Product Differentiation, Benchmarking, and Corporate Fraud

Audra Boone[y] William Grieser[y] Rachel Li[z] Parth Venkat[x]

March 15, 2019

Abstract

We find that product market differentiation is an economically meaningful and sta-

Our analyses reveal that the incidence of fraud is significantly lower for firms with a less differentiated product mix. Specifically, a one standard deviation increase in average product similarity score is associated with a 14.8-23.7% decrease in the rate of settled SEC enforcement actions and fraud-driven class action lawsuits.

nancial reports and industry trends to uncover misrepresentations" (Dyck et al., 2010).

decrease outsiders ability to detect reporting manipulations. Thus, our findings suggest that managers rationally respond to enhanced detection rates by committing less fraud. We find corroborating evidence for the commission effect according to the bivariate probit (partial observability) model highlighted in Wang et al. (2010).

In falsification tests, we replace settled fraud-driven misstatements with unintentional accounting restatements and fail to observe a discernable pattern with product market similarity. This finding increases our confidence that we have identified an economically meaningful link between fraud and product market similarity in our primary analysis. Additionally, our primary results remain qualitatively similar to a variety of non-linear model specifications, as well variations in our set of control variables and the time period used in our estimation. Our results are also robust to a) variations in the level of winsorization, b) to removing outliers, rather than winsorizing, and c) eliminating any winsorization of the data.Further, our results are not sensitive to the specific construction of our independent variable. In particular, our results are qualitatively unchanged when we average product similarity scores across a firms top 5, 10 or 15 closest rivals, rather than averaging across the firms entire competitor network, when we use a precision weighted average, and when we use the number of rivals above a given percentile of product market similarity (75th, 90th and 95th).

In summary, we find a strong and robust link between product market differentiation and corporate fraud. Indeed, our estimates suggest that product market similarity has an economically larger effect on fraud than any factor, other than firm size, previously documented in the literature. Our initial results suggest that product market similarity imposes a strong disciplining effect on financial reporting misconduct. Further, while none of our follow-up analyses provides incontrovertible evidence in isolation, the preponderance of evidence suggests that the disciplining effect stems through a benchmarking channel. These results indicate that market-based mechanisms, particularly through enhanced information environment, play an important role in both the incentive to commit fraud and the ability of external parties to uncover fraudulent activities.

II Literature Review

Our paper relates to the literature examining the effect of various measures of competition on managerial discipline. On one hand, competition can diminish conflicts of interest by incentivizing managerial effort (Nickell, 1996) or by reducing resources available for rent extraction (Ades and Di Tella, 1999; Schmidt, 1997). On the other hand, competition has been argued to pressure managers to distort the perceived performance relative to rivals

(Shleifer, 2004; Tirole, 2010; Andergassen, 2016). Until recently, only coarse industry-level measures of competition have been available to researchers, which has introduced challenges in identifying the potential e ect of these opposing forces. Indeed, the existing empirical evidence on the link between competition and fraud is often contradictory and inconclusive (e.g. Holmstrom, 1982; Nalebu and Stiglitz, 1983; Karaoglu et al., 2006). We shed light on this relationship by exploiting newly developed, rm-level measures of product di erentiation that allow us to conduct more powerful tests. Consistent with a disciplining channel of competition, we document that product market similarity is strongly associated with a lower incidence of fraud.

In addition, our work suggests that benchmarking is an important factor to consider in studying competition, as it enhances information, and therefore facilitates monitoring ability
(

reassuring for the disciplining channel of product market similarity.

Our work also relates to the literature on corporate governance and corporate fraud (e.g. Beasley, 1996; Farber, 2005; Khanna et al., 2015). Several papers propose that corporate governance mechanisms are endogenous responses to the cost 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 findings suggest an alternate source of external discipline: product market competition, which compliments recent work suggesting that product market competition can substitute for other formal corporate governance mechanisms (Giroud and Mueller, 2010; Chhaochharia et al., 2016). The remainder of the paper is organized as follows. Section II discusses the data and our set of control variables. Section III contains the empirical measures of competition. Section IV covers the results, and Section V concludes.

III Data

We follow recent empirical work of Donelson et al. (2017) by defining corporate accounting fraud as, the intentional, material misstatement of financial statements that causes damages to investors. Donelson et al. (2017) advocate using a combination of public and private enforcement actions through AAER and class action lawsuits to capture financial reporting fraud to mitigate measurement error. While regulatory enforcement is important, other participants, such as the media, industry regulators, and employees, serve as important actors in this arena (Dyck et al., 2010).

We obtain AAER data for the sample period 1996-2010. According to the Center for Financial Reporting and Management, the U.S. Securities and Exchange Commission (SEC) issues AAERs during, or at the conclusion of, an investigation against a company, an auditor, or an officer for alleged accounting or auditing misconduct. 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. The misstatement investigations in our sample occur because of bribery, fraud, inflated assets, financial reporting related enforcement actions concerning civil lawsuits brought by the in federal court, and orders concerning the institution and/or settlement of administrative proceedings.

We construct our sample of class action lawsuits following the work of Choi et al. (2009), Griffin et al. (2004), Jayaraman and Milbourn (2009), and Thompson and Sale (2003). We start by downloading all class action lawsuits from the SCAC hosted by Stanford University for 1996 through 2011 and scan each filing to only include 10-b5 class action lawsuits, which eliminates those lawsuits that occur for non-financial reasons[6].

We define each firm-year as an AAER year, a SCAC year, both, or neither. Our primary independent variable, fraud, is a binary variable equal to one for all firm years in which there is an AAER or SCAC. We exclude firms in the financial and utilities industries and firms headquartered outside the United States. Further, we drop ADRs, firms with assets less than \$1M, and firms with missing assets or sales items in Compustat. Our final sample of corporate fraud events includes 935 firm-years that are affected by AAER misstatements in at least one quarterly or annual financial statement from 322 unique firms from 1996 to 2010. In addition, our sample includes 311 class action lawsuits affecting 299 firms from 1996 to 2011. In total, our sample contains 498 firms and 1,217 firm-years, flagged as years with fraudulent reporting. These figures are very closely in line with those of (Dyck et al., 2010). As shown in Table 1, the overall incidence of fraud in our sample is 1.9%.

To construct our set of control variables, we follow work in the finance and accounting literature related to corporate fraud (variable definitions are reported in Table A.1). We include predictors of accounting misstatements from Dechow et al. (2011), which include Richardson et al. (2005) (RSST) accruals, change in accounts receivable (ΔAR), change in inventory (ΔInventory), the percentage of soft assets (% Soft Assets), change in cash sales (ΔCash Sales), change in ROA (ΔROA), change in employees (ΔEmployees), and a dummy for security issuance (dSecurity Issue).

The variable RSST accruals 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 change in inventory (ΔInventory) are associated with incentives to improve sales growth and gross profit margin. A firms soft assets as a percentage of total assets (% soft assets) is associated with more discretion for earnings management. We define % soft assets as total assets minus property plant and equipment and cash and cash equivalents, all scaled by total assets. Change in cash-based sales (ΔCash Sales) excludes accrual-based sales to measure the portion of sales that are not subject to discretionary accrual management. Change in ROA (ΔROA) controls for changes in earnings growth. The variable ΔEmployees is the percentage change in employees less the percentage change in total assets. This measure is associated with labor costs and must be expensed as incurred. Reducing the number of employees can boost a firms short-term financial performance by immediately lowering expenses. Finally, we in-

[6] Karpoff et al. (2017) note the importance of additional checks of the sources to ensure that they contain instances of fraud.

clude a dummy variable (dSecurity Issue) equal to one for firm years in which a firm issues debt or equity, which can increase incentives to manage earnings (Rangan, 1998). We refer to specifications including only the controls from Dechow et al. (2011) as the Dechow set of controls.

We also include specifications that contain proxies for monitoring mechanisms and corporate opaqueness, which could potentially influence the marginal impact of our proposed benchmarking channel. We include Institutional Ownership, the natural log of the number of analysts covering a firms stock (Ln Num Analysts), research and development expenses (R&D), and industry stock return r-squared (Ind R2).[7] To construct the industry r-squared, we follow Wang and Winton (2014) and 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 a given three-digit SIC code. Managers may feel pressured to commit fraud when they require capital from outside sources (Teoh et al., 1998; Wang and Winton, 2014). Thus, we include the Whited and Wu (2006) Index for financial constraints.[8]

We include the natural log of total assets (ln assets) as a measure of firm size. We also include the variable book leverage, which is defined as long and short-term debt over total assets. Highly levered firms have greater probabilities of financial distress, which has been shown be associated with financial misreporting (e.g. Healy and Wahlen, 1999). Alternatively, leverage can have a disciplining effect by either mitigating agency issues between managers and shareholders (Grossman and Hart, 1982), or providing an additional source of external monitoring vis-a-vis debtholders.

Product differentiation is likely related to relative performance evaluation (RPE). Firms with less product market differentiation might naturally have better benchmarks, and therefore, be more prone to RPE, which could pressure some managers to cut corners or misstate earnings to outperform benchmarks (Cheng, 2011). This effect would work against our hypothesis and findings. Thus, to increase the power of our tests, we control for RPE following the work of Wang and Winton (2014) who construct an indicator variable RPE. First, the authors estimate the following regression equation:

$$prob(\text{CEO Turnover}_{i;t \; 1}) = \beta_1 RP^{+}_{i;t} + \beta_2 RP_{i;t} + \epsilon_{i;t} \quad (1)$$

[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).

[8]In unreported analysis, we use an alternative proxy for equity finance needed (EFN) defined by Demirg-Kunt and Maksimovic (1998) as ROA/(1-ROA), which measures a firms asset growth rate in excess of the maximum internally financeable growth rate. We find qualitatively similar results.

where $RP^{+}_{i;t}$ is equal to relative performance when relative performance is above 0, and zero otherwise; and $RP_{i;t}$ is equal to relative performance when relative performance is below 0, and 0 otherwise. Relative performance is measured as the di erence in performance between rm i and the weighted average of rm is rivals according to its three-digit SIC code. Following Wang and Winton (2014), we estimate equation (1) separately for each industry (three-digit SIC) and de ne the binary variable RPE equal one for industries where $\hat{}_{2} < 0$. We refer to speci cations that include all our control variables as the full set of controls.

Table 1 provides the number of observations, mean, standard deviation, 10th percentile, and 90th percentile value for our control variables. We estimate all speci cations for both features of competition from industry-level characteristics in a regression framework.

The TNIC approach also improves upon some basic inaccuracies of other classification schemes. For example, the Coca-Cola Company and PepsiCo are not considered competitors according to their four-digit SIC code, or their Fama-French 48 industry classification, but have a high similarity score (80th percentile). Furthermore, TNIC industry classifications are updated annually, which provides more flexibility and accuracy in empirical design. 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 product similarity score (from 0.035 to 0.012). However, the divestment from Exxon was not reflected by a change in its SIC code or other industry classifications. As a result, the level of competition that Exxon faced according to SIC code-based Hirschman-Herfindahl Index (HHI) measures did not change in response to its large divestment. The measurement error imposed by traditional competition measures can bias results and limit the power to detect existing relationships between fraud and various aspects of competition.

Using the TNIC competitor classification and product similarity scores, we create our main variable of interest; Average Similarity Score, as the average pairwise similarity score of all competitors within a firms TNIC-3 classification in each year. As shown in Table 1, the firms in our sample have 49 competitors on average, with an Average Similarity Score of 0.03 above the threshold set by Hoberg and Phillips (2016).

A potential issue with the Average Similarity Score based on the TNIC classification is that it only includes firms over a certain threshold of similarity. While imposing this threshold allows us to focus on closely-related rivals, there can be substantial variation in the number of competitors being averaged across for each firm. The wide variation in both the number of competitors each firm has and the degree of similarity with each competitor, can obfuscate the association between fraud and product differentiation. Two firms, for example, could have the same average product market similarity scores for different reasons. One firm could have several moderately close rivals, while another firm could have a mix of some nearly identical rivals and some that are barely related. While both firms could have the same average product similarity score, we would expect the firm with the near identical rivals to provide more precise information about a firm's competitive landscape and factors affecting performance.

To address such concerns, we implement a series of alternate methods for aggregating product similarity scores. Rather than averaging across all competitors in a firm's TNIC, we average across the top 5, 10 or 15 closest competitors. This process creates more homogeneity by utilizing the same number of competitors for each firm and focuses on each firm's closest rivals, which should provide the greatest information externalities. As an alternative approach, we count of the number of competitors each firm has that are in the top percentile (95th, 90th and 75th) of the overall distribution of similarity scores across all firms in the sample. This process allows us to count the number of rivals that each firm has that are very similar relative to the complete cross-section of firms.

Additionally, we develop a measure that emphasizes the degree of similarity between rivals. In particular, rivals provide signals about similar firms, with greater similarity between two rivals producing a less noisy signal. It follows that both the similarity with a given rival, as well as the number of rivals, impact the total signal provided by a firms product market competitors. If signal noise is normally distributed, then there is an inverse squared relationship between product market similarity and the quality of the signal. We define a measure of precision as:

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

where N_i is the number of competitors in firm is TNIC, and $score_{i;j;t}$ is the product similarity score between firm i and competitor j in year t.[10] Higher precision is indicative of a greater signal provided by a firms product market rivals. Rather than an average, we also create a sum of the similarity scores, which would be higher when a firm has more rivals that are more similar.

Table A.2 of the Internet Appendix reports the correlations for our main measure, Average Similarity Score, and the alternative similarity score measures noted above. These measures are highly correlated with each other and with the main measure that averages across all competitors (around 75%), which mitigates concerns regarding the distribution of

IV Empirical Results

IV.1 Corporate Fraud and Product Market Differentiation

IV.1.1 Product Differentiation

In this section, we discuss results from firm-level regressions that examine the association between corporate fraud and product market differentiation. We first explore associations in a standard panel data framework before exploring an instrumental variables approach.

We report OLS estimates for the association between average product similarity score (Average Similarity Score) and corporate fraud in Table 2.[11] The firm-year is the unit of observation in all reported specifications in this section. The specification in Column 1 only includes year fixed effects. Column 2 includes the natural log of total assets (Ln Assets) as well as the Dechow set of controls (i.e. accruals, change in accounts receivable (ΔAR), change in Inventory (ΔInventory), the percentage of soft assets (% Soft Assets), change in cash sales (ΔCash Sales), change in ROA (ΔROA), change in employees (ΔEmployees), and a dummy for security issuance (Security Issue)).

In Column 3, we also include R&D, a dummy for positive R&D, the natural log of the number of analysts (Ln number analyst), Institutional Ownership, the Whited-Wu Index, Industry Stock Return R-squared, a flag for relative performance evaluation (RPE flag), and the number of competitors based on TNIC classification (TNIC NCOMP). Including Institutional Ownership results in a large drop in the number of observations and does not appear to have a meaningful effect on the detection of fraud. Furthermore, inclusion of Institutional Ownership only seems to intensify the relationship between fraud and Average Similarity Score. Considering these issues, we drop Institutional Ownership from the remaining specifications. We also exclude the RPE flag due as it is an industry level measure, which is absorbed by the industry fixed effects. Thus, the specification from Column 4 is our primary specification throughout the remainder of our analysis. Henceforth, we refer to the specification of control variables in Column 4 as our Full set of control variables. All explanatory variables are lagged by one year.

The granularity of our data enables us to control for unobserved heterogeneity at the industry and industry-year level. The specification in Column 4 includes industry (three-digit SIC code) and year fixed effects, and the specification in Column 5 includes industry-year fixed effects. The inclusion of fixed effects improves upon existing studies that are typically

11

unable to account for unobserved industry heterogeneity because the variables they deploy are often constructed at the industry level. In particular, inclusion of industry or industry-year xed e ects accounts for pervasive di erences in the propensity to commit fraud across

all rival rms issuing an IPO in year t. Already public rivals in year t are excluded from this calculation. Next, we create a variable, Num Competitor IPO, that is the total number of rivals rm i competes with that underwent an IPO in year t. We control for the number of rival IPOs to help isolate the e ect due to changes in Average Similarity Score, rather than the extent of rival IPO activity. For robustness we also create an indicator variable, Competitor IPO, that is equal to 1 if any of a rms rivals underwent an IPO in year t, and 0 otherwise. On average, there are 3.2 rival IPOs per rm-year in our sample, with a median of 0 (43% of rms have at least one rival IPO). Contingent on having at least one IPO rival, each rm has an average of 7.6 rivals launching IPOs.

In our rst stage results, reported in Column 1-3 of Table 3, we nd a strong positive relationship between Rival-IPO Similarity and Average Similarity Score. The positive sign indicates that rivals undergoing an IPO that are more similar to rm i, increases rm is overall similarity score, on average. The smallest F-statistic that we observe in the rst stage is 35.52 (5.96^2) and all others are above 70.96 (8.42^2). These F-statistics are all substantively larger than 10 (the typical rule of thumb threshold), so it does not appear that we have a weak instrument problem. The reported t-statistics are calculated using standard errors clustered at the three-digit SIC code (SIC3).

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 Average Similarity Score. We nd strong corroborating evidence that the incidence of corporate fraud is signi cantly lower for rms operating in less di erentiated product markets. In particular, the coe cient estimates range from 0.495-0.670 across all speci cations, suggesting a consistent and economically meaningful e ect. Importantly, these ndings persist with the inclusion of industry xed-e ects in Column 6, which further helps to mitigate endogeneity concerns.

The coe cient estimates in the IV analysis are roughly twice as large as the OLS coef- cients. The larger estimates could imply that there are omitted variables working against our observed e ect in our initial analyses, and that the actual impact of product market di erentiation on fraud is indeed larger than our initial estimates suggest. Alternatively, the larger coe cient estimates could be capturing a local average treatment e ect. That is, the larger partial e ect could be concentrated in rms with rival IPO activity. However, the estimates are in line with those of the top quartile in our complexity analysis presented in Table 6. While we cannot entirely rule out the potential for omitted variables to jointly determine a rms fraudulent reporting and the similarity score of rivals who undergo IPOs, the IV results are suggestive of a causal relationship between product di erentiation and corporate fraud.

IV.1.3 Alternative Measures of Competition

In this section, we illustrate that product differentiation captures a particular aspect of competition not explained by traditionally-used measures. Our alternative measures of competition include those widely utilized in prior literature, such as: HHI (Hirschman, 1945; Herfindahl, 1950), profit margin (Bain, 1951), and the sales concentration ratio of the largest four firms in an industry (Hefflebower, 1957).

The HHI based on SIC code is the most extensively used measure of competition in studies related to product market competition. The HHI for industry j is calculated as:

$$HHI_j = \sum_{i=1}^{N_j} (MS_i)^2 \tag{3}$$

Where MS_i is the sales-based markets share of firm i in industry j, and N_j is the number of firms in industry j. HHI has a maximum value of 1 that is attained if a single firm makes up an entire industry, and a minimum value of $1/N_j$ if each firm has equal weight in industry j. HHI was originally designed to measure concentration in the U.S. steel industry, a relatively homogeneous industry. Thus, this measure can better capture the competitive landscape where industries are well defined (e.g. Faccio and Zingales, 2017). It is less useful, however, in instances where firms have diversified baskets of differentiated products and are therefore more difficult to delineate. To allow for more accurately defined product markets, we also create a TNIC based Td1(based)-435(Td1(based)-Td [(.)-421(3 0n1(base(theate)-J/F21 5.6(66e193

alternative measures are all designed to capture the degree of competition in an industry in various ways, our results suggest that there is something unique about the relation between fraud and product market similarity.

IV.1.4 FedDtsv Cisn

One concern with empirical studies on fraud is that only detected fraud rather than all committed fraud is observed (Dyck et al., 2013; Dimmock and Gerken, 2012). The empirical measures of fraud captures the joint outcome of a rm committing fraud and being caught. This can make studies di cult to interpret because partial e ects can be due to changes in the probability of detection and / or changes in commission rates. Lower product market di erentiation could enhance any combination of the two pairs: detection or evasion and / or pressure or discipline.

If product market di erentiation decreases the amount of fraud committed or decreases detection rates without impacting commission rates, we would observe less detected fraud for highly di erentiated rms. However, we nd the opposite, that rms with higher product market di erentiation have higher rates of detected fraud. This indicates that managers of di erentiated rms either are less likely to be caught or engage in less fraud. In our opinion, it is di cult to ascertain a plausible explanation for why the presence of similar rivals would decrease outsiders ability to detect reporting manipulations. On the other hand, lower product market di erentiation with competitors should be more informative regarding common shocks to production costs and demand (e.g. Tirole, 2010) which govern rm performance and nancial reporting incentives. This argument is consistent with evidence suggesting that benchmarking informs boards regarding CEO ability (Murphy, 1986) as well as market- and industry-wide conditions when determining CEO pay (Oyer, 2004). Thus, because we do not believe higher di erentiation could make it easier to detect fraud (it should have no e ect or a positive e ect on detection) and because we nd more detected fraud among more di erentiated rms, the only possibility is that managers of di erentiated rms commit more fraud.

The next important question is why di erentiated rms commit more fraud. There could be a direct impact of di erentiation on fraud commission or an indirect impact through increased detection probability. Our hypo Td [(and)-r08l681(n)-448(indirect)-447(impa47(48(indirect)-4

As a rst step to provide evidence for the indirect channel, we estimate a bivariate probit model employed by Wang (2011). This is a latent variable model that aims to exploit the timing di erences in detection and commission (with commission being prior to detection). The model solves two simultaneous probit speci cations and achieves identi cation through exclusion restrictions: namely, that some variables are only associated with detection while others are only associated with commission. Following Wang (2011) we include Relative Performance Evaluation, 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 Dummy only in the detection regression. All other controls are included in both regressions. In Table A.3 of the Internet Appendix, we report coe cient estimates from the partially observable bivariate probit model, P(Z=1) = P(F=1)P(D=1|F=1). This table provides evidence that Average Similarity Score (as well as top 5, top 10, and top 15 Average Similarity Score) are strongly associated with a decline in fraud commission and weakly related to an increase in fraud detection. These ndings are consistent with the indirect channel, that managers understand they are more likely to get caught if they have close benchmarks and respond by committing less fraud.

As an additional test on this front, we exploit the granularity of the data at the competitor-pair level (pairwise observations) and examine the incidence of fraud, conditional on a similar product market rival getting caught. We report results from this analysis in Table A.6 of the Internet Appendix. The ndings indicate that a rm is more likely to be accused of fraud if a rival was recently charged with fraudulent reporting practices. This nding is also consistent with prior work indicating that fraud occurs in industry waves (Povel et al., 2007; Wang et al., 2010) or that there is contagion in nancial misconduct Dimmock et al. (2018).

IV.2 PtdDi teadDlpsCha

Thus far, we have documented a strong mitigating e ect of product market similarity (lack of di erentiation) on corporate fraud. In this section, we explore two primary channels through which this discipline can originate: managerial slack and benchmarking. We then discuss potential alternatives explanations.

Economists have long argued that product market competition can impose discipline by reducing managerial slack (Machlup, 1967). For example, concept to corporate fraud, competition potentially reduces the economic profits that may be extracted through financial reporting manipulation.[16]. 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.

Second, information conveyed by close product market rivals can yield insight about common shocks to production costs and demand, enabling more precise signals of firm-specific performance (Holmstrom, 1982; Nalebuff and Stiglitz, 1983). Along these lines, evidence suggests that closer benchmarks inform boards regarding CEO ability (Murphy, 1986) as well as market-and industry-wide conditions when determining CEO pay (Oyer, 2004). Further, public firms provide a large amount of information through disclosures, which reduces uncertainty (Badertscher et al., 2013). Through a similar process, rivals with significant product market overlap can facilitate monitoring for investors, regulators, and auditors by providing contexts to interpret financial statement (e.g. Hart, 1983; Dyck et al., 2010). For instance, a survey of CFOs by Dichev et al. (2013) indicates that comparability between rival firms is an important means for identifying financial reporting abnormalities.

IV.2.1 Tariffs

To explore whether the discipline effect of product market similarity is driven by variation in managerial slack, we exploit large tariff reductions at the industry level. Tariff reductions 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, changes in tariff rates directly affect the short- and intermediate-term ability of foreign rivals to offer competitive prices. Changes in tariffs, however, do not 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 short-term and intermediate effects of changes in competition that are likely independent of the benchmarking channels.

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. The duties and customs value are collected from the U.S. International Trade Commission. We then 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)Tariff shock is an indica-

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

tor variable that takes value of 1 if the 4-year percentage change in tariff rate is the bottom (tercile/quartile/quintile), 0 otherwise. [17]

Table 5 shows the tariff results. Each tariff reduction specification with and without average score. For the quartile specification (column 3), there is some evidence that an increase in foreign competition pressures managers to commit more fraud. For the tercile and quintile specifications (column 1 and 5) the coefficient is also positive but not significant. This is likely a power issue as this industry level measure trades-off having stronger shocks and not having enough industries. Importantly, when we add Average Score, the coefficient is still negative and significant in all three specifications (columns 2,4,6). This tells us two things. First, Average Score is operating through a different channel than tariff reductions and works in the opposite direction. Thus, even controlling for changes to foreign competition, and thus the levels of managerial slack, managers still face discipline from an alternate channel related to product market differentiation.

IV.2.2 Firm Complexity

To explore the benchmarking channel, we study the disciplining effects of product market similarity and a measure of firm complexity. Cohen and Lou (2012) argue complicated firms require more complicated analysis to impound the same piece of information into the price of a firm with multiple operating segments. It stands to reason that regulators, media, and employees can more easily disseminate information for firms with a simple organizational structure, and are therefore, more likely to detect abnormal performance or financial reporting. Thus, for firms with a very simple organizational structure and product mix, the information provided by having similar rivals (benchmarks) would have a lower marginal effect on outsiders monitoring ability. In contrast, complex firms can be very difficult to understand and detect abnormal behavior without a clear benchmark. Thus, having close rivals for complex firms should intuitively provide a larger marginal effect on the ability to detect earnings manipulations.

All else equal, a firm that operates in several product markets has greater scope to conceal financial information. Operating across a multitude of product markets reduces substantive analytic procedures that auditors can perform and will require more subjective and detailed testing. This notion is reflected in the higher audit fees for firms with many segments (Brinn et al., 1994). For example, a firm that competes in pharmaceuticals, manufacturing, and consumer durables, could hide information by shifting resources across segments or using

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

complex transactions. Furthermore, monitors would need to understand all three industries to confidently detect reporting abnormalities.

As such, we define complexity as the unique number of 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 markets. A higher score on complexity indicates that a firm operates in an environment where rivals are from many different industries, and thus the firm is likely more diversified and has a complex basket of products that compete across several markets. Our measure of complexity builds on the intuition provided by Cohen and Lou (2012) who measure complexity as whether a firm operates in multiple markets.

We split our sample into quartiles according complexity rankings. Then, we estimate our main specification for the relationship between corporate fraud and product market similarity separately for each quartile. The results are presented in Table 6. 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 full set of control variables, described in Section II and in our analysis of Table 2. We estimate regressions separately for each complexity quartile in Panel B.

Consistent with the benchmarking channel, we find that the disciplining effect of product 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 Average Similarity Score for the least complex firms leads to a decrease in propensity of fraud from 1.9% to approximately the other 3 quartiles (tested jointly) and the 3rd quartile partial effect tested independently. Additionally, we find qualitatively similar results in untabulated probit and logit specifications.

IV.2.3 Institutional Ownership and Analyst Coverage

Next, we examine the role that product differentiation plays for two particularly important corporate governance actors with varying degrees of reliance on public financial statements: institutional investors and sell-side analysts. Specifically, we explore whether the marginal disciplining effect of product market similarity is a function of the intensity of institutional ownership and analyst coverage.

The negative effects of corporate fraud on shareholder value can be particularly costly to investors with large ownership stakes, such as institutional investors, thus creating strong

activity by a rms rivals. In particular, both events plausibly lead to a shock to the rms information environment. We rst study the event of IPOs by a rms rivals. These events increase the publicly-available nancial information of previously existing, private competitors, which in turn, enhances the ability to assess, compare, and scrutinize a rms own nancial statements. Consistent with this view, Bauguess et al. (2018) provide evidence that IPOs lead to in an increase EDGAR tra c for rival rms that are already publicly traded. Next, we study acquisitions by a rms rivals. Acquisitions are material events that can draw considerable scrutiny from investors, analysts, regulators and the media, thus increasing the saliency of existing information in the industry. For instance, acquisitions often occur in waves, suggesting an increase in attention for other rms that could potentially be involved in a deal (Song and Walkling, 2000).

For the IPO tests, we take each pairwise observation of competitors, i and j, and ag whether rm j underwent an IPO in year t. We then aggregate the data to the rm-year level for rm i, counting the number of rivals that underwent an IPO in year t. For robustness we also de ne a dummy variable (Competitor IPO) equal to 1 if any of a rms rivals underwent an IPO in year t, and 0 otherwise. There are 3.2 rival IPOs per rm-year in our sample, with a median of 0 (43% of rms have at least one rival IPO). Contingent on having at least one IPO rival, each rm has an average 7.6 rivals undertaking IPOs, which is consistent with the documented evidence that IPOs occur in waves (e.g. Lowry and Schwert, 2002).

A competitors IPO is a shock to competition via two channels. First, as discussed, more information about economic conditions becomes publicly available for the rival, as well as propensity. The positive e ect of rival rm IPOs on fraud suggests a shock to detection. In particular, IPOs by rivals change the available information for comparison rather abruptly, before a rm has time to fully wind down nancial misconduct.

The coe cient estimate for the interaction term is negative (lower for rms with greater pre-existing product market similarity). This nding suggests that the increased detection resulting from rival IPOs is signi cantly more pronounced for rms with more product di erentiation prior to the rivals IPO (i.e., the e ect is more pronounced for ex ante undisciplined rms).[18] We also split the sample between pre-IPO year high and low Average Similarity Score rms to verify that the IPO-detection e ect is greater for rms that had lower ex ante in Section III, we implement alternative constructions of using product market similarity to ensure our results are not driven by our main construct.

In Panel A of Table 9, we re-estimate our main specification with each of the alternative aggregation schemes. Given the high correlation between these measures (shown in Table A.2 of the Internet Appendix), it is unsurprising that all variations yield a highly significant negative relationship with fraud. These results mitigate concerns that the equal weighted average potentially obscures the association between fraud and product differentiation. To facilitate comparison the economic magnitudes of the various measures, we estimate a specification in which we standardize all variables and report the results Panel B. The coefficient estimates exhibit monotonicity based on the number of competitors used to compute each firms the particular model specifications that we have chosen, we have explored many variations and failed to find an association in any of the variations that we tried.[20] Overall, this analysis increases our confidence that we have identified an economically meaningful link between fraud and product market similarity in our primary analysis.

V Conclusions

Our paper examines the relationship between an individual firm's competitive landscape and the incidence of corporate financial fraud. Empirically examining such associations is challenging due the multifaceted nature of product market competition, which leads to difficulty in both capturing firm-level, rather than industry-level, values and picking up the aspects of competition that likely play the largest role in the detection and commission of fraud. As the degree of competition is predicated on the substitutability, we use pairwise product market similarity scores developed by Hoberg and Phillips (2010, 2016). This firm-specific measure is based on 10-K descriptions of public rivals, and thus captures publicly-available information about a firm's differentiation.

We find that firms with lower product market differentiation exhibit significantly lower incidences of fraud. The economic magnitude of product market similarity is large compared to many other factors that have previously been explored in the fraud literature. We corroborate these findings using an instrumental variables analysis using rival firm IPOs, which helps mitigate concerns about selection effects and industry heterogeneity.

To ascertain whether the publicly available information about product market rivals captures unique informational content of that is related to fraud, we explore a battery of other measures of competition. These include both traditional measures that have a strong industry component and newer measures that encapsulate intensity of competition, but less about the information gleaned from rivals. Individually, these measures have little predictive power on the incidence of corporate fraud, and using such measures as control variables does managerial discipline by reducing managerial slack. To differentiate these channels, we

Rfs

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

Andergassen, R. 2016. Managerial compensation, product market competition and fraud.International Review of Economics & Finance 45:1{15.

Badertscher, B., N. Shro , 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 Pro t Rate to Industry Concentration: American Manufacturing, 1936-1940. The Quarterly Journal of Economics 65:293{324.

Balakrishnan, K., and D. A. Cohen. 2013. Competition and Financial Accounting Misreporting. Working Paper .

Bauguess, S. W., J. Cooney, and K. W. Hanley. 2018. Investor Demand for Information in Newly Issued Securities. 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.

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.

Nalebu , 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.

Oyer, P. 2004. Why Do Firms Use Incentives That Have No Incentive E ects? The Journal of Finance 59:1619{1650.

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.

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 o erings. Journal of Financial Economics 50:101{122.

Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The in uence of large clients on o ce-level auditor reporting decisions. Journal of Accounting and Economics 30:375{400.

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.

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.

Shro , N., R. S. Verdi, and B. P. Yost. 2017. When does peer information matter. Journal of Accounting and Economics .

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

Song, M. H., and R. A. Walkling. 2000. Abnormal returns to rivals of acquisition targets: A test of the `acquisition probability hypothesis'. Journal of Financial Economics 55:143{171.

Teoh, S. H., I. Welch, and T. Wong. 1998. Earnings Management and the Long-Run Market Performance of Initial Public O erings. The Journal of Finance 53:1935{1974.

Thompson, R. B., and H. A. Sale. 2003. Securities Fraud as Corporate Governance: Re ections upon Federalism. Vanderbilt Law Review

Wang, T. Y., and A. Winton. 2014. Product Market Interactions and Corporate Fraud. Working Paper .

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.

Table 1: Summary Statistics

This table reports summary statistics of rm characteristics at the rm-year level. Variable de nitions are provided in the Appendix. Our sample spans 1996 through 2011.

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
Avg Similarity Score	55,381	0.030	0.023	0.012	0.055
Avg Top5 Similarity	55,381	0.080	0.058	0.017	0.156
Avg Top10 Similarity	55,381	0.066	0.050	0.014	0.135
Avg Top15 Similarity	55,381	0.059	0.047	0.013	0.123
Avg Score 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 Pro t 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

Table 2:

Table 3: Shock to Product Market Di erentiation from Rival IPO

This table reports 2SLS estimates for the relationship between product market similarity and corporate fraud. In columns 1-3, we report the rst stage result for 2SLS regression namely the relationship between a rm's IPO-rival's lagged similarity score on the rm's overall average score. In columns 4-6, we use similarity scores with competitors undergoing an IPO as an instrument for the rms Average Similarity Score on fraud. Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. T-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

Table 4: Product Market Di erentiation and Corporate Fraud
(Controlling for Alternative Measures of Competition)

This table reports OLS estimates for the incidence of fraud on Average Similarity Score, while controlling for alternative 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 Her ndahl-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 pro t margin and an industry concentration measure. In Column 4 we include the sales based HHI according to the rm's TNIC. Column 5 also includes the number of competitors within a rm's TNIC. Column 6 also includes the sum similarity score. 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. The speci cations include the full set of controls as described in Section II. All speci cations are run at the rm-year level, include year and SIC3 xed e ects, and explanatory variables lagged by one year. The t-statistics, calculated from standard errors clustered at SIC3 level, are reported in parentheses below coe cient estimates Statistical signi cance (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
Avg Similarity Score	-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)
Controls	Full	Full	Full	Full	Full	Full	Full
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
Cluster	sic3	sic3	sic3	sic3	sic3	sic3	sic3

Table 5: Product Market Di erentiation, Tari Reductions and and Corporate Fraud

This table reports OLS estimates for the incidence of fraud on the average similarity of each rms rivals combined with large industry level tari reductions. Our proxy for corporate 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 nd year-SIC3 level tari levels and identify large year over year tari reductions. Columns 1 and 2 use the top96S] the3(2)-336(useLu5(the)-303p96S])3l

Table 6: Product Di erentiation and Fraud by Complexity Quartiles

This table reports OLS estimates for the incidence of fraud on the average similarity of each rms rivals split into complexity quartiles. Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. We de ne complexity as the number of unique SIC codes spanned by a rms set of competitors according to the TNIC developed by Hoberg and Phillips, 2016. Panel A reports competitor and fraud classi cations for each quartile. Panel B reports OLS estimates for each quartile including our full set of control variables described in Section II. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

Complexity	Low			High
	Q1	Q2	Q3	Q4
Panel A				
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
Avg Similarity Score	2.7	2.8	3.3	3.5
Panel B				
Avg Similarity Score	-0.168***	-0.197**	-0.201	-0.683***
	(-3.418)	(-2.029)	(-1.508)	(-5.053)
Observations	9,995	9,628	9,018	8,503
R-squared	0.016	0.017	0.024	0.026
FE	Year	Year	Year	Year
Controls	Full	Full	Full	Full

Table 7: Product Di erentiation and Fraud by Institutional Ownership and Analyst Coverage Splits

This table reports OLS estimates for the incidence of fraud on the average similarity of each rms rivals split into groups based on Institutional Ownership HHI and Analyst Coverage. Our proxy for corporate 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 including our full set of control variables described in Section II. Panel B contains the same but splits by quartiles based on the number of sell-side analysts covering a rm in a given year and whether that rm is covered by at least one star analyst or not. Panel's C and D repeat the analysis in Panel B but look only at observations either before or after REG FD in 2000. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	Panel A Institutional Ownership HHI					
	Low			High		
	Q1	Q2	Q3	Q4		
Avg Similarity Score	-0.425***	-0.298***	-0.107**	-0.066		
	(-3.373)	(-3.145)	(-2.175)	(-1.217)		
Observations	7,943	8,225	8,085	7,153		
R-squared	0.023	0.022	0.012	0.007		
	Panel B Analyst Coverage					
	Low			High	Non Star	Star
	Q1	Q2	Q3	Q4		
Avg Similarity Score	-0.001	-0.204**	-0.456***	-0.487***	-0.133*	-0.388***
	(-0.014)	(-2.417)	(-3.940)	(-4.378)	(-1.737)	(-3.985)
Observations	7,742	7,179	6,420	6,450	12,944	7,221
R-squared	0.006	0.014	0.038	0.028	0.011	0.028
	Panel C Analyst Coverage - Pre 2000					
Analyst Coverage	Low			High	Non Star	Star
	Q1	Q2	Q3	Q4		
Avg Similarity Score	0.058	-0.267**	-0.786***	-0.581**	-0.336**	-0.508*
	(1.037)	(-2.170)	(-3.296ilarity Sc4p)		(-3Sc442051	(-1.7374051

Table 8: IPOs and Acquisitions of Rivals as Change to Information Environment

This table reports OLS estimates for the association between fraud and rival IPOs (M&A) activity including our full set of control variables described in Section II.The speci cations include rival rm IPO activity in columns 1 and 2 and M&A activity in panels 3 and 4. We split the data by high and low non-IPO (non-acquired) similarity scores in year t-1. All speci cations include year xed e ects and three-digit SIC code (SIC3) xed e ects, and all control variables are lagged one year. The t-statistics, calculated from standard errors clustered at the SIC3 level, are reported in parentheses below coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)
	Low Non-IPO Rival Score Fraud	High Non-IPO Rival Score Fraud	Low Non-M&A Rival Score Fraud	High Non-M&A Rival Score Fraud
Ln Num Competitor IPO	0.005***	0.004		
	(2.820)	(1.440)		
IPO Size ($)	0.000	0.000		
	(-0.02)	(-0.72)		
Ln Num Competitor Target			0.103***	0.000
			(3.141)	(-0.024)
Ln Target MarketCap			-0.008***	0.002
			(-2.829)	(1.380)
Observations	18858	18279	18672	18449
R-squared	0.05	0.037	0.052	0.039
Controls	Full	Full	Full	Full
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3

Table 9: Product Market Di erentiation and Corporate Fraud
(Alternate Constructions of Independent Variable)

This table reports OLS estimates for the incidence of fraud on alternative constructions of our primary independent variable. Our proxy for corporate 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 Average Similarity Score with a rms product market similarity score averaged across its closest 15, 10, and 5 competitors, respectively. In Column 5, we replace Average Similarity Score with the Precision measure outlined in section III. The unit of observation in this analysis is the rm-year. All speci cations include the full set of controls as described in Section II. They include year and SIC3 xed e ects, 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 coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)
	Fraud	Fraud	Fraud	Fraud	Fraud
Panel A OLS Regression					
Avg Similarity Score	-0.171***				
	(-3.946)				
Avg Top 15 Similarity		-0.089**			
		(-2.020)			
Avg Top 10 Similarity			-0.091**		
			(-2.244)		
Avg Top 5 Similarity				-0.086***	
				(-2.730)	
Avg Score 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 Regression					
Avg Similarity Score	-0.004***				
	(-3.946)				
Avg Top 15 Similarity		-0.004**			
		(-2.020)			
Avg Top 10 Similarity			-0.005**		
			(-2.244)		
Avg Top 5 Similarity				-0.005***	
				(-2.730)	
Avg Score 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
Cluster	sic3	sic3	sic3	sic3	sic3

Table 10: Product Similarity and Restatements
(Falsi cation Test)

This table reports OLS estimates for the incidence of accounting restatement on the average similarity of each rms rivals. Our primary dependent variable accounting restatement (RE AA) is obtained from Audit Analytics database from 1996-2012. We do not perform any screens for our restatement variable. The speci cation in Column 1 does not include control variables. The speci cation 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 II and Column 3 also includes Institutional Ownership. In Column 6 we report the standardized regression. All speci cations are run at the rm-year level, include year xed e ects, and include explanatory variables are lagged by one year. Column 4 also includes three-digit SIC code (SIC3) xed e ects, Column 5 adds year SIC3 xed e ects. In Column 6, we run the speci cation from Column 4 but with standardized regressors.

Table 11: Alternative Measures of Competition and Corporate Fraud

This table reports OLS estimates for the incidence of fraud on commonly used industry-level proxies for competition. Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Our measures of competition include: sales based HHI, average pro t margin, top-4 sales concentration and number of competitors constructed using three-digit SIC code, sales based HHI and number of competitors according to TNIC3, product market uidity and the 10-K-based competition word measure. Columns 1-8 include the full set of controls as described in Section II. The rm-year is the unit of observation in this analysis. All speci cations include year and SIC3 xed e ects, and control variables lagged by one year. The t-statistics, calculated from standard errors clustered at the SIC3 level, are reported in parentheses below coe cient estimates. Statistical signi cance (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	(8) Fraud
SIC3 HHI	0.019							
	(0.903)							
SIC3 PM sale		-0.021						
		(-1.453)						
SIC3 Top 4 Concentration			0.035*					
			(1.770)					
SIC3 NCOMP				0.013*				
				(1.785)				
TNIC HHI					0.000			
					(0.008)			
TNIC NCOMP						0.001		
						(0.633)		
prodmkt uid							0.001	
							(0.951)	
Competition 10K								0.004
								(1.065)
Observations	37,144	37,144	37,144	37,144	37,144	37,144	35,999	18,696
R-squared	0.033	0.034	0.034	0.034	0.033	0.033	0.034	0.048
Controls	Full	Full	Full	Full	Full	Full	Full	Full
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3
Cluster	sic3	sic3	sic3	sic3	sic3	sic3	sic3	sic3

IteoA jpxfb

PtdDi ta Ifpnd

Fed

Table A.1: Variable Description

Variable	De nitions
AAER Misstatement	Equal to 1 for rm-years for which rms 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 rm-years for which rms settle a securities class action lawsuit for an alleged 10B-5 fraud allegation.
Fraud	Equal to 1 for all rm-years with an AAER or SCA.
SIC3 HHI	Her ndahl-Hirschman index based on rm sales and three-digit SIC code industry classi cations
TNIC HHI	Her ndahl-Hirschman index based on rm sales Text-based Network Industry clas-si cations (TNIC) from Hoberg and Phillips.
Avg Similarity Score	Mean Hoberg and Phillips Similarity Score for all rivals within each rm-years TNIC
Precision	De ned as $(\frac{1}{NCOMP_{TNIC}} \sum \frac{1}{(1 - score)^2})^{\frac{1}{2}}$
Pro t Margin	Average EBITDA/sales ratio for rms within each three-digit SIC code
Top 4 Concentration	Proportion of sales within a three-digit SIC code attributable to the four largest rms within an industry
Age	Number of years the rm has been in Compustat
Analyst Num	Number of analysts covering the rm 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
Industry	

Table A.2: Correlations

Correlation coe cients are reported for various measures of product market similarity and competition. Our sample covers 1996 through 2011.

	Avg Sim. Score	Top 15 Similarity	Top 10 Similarity	Top 5 Similarity	Sim. Precision	Sum Similarity	SIC3 HHI	SIC3 PM	SIC3 NCOMP	TNIC HHI	TNIC NCOMP	Product Fluidity	10-K Competition
Avg Similarity Score	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									
Sim. 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 Pro t 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		
Product Fluidity	0.201	0.47	0.468	0.437	0.206	0.515	-0.222	-0.186	0.246	-0.308	0.518	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.3: Product Market Differentiation and Corporate 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 (2014). In specifications 2-4, we replace Average Similarity Score with the average from 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)
Avg Similarity Score	-7.378*	4.221						
	(-1.938)	(0.497)						
Avg Top 15 Sim.			-3.209**	4.603				
			(-2.385)	(1.582)				
Avg Top 10 Sim.					-2.970**	4.111		
					(-2.444)	(1.580)		
Avg Top 5 Sim.							-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 dummy	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)

Table A.4: Product Di erentiation and Corporate Fraud
(Alternate Constructions of Independent Variable)

Table A.5: Product Market Di erentiation and Financial Statement Comparability

This table reports estimates for the incidence of fraud on the average similarity of each rms rivals using ordinary least squares (OLS) regressions adding in the output-based measure of accounting comparability from De Franco et al. (2011). Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. Panel A presents the mean Average Similarity Score and % Fraud for rm-years with below and above median Accounting Comparability and the correlation between the two measures. Panel B includes regressions

Table A.6: Rival Fraud, Product Di erentiation, and Fraud Detection

This table reports estimates for the incidence of fraud on the average similarity of each rms rivals and rival rm fraud activity using ordinary least squares (OLS) regressions. Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from

Table A.7: IPOs and Acquisitions of Rivals as Change to Information Environment

This table reports OLS estimates for the association between fraud and rival IPOs (M&A) activity. The speci cations are the same as model (4) of Table 2, but also include rival rm IPO (M&A) activity. For each rm-year, include the natural log of the number of rms that compete with rm i and that underwent an IPO or were acquired in year t, and an interaction term Ln Num Competitor IPO Avg Similarity Score or Ln Num Competitor Target Avg Similarity Score. In Column 2 and 5, we control for IPO (M&A) Size ($) which is the sum of all-capital raised by IPO rivals (total market capitalization of Target rivals). All speci cations include year xed e ects and all control variables are lagged one year. Columns 3 and 6 also include three-digit SIC code (SIC3) xed e ects. The t-statistics, calculated from standard errors clustered at the SIC3 level, are reported in parentheses below coe cient estimates. Statistical signi cance (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
Avg Similarity Score	-0.149***	-0.149***	-0.134***	-0.185***	-0.185***	-0.162***
	(-3.488)	(-3.473)	(-3.382)	(-4.060)	(-4.061)	(-3.927)
Ln Num Competitor IPO	0.009***	0.009***	0.008***			
	-4.767	-4.909	-4.565			
Av Score Ln Num Comp IP	-0.119***	-0.123***	-0.109***			
	(-3.370)	(-3.460)	(-2.660)			
IPO Size ($)		0.000				
		(-0.498)				
Ln Num Competitor Target				0.052***	0.061**	0.048***
				-3.373	-2.436	-3.38
Avg Score Ln Num Comp Target				-0.844***	-0.874**	-0.685**
				(-2.693)	(-2.593)	(-2.430)
Ln Target MarketCap					-0.001	
					(-0.588)	
Observations	37144	37144	37144	37144	37144	37144
R-squared	0.017	0.017	0.034	0.017	0.017	0.035
Controls	Full	Full	Full	Full	Full	Full
FE	Year	Year	Year+Sic3	Year	Year	Year+Sic3

Table A.8:

Table A.9: IPO Rival Similarity as a shock to product market di erentiation

This table reports 2SLS estimates (second stage) for the relationship between product market similarity and corporate fraud. In columns 1-4, we report the rst stage result for 2SLS regression. In columns 5-8, we use similarity scores with competitors undergoing an IPO (being acquired) as an instrument for the rms Average Similarity Score (Avg Top 15/10/5 Similarity). Column 1 and 5 only includes subsample with number of competitor IPO > 0. Our proxy for corporate fraud includes a combination of AAER misstatements from the AAER dataset and Securities Class Actions from the Stanford University Lawsuit Database. The t-statistics, calculated from standard errors clustered at the three-digit SIC code (SIC3) level, are reported in parentheses below coe cient estimates. Statistical signi cance (two-sided) at the 10% 5%, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Avg Sim. Score	Avg Top 15 Sim.	Avg Top 10 Sim.	Avg Top 5 Sim.		Fraud		
IPO Avg Score	0.258***	0.397***	0.430***	0.476***				
	(8.859)	(8.501)	(10.065)	(12.796)				
Ln Num Comp IPO	-0.000	-0.001	-0.002***	-0.004***	0.003	0.003**	0.003*	0.002*
	(-1.133)	(-1.098)	(-2.820)	(-5.414)	(1.092)	(2.042)	(1.951)	(1.763)
IPO Size ($)	0.016***	0.088***	0.071***	0.045**	0.127***	0.105***	0.100***	0.094**
	(3.055)	(5.368)	(4.050)	(2.291)	(3.025)	(2.674)	(2.579)	(2.432)
$\widehat{\text{AvgSim.Score}}$					-0.314*			
					(-1.673)			
$\widehat{\text{AvgTop15Sim.}}$						-0.219*		
						(-1.683)		
$\widehat{\text{AvgTop10Sim.}}$							-0.203*	
							(-1.657)	
$\widehat{\text{AvgTop5Sim.}}$								-0.183
								(-1.619)
Constant	0.003	-0.021***	-0.026***	-0.029***	-0.021	0.009	0.008	0.008
	(1.446)	(-6.495)	(-7.902)	(-7.999)	(-1.079)	(0.666)	(0.614)	(0.589)
Observations	14,823	28,786	28,786	28,786	14,823	28,786	28,786	28,786
R-squared	0.679	0.682	0.681	0.630	0.046	0.040	0.040	0.040
FE	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3	Year+Sic3
Cluster	sic3	sic3	sic3	sic3	sic3	sic3	sic3	sic3

Table A.10: Product Di erentiation and Fraud by Size Quartiles

This table reports estimates for the incidence of fraud on various competition measures using ordinary least

Table A.11:

Table A.12: Arthur Andersen Auditor Tests

For our sample of rms, we identify if that company used Arthur Andersen as its auditor and whether such a rm was detected for fraud after it switched auditors (2002-2004 time period) based on conduct that occurred while it was an Arthur Anderson client (2000 to 2002 time period). High and low Average Similarity Score splits is based on their 1999 values.

	Obs	Detected Fraud	t-stat
Low Avg Score	39	0.231	
High Avg Score	43	0.558	
Di		-0.327***	-3.126

Table A.13: Whistle Blowers

We use the data from Dyck et al. (2010) on the whistleblower types. The whistleblowers for Internal Crime come from people within the rm. The whistleblowers for External Crime are analysts, auditors, clients or competitors, equity holders, industry regulators, law rms, newspapers, the SEC, and the short-sellers. We compare the Average Similarity Score of year rm rst has SCAC case between the two groups.

	Obs	Average Score	t-stat
External Crime	86	0.033	
Internal Crime	31	0.025	
Di		-0.007**	2.047

Table A.14: Product Market Di erentiation and Corporate Fraud - Non-Linear Speci cations