
TNIC NCOMP			0.000	0.000	-0.001	0.000	-0.001
			(0.251)	(0.116)	(-0.363)	(0.087)	(-0.327)
Ln Asset		0.005***	0.005***	0.006***	0.007***	0.008***	0.007***
		(4.651)	(2.755)	(2.919)	(2.994)	(3.322)	(3.013)
Constant	0.014***	-0.009*	-0.011	-0.008	-0.003	-0.006	-0.002
	(4.250)	(-1.785)	(-1.075)	(-0.665)	(-0.272)	(-0.577)	(-0.190)
Observations	41,436	32,413	30,435	30,435	30,425	31,847	31,690
R-squared	0.005	0.015	0.015	0.036	0.034	0.092	0.094
FE	Year	Year	Year	Year Sic3	Year Tnic3	Year#Sic3	Year#Tnic3

Table A.6: Product Market Similarity and Fraud - Bivariate Probit

This table reports coefficient estimates from the partially observable bivariate probit model, $P(Z = 1) = P(F = 1)P(D = 1|F = 1)$, used in Wang and Winton (2021) . In specifications 2-4, we replace *Product Market Similarity* with the similarity with the firms most similar 5, 10, and 15 peers respectively.

	(1)		(2)		(3)		(4)	
	P(F)	$P(D\|F)$	P(F)	$P(D\|F)$	P(F)	$P(D\|F)$	P(F)	$P(D\|F)$
Product Market Similarity	-7.378*	4.221						
	(-1.938)	(0.497)						
Top 15 Similarity			-3.209**	4.603				
			(-2.385)	(1.582)				
Top 10 Similarity					-2.970**	4.111		
					(-2.444)	(1.580)		
Top 5 Similarity							-2.856***	3.812*
							(-2.695)	(1.665)
SIC3 NCOMP	-0.070**	0.303***	-0.076**	0.300***	-0.076**	0.304***	-0.079**	0.312***
	(-1.987)	(3.875)	(-2.088)	(3.393)	(-2.095)	(3.570)	(-2.148)	(3.870)
MA	0.367**	-0.514	0.371**	-0.518*	0.371**	-0.515*	0.371**	-0.505
	(2.055)	(-1.619)	(2.032)	(-1.672)	(2.034)	(-1.650)	(2.034)	(-1.592)
Stock Ind Return R2	0.642**	-2.082***	0.706**	-2.165***	0.707**	-2.176***	0.716**	-2.208***
	(1.974)	(-3.444)	(2.177)	(-3.598)	(2.173)	(-3.699)	(2.188)	(-3.881)
R&D	-0.123	-0.338	0.144	-0.931	0.125	-0.887	0.138	-0.924
	(-0.209)	(-0.199)	(0.215)	(-0.502)	(0.188)	(-0.482)	(0.210)	(-0.503)
R&D Indicator	0.257***	-0.664***	0.315***	-0.710***	0.317***	-0.718***	0.324***	-0.740***
	(3.052)	(-3.879)	(3.509)	(-3.872)	(3.512)	(-3.962)	(3.554)	(-4.125)
Capx	0.212	-0.606	0.265	-0.724	0.256	-0.706	0.260	-0.716
	(0.468)	(-0.653)	(0.580)	(-0.781)	(0.560)	(-0.763)	(0.570)	(-0.774)
Ln number analysts	-0.034	0.113	-0.023	0.089	-0.024	0.092	-0.027	0.097
	(-0.744)	(1.347)	(-0.504)	(1.027)	(-0.523)	(1.066)	(-0.578)	(1.137)
Inst Ownership	0.522***	-0.521	0.517***	-0.551*	0.516***	-0.541	0.520***	-0.532
	(3.438)	(-1.473)	(3.374)	(-1.691)	(3.367)	(-1.639)	(3.396)	(-1.555)
Ln Asset	0.228***	-0.239***	0.229***	-0.249***	0.230***	-0.247***	0.233***	-0.244***
	(7.600)	(-2.705)	(7.514)	(-3.360)	(7.536)	(-3.222)	(7.588)	(-2.899)
Ln Age	-0.563***	0.834***	-0.576***	0.863***	-0.577***	0.858***	-0.576***	0.841***
	(-8.257)	(4.597)	(-8.317)	(5.667)	(-8.326)	(5.402)	(-8.355)	(4.802)
Abnormal ROA		-1.079		-1.001		-1.033		-1.081
		(-1.629)		(-1.444)		(-1.521)		(-1.614)
Abnormal AAER		0.130***		0.116***		0.118***		0.121***
		(3.373)		(2.694)		(2.896)		(3.151)
Abnormal Volatility		-1.015		-0.816		-0.835		-0.862
		(-1.333)		(-1.146)		(-1.167)		(-1.187)
Abnormal Turnover		2.359		2.233		2.300		2.419
		(1.062)		(1.048)		(1.070)		(1.102)
Disastrous Return		0.489**		0.433**		0.440**		0.449**
		(2.358)		(2.101)		(2.186)		(2.276)
Dummy Security Issue	-0.249	0.683*	-0.229	0.621	-0.227	0.619	-0.215	0.596
	(-1.310)	(1.839)	(-1.176)	(1.644)	(-1.166)	(1.641)	(-1.104)	(1.571)
EFN	-0.111		-0.112		-0.113		-0.116	
	(-0.489)		(-0.524)		(-0.525)		(-0.528)	
ROA	0.064		0.061		0.068		0.077	
	(0.285)		(0.280)		(0.309)		(0.345)	
Book Leverage	-0.087		-0.103		-0.104		-0.106	
	(-1.140)		(-1.391)		(-1.393)		(-1.389)	
Pct Soft Assets	0.548***		0.537***		0.544***		0.554***	
	(3.577)		(3.433)		(3.532)		(3.557)	
Constant	-1.683***	-0.737	-1.733***	-0.479	-1.734***	-0.551	-1.736***	-0.661
	(-4.896)	(-0.454)	(-4.996)	(-0.307)	(-5.015)	(-0.357)	(-5.040)	(-0.422)
Observations	30,117	30,117	30,117	30,117	30,117	30,117	30,117	30,117
FE	Year	Year	Year	Year	Year	Year	Year	Year

Table A.7: Product Market Similarity and Restatements

This table reports OLS estimates for the restatement on *Product Market Similarity*. *Restatement with Misstatement* is an indicator variable that equals one if the restatement is accompanied by a misstatement (what we use as a proxy for fraud) in the same year. All Columns include our "Full" set of controls as described in Section III. All specifications are run at the firm-year level, include year fixed effects and three-digit SIC code (SIC3) fixed effects. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coefficient estimates. Statistical significance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)
	Restatement	Restatement with Misstatement
Product Market Similarity	0.108	-0.122***
	(0.411)	(-2.924)
Observations	27,790	27,790
R-squared	0.062	0.032
Controls	Full	Full
FE	Year+Sic3	Year+Sic3
Cluster	sic3	sic3

Table A.8: Institutional Ownership and Analyst Coverage - Chow Test

This table reports Chow test for coefficient on *Product Market Similarity* in Table 6. χ^2 shows the difference in estimates on *Product Market Similarity* between group Q1 (bottom quartile) and Q4 (top quartile), and between group Star and Non Star. Statistical significance at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	Panel A Institutional Ownership HHI
	Q4 - Q1 (High - Low)
χ^2	7.97
Prob $>\chi^2$	0.005***

	Panel B Analyst Coverage	
	Q4 - Q1 (High - Low)	Star - Non Star
χ^2	17.33	6.56
Prob $>\chi^2$	0.000***	0.010**

	Panel C Analyst Coverage - Pre 2000	
	Q4 - Q1 (High - Low)	Star - Non Star
χ^2	7.57	0.35
Prob $>\chi^2$	0.006***	0.552

	Panel D Analyst Coverage - Post 2000	
	Q4 - Q1 (High - Low)	Star - Non Star
χ^2	5.68	3.99
Prob $>\chi^2$	0.017**	0.046**