# Do managers tacitly collude to withhold industry-wide bad news?

Jonathan L. Rogers
University of Colorado

Catherine Schrand
University of Pennsylvania

Sarah L. C. Zechman
University of Chicago

That managers would choose to withhold firm-specific bad news is not only intuitive, but supported by theory, observed disclosure patterns, and survey responses. When the bad news is industry-wide, however, explaining withholding as a sustainable equilibrium is more complicated. If any one firm chooses to disclose, the news effectively becomes public, creating incentives for other firms to disclose. Withholding is only sustainable if all firms cooperate ("tacitly collude"), which depends on their own incentives, and their conjectures about the incentives of other firms in the industry to cooperate. We document cases of increased intra-industry obfuscation in the annual 10-K, controlling for changes in fundamentals, consistent with tacit collusion to hide news. Tacit collusion is more likely in industries with more significant equity incentives and greater litigation risk and less likely in industries in which observable/public macro-economic data relevant to firm valuation is available. The results have implications for understanding when market forces are sufficient to generate voluntary disclosure of industry-wide news.

First draft: February 2013

This draft: September 2013

Preliminary and incomplete

The authors thank Paul Fischer, Mirko Heinle, and audiences at Duke, MIT, and Northwestern for helpful comments. The authors thank Sehwa Kim for research assistance with this project. The authors gratefully acknowledge the financial support of the University of Chicago Booth School of Business (Jonathan Rogers and Sarah Zechman) and the Neubauer family (Sarah Zechman).

## *1. Introduction*

The idea that firms would withhold bad news is intuitive, and theoretical models predict withholding when traders are uncertain about whether the manager has received a signal (e.g., Dye, 1985; Jung and Kwon, 1988; Shin, 2003 and 2006).[1] Firms with relatively adverse news are able to hide in the pool of firms that have no news. But these models assume that firms' adverse private signals are uncorrelated. The withholding prediction in these models does not necessarily translate into a prediction that firms will withhold *industry-wide* bad news, which is by definition correlated across firms in an industry. If at least one firm in the industry receives a firm-specific valuation signal that is higher than its current stock price, despite the adverse industry-wide component, that firm will disclose, and the news effectively becomes public. The uncertainty about news arrival, which generates the partial disclosure equilibrium, disappears and adverse selection should lead to the classic unraveling "full disclosure" equilibrium result (Grossman, 1981; Milgrom, 1981).

We investigate whether and, if so, when firms withhold industry-wide bad news. We assume that withholding industry-wide bad news when another firm in the industry discloses it first imposes externality costs on the non-discloser, such as increased litigation or reputational costs. With externality costs, a withholding equilibrium is not sustainable if the manager views the disclosure decision as a static problem about the disclosure of an adverse signal. However, we add the possibility that managers anticipate receiving revised signals in subsequent periods. Thus, adverse news, if withheld, may never materialize. The consideration of subsequent valuations in evaluating current-period disclosure decisions is realistic. Graham et al. (2005) report that several surveyed CFOs indicate they delay bad news "…in hopes that the firm's status will improve before the next required information release, perhaps saving the company the need to ever release the bad information…".[2] In the anticipated repeated setting, collusion can be a rational disclosure strategy.

The managers' disclosure decisions regarding industry-wide adverse news depend on their conjectures about the probability of new (and improved) subsequent signals of industry-wide news, the impact of the industry-wide news on other firms' valuations, the benefits firms would expect to receive from withholding adverse news (e.g., an elevated stock price, reduced borrowing costs), the expected duration of the period that the news remains hidden, the costs of withholding, and the

---

[1] See empirical evidence that firms withhold bad news in Kothari, Shu, and Wysocki (2009) and survey evidence in Graham, Harvey and Rajgopal (2005). See Sletten (2012) and Tse and Tucker (2010) for evidence of clustering in bad news announcements, which theoretical models predict when firms withhold bad news (Dye and Sridhar, 1995; Acharya et al., 2011).

[2] The outcome that collusion is not achieved in a single-period game, but that it is a possible equilibrium with multi-periods, is analogous to the conclusions in product-market collusion games (Pindyck and Rubinfeld, 1989).

extent to which the conjectures about each of these elements in firms' disclosure decisions are common knowledge. Because the withholding equilibrium requires a reliance on conjectures about the strategic behavior of other firms, we refer to the cooperative behavior that leads to withholding of industry-wide bad news as "tacit collusion." [3]

Understanding whether and when firms withhold industry-wide bad news has important implications for debates about mandating disclosures. Proponents of mandated disclosure argue that firms left to their own devices will withhold private information, in particular adverse news. Opponents argue that disclosure mandates impose unnecessary costs on firms because market forces are sufficient to encourage voluntary disclosure (Easterbrook and Fischel, 1984). Capital market price pressure due to adverse selection, described previously, and the threat of litigation both encourage disclosure. Our analysis brings real data to the question of when these and other market forces are sufficient to encourage voluntary disclosure of industry-wide bad news, which informs the debate on the necessity for mandated disclosures.

In the empirical analysis, we operationalize withholding of bad news by increased industry-level opacity or obfuscation. We expect firms will withhold industry-wide bad news by obfuscating the discussion in their reports rather than excluding it for two reasons. First, extant studies show that termination of previously disclosed information is costly (e.g., Chen, Matsumoto, and Rajgopal, 2011), and second, Reg. S-X requires firms to discuss various topics, but the content is discretionary. Our measure of increased obfuscation is a decrease in the readability of disclosures. Li (2008) establishes the FOG score as a measure of readability, which he uses as a proxy for management obfuscation, and documents an association between firms' FOG scores and both future performance and performance persistence.

We use two methods to identify collusion episodes. We first document periods in which firms within an industry, on average, collectively increase disclosure obfuscation as measured by the FOG score. We identify "collusion episodes" for a given industry as a year with a significant increase in the intra-industry average level of annual report FOG scores after controlling for fundamental firm characteristics that prior literature has shown to be associated with FOG scores (Li, 2008). The measured increase attempts to isolate the discretionary increase in industry-level opacity. The number of industry-year collusion episodes is small. Across the Fama-French 49

[3] As described in Ivaldi et al. (2003), footnote 2: ' "Tacit collusion" need not involve any "collusion" in the legal sense, and in particular need involve no communication between the parties. It is referred to as tacit collusion only because the outcome … may well resemble that of explicit collusion or even of an official cartel. A better term from a legal perspective might be "tacit coordination".'

industry groups over the 14 year period from 1995 to 2008, we find between 12 and 17 collusion episodes (depending on the sample) in which the average increase in intra-industry disclosure opacity suggests tacit collusion. Our second method identifies a subset of these collusion episodes in which at least 60% of influential firms in the industry exhibit positive abnormal FOG levels. This restriction yields between 9 and 12 episodes.

We provide weak evidence that the cumulative adjusted returns following the episodes are significantly more negative for the colluders than for the non-colluders starting in the second year following the collusion episode. This pattern is consistent with the collusion episodes representing periods in which firms in the industry withheld bad news that eventually materialized. The fact that the poor performance comes approximately two years after the disclosure obfuscation is sensible because the duration of the withholding period is predicted to be positively related to the likelihood of withholding. If firms had rationally anticipated that the news would be made public quickly, withholding should not have been a dominant equilibrium. We attribute the weakness of the evidence to the fact that we analyze average *realized* returns. These returns understate the amount of bad news that was hidden if the bad news is either partially mitigated or never realized, either of which is reasonably likely given that the firms in the industry have chosen to withhold the news.

The main analysis in the paper investigates the industry characteristics associated with the identified collusion episodes. Three main results emerge. First, industries in which the news is likely to be public or become public quickly are less likely to collude. We identify such industries based on the incremental explanatory power of observable/public macro-economic signals (i.e., short-term interest rates, default spreads, term spreads, foreign exchange rates, producer price index, and the Fama-French SMB and HML factors) for returns. When news is public, or will become public in the near future, each firm in the industry has little individual incentive to withhold. Also, each firm will conjecture that other firms have less incentive and this will be common knowledge. Thus, firms that withhold will face a higher risk of externality costs given that other firms are more likely to disclose and we observe less collusion.

Second, when the proportion of firms in the industry that trade on a major exchange (i.e., in the CRSP database) is high relative to the proportion filing annual reports (i.e., in the Compustat database), the industry is more likely to experience a collusion episode. We provide two interpretations of this finding. One interpretation is that firms in industries with a higher proportion of publicly traded firms have greater equity incentives to withhold adverse news to maintain an elevated stock price, and this would be common knowledge, making a withholding equilibrium more sustainable. A second interpretation is that firms in industries with a lower proportion of publicly traded firms (i.e., firms file annual reports despite not having publicly traded equity) are subject to significant disclosure regulation, perhaps from regulators or other users that demand financial statements. Assuming mandated disclosure is common knowledge, firms in such industries will conjecture that other firms will disclose, making a withholding equilibrium less sustainable. Supplemental analysis favors the first interpretation – that greater equity incentives are associated with collusion – but the tests are weak due to small sample sizes in the partitioned sample.

Third, industry-level litigation risk is positively associated with collusion episodes. This finding contrasts with our prediction that litigation risk should reduce the likelihood that a withholding equilibrium is sustainable, consistent with the traditional view that firms disclose adverse news to avoid litigation costs (e.g., Skinner, 1994). This result could suggest that our litigation risk proxy captures uncertainty in the industry-level propensity to incur large adverse shocks. Greater uncertainty implies that firms are more likely to believe that future valuation signals may increase (i.e., the news will not be as bad as anticipated or may never materialize), and that this is common knowledge, increasing the sustainability of a withholding equilibrium. Under this interpretation, litigation risk is not a sufficient market force to encourage market disclosure for industry-wide bad news.

We examine two other industry characteristics that we predict affect withholding sustainability with mixed results. We examine whether a withholding equilibrium is more sustainable when an industry is subject to greater uncertainty about its propensity to experience a large adverse event. We use a measure of implied negative expected return skew from option prices ("tail risk") to identify high-risk industries. We do not find evidence supporting an association between high risk industries and collusion.

We also investigate the impact of within-industry heterogeneity in operations, measured as the average value of intra-industry idiosyncratic risk. We find no association with collusion episodes, potentially because of offsetting predicted associations. Industries in which firm returns have a greater idiosyncratic component are less likely to have industry-wide adverse news of a magnitude that will sustain a withholding equilibrium, but intra-industry heterogeneity also makes it less likely that investors will infer industry-wide news from the disclosure by one firm, decreasing the costs to withholding and making a collusive equilibrium more sustainable.

The model includes industry concentration as a control variable and we provide some evidence that collusion episodes are more likely for more concentrated industries. Industry concentration could be related to several fundamental industry characteristics that positively affect collusion likelihood. Greater concentration could increase the probability that firms conjecture that other firms have received a similar negative signal, such that they have individual incentives to withhold. Greater concentration could also be correlated with fewer influential firms in the industry, which provides greater opportunities to form beliefs that the other influential firms will cooperate, perhaps because of explicit collusion or simply greater interpersonal connections, or perhaps because a focal point is more likely available to facilitate cooperation. Greater concentration could also imply more significant existing barriers to entry, which decreases each firm's incentives to disclose bad news relative to industries with low entry barriers.

In section 2 we discuss firms' incentives to withhold industry-wide bad news and the conditions for a collusive equilibrium. We describe our measurement of collusion episodes in Section 3, including the subsequent performance tests meant to provide validation of this measure. Section 4 provides the main analysis of the determinants of the collusion episodes and section 5 concludes.

## *2. Hypothesis development*

### *2.1 Related non-disclosure literature*

The starting point for a discussion of non-disclosure of non-proprietary information is Grossman (1981) and Milgrom (1981) who predict a full disclosure equilibrium. The basic idea is that adverse selection leads each firm with a signal that is greater than the expected value conditional on non-disclosure to report its private signal. As each firm with a private signal above the conditional expected value discloses, the conditional expected value ratchets down and the classic unraveling results. Withholding adverse news is not an equilibrium in their settings.

Subsequent models explained non-disclosure by assuming that managers receive a private signal at an interim period $t = 1$ with probability $< 1$. Traders are uncertain about whether the manager has received a private signal. The manager makes a disclosure decision that maximizes stock price at $t = 1$. The valuations of all firms are revealed at time $t = 2$. The result is an endogenous threshold level below which firms withhold adverse news (Dye, 1985; Jung and Kwon, 1988; Shin, 2003 and 2006). The basic intuition is that firms with relatively adverse news are able to hide in the pool of firms that have no news. Thus, firms with sufficiently adverse news choose to withhold it.

These "threshold" models, however, cannot be used to make predictions about withholding of industry-wide news as it is defined in Section 2.2. If even one firm in the industry receives a firm-specific valuation signal that is higher than its expected valuation in the case of non-disclosure, despite the adverse industry-wide component, that firm will disclose, and the news effectively becomes public. The uncertainty about news arrival, which generates the partial disclosure equilibrium in the threshold models, disappears and adverse selection should lead to the classic unraveling. The result will be full disclosure if the private signals contain only an industry-wide component or a lower disclosure threshold (i.e., less withholding) if the signals contain an idiosyncratic component.

Three voluntary disclosure studies have analyses that relate to industry-wide news. Dye and Sridhar (DS, 1995) analyze disclosure choice when the timing of the arrival of private signals is correlated, but the signal values are uncorrelated. In the special case that they call "industry-wide common knowledge," firms also learn whether other firms in the industry have received a signal. As DS note, if any firm "knows" that another firm will disclose, unraveling and a full disclosure equilibrium will occur (p. 166). Withholding is a possible equilibrium, but only if firms "know" that other firms will not disclose. DS do not analyze disclosure choice when firms must form beliefs about other firms' signals as well as conjectures about whether these beliefs are common knowledge.

Acharya, DeMarzo, and Kremer (ADK, 2011) examine disclosure when signal values are correlated. They operationalize this assumption by assuming firm-specific news is correlated with market-wide news, thus their analysis essentially analyzes disclosure decisions when the news contains a market-wide component. If the market-wide news is already public at the time firms receive their private signals, ADK predict a full disclosure equilibrium because investors are certain that the firm has received a signal. If the market-wide news is not public, such that investors remain uncertain about whether the manager has received a signal, disclosure of bad news is delayed relative to the public information case. When the news is subsequently made public, the release can trigger disclosure by multiple firms that had withheld news. In their model, the subsequent release of the news that makes it public is exogenous. ADK discuss the implications for their equilibrium if the news is made public through the endogenous disclosure by another firm, which is the primary focus of our analysis. However, they only conjecture that the results would hold, stating: "The construction of the equilibrium presents a significant computational challenge." Although ADK's setting has multiple periods, the periods after the signal arrives only provide another opportunity for the manager to disclose. No new signals arrive or are anticipated.

ADK predict a clustering of bad news announcements after a public disclosure and DS predicts disclosure waves following the disclosures by other firms. The clustering occurs because multiple firms had withheld bad news until these events, which then trigger disclosure. Sletten (2012) documents increases in management forecasting following stock price declines associated with a restatement by an industry peer, suggesting that firms withheld bad news, but disclose it when the news becomes favorable relative to the new lower stock price. Tucker and Tse (2010) provide evidence on clustering in warnings about earnings shortfalls. These studies are consistent with firms withholding firm-specific private information, but they do not have implications for whether firms withhold industry-wide bad news.[4]

Jorgensen and Kirschenheiter (2012) analyze disclosure of industry-wide news with an exogenous leader and follower firm. The leader firm anticipates the disclosure decision of the follower. Their analysis focuses on how costs of disclosure, not costs of withholding, affect the leader's disclosure choice. The follower does cannot receive a revised signal after the leader discloses.

In the remainder of this section, we describe when and whether withholding private signals of adverse industry-wide news is a sustainable equilibrium. Section 2.1 describes the setting. Section 2.2 summarizes our hypotheses about when industry-wide news might be withheld and describes the variables we use to test these hypotheses. Section 2.3 provides a summary of related literature on tacit collusion in product markets.

*2.2 The setting*

Each manager of a firm in industry *i* receives a private signal about its firm value. The valuation contains an idiosyncratic component and a common industry-wide component. Although the signals can contain an idiosyncratic component, we refer to such signals as "industry-wide" news. Examples of common components are demand shocks to products sold by firms in the industry or average credit quality shocks to customers in the industry. The assumption that the signal contains a common intra-industry component accords with intuition about within-industry valuations. Firms within an industry are likely to have relatively homogeneous production and cost functions, thus a single piece of news will affect the valuations of all firms in the same direction. This assumption is consistent with the idea that firms' valuations are determined by a common

[4] See also empirical evidence that firms withhold bad news in Kothari, Shu, and Wysocki (2009) and survey evidence in Graham, Harvey and Rajgopal (2005).

industry wide component. O'Brien (1990) shows that investment analysts tend to specialize in one industry, and Dunn and Nathan (2005) show that analysts who focus on one industry produce better forecasts than analysts focusing on multiple industries.

Consistent with a common industry component, if one firm in the industry receives a signal, then all firms in the same industry also receive a signal about their own firm-specific values. Firms observe their private signals but not the signals received by other firms. They know, however, that the other firms have received a signal and are able to form beliefs about the expected values of signals for other firms in the industry.

Traders, however, remain uncertain about whether firms in an industry have received a signal. Traders only learn about the news arrival if a firm, any firm, in an industry discloses its valuation. In that case, traders have expectations about the industry-wide implications and infer that other firms in the industry also received a signal. The common component of the valuation effectively becomes public upon disclosure by any firm.

If a manager receives an adverse industry-wide signal at time 1, she chooses whether to disclose or withhold it. Firms make disclosure decisions contemporaneously. In the empirical analysis, we examine disclosure opacity in annual 10-K filings. While 10-Ks are not released on the same day, there is a large degree of clustering within calendar time for firms within an industry and it is unlikely that one firm would see the release of a peer firm's report and be able to adjust its own 10-K in response before filing. In this sense, we consider firms' annual 10-K disclosure decisions to be contemporaneous.

Traders respond to a disclosure and to silence (i.e., non-disclosure), taking into account the strategic behavior of managers. The stock price of a firm that discloses its valuation will be the disclosed amount, which is the private signal. The stock price of a non-disclosing firm will be the expected value conditional on its non-disclosure and the disclosure decisions of other firms.

When the manager makes her decision to disclose the news or to withhold it, she anticipates that she will receive revised signals in subsequent periods, consistent with survey evidence in Graham et al. (2005). This assumption implies that adverse news may never materialize if it is successfully withheld, in which case the firm will not be subject to stock price fluctuations associated with fundamental uncertainty. This assumption is unique relative to Dye and Sridhar (1995) and Acharya, DeMarzo, and Kremer (2011), described previously. Those models have multiple periods, but the additional periods only represent additional opportunities to disclose the initial signal; revised signals do not arrive.

The self-interested manager makes a disclosure decision that maximizes her utility. We assume the manager's utility is increasing in the firm's stock price in the current period and in the expected stock prices at future dates. The manager's utility is decreasing in expected costs associated with withholding of adverse news, such as increased litigation and reputation costs.

The individual manager's withholding decision is affected by her conjectures about the strategic behavior of other firms. As noted, if any firm in the industry discloses, traders will infer that all firms in the industry received a signal, which in turn, affects the expected conditional value. The uncertainty about news arrival, which generates the partial disclosure equilibrium in the threshold models, disappears and adverse selection should lead each manager to disclose, generating a classic unraveling result.

One might contend that firms could still withhold the news, hoping other firms do not disclose it. We assume, however, that withholding industry-wide bad news when another firm in the industry discloses it first imposes externality costs on the non-discloser that could decrease firms' willingness to follow this strategy. [5]

Litigation costs are one example of an externality cost. One element of a 10b-5 case is that the defendant omitted a material fact. The disclosure of industry-wide news by a peer firm in the industry can provide useful evidence to support this claim. If one firm in the industry discloses, plaintiffs are more likely able to establish that the non-discloser knew the information (or was reckless in not knowing it) and *intentionally* withheld it, which is otherwise difficult to prove. It may also be easier to establish materiality, which is a subjective evaluation, given that the disclosing firm assessed the information as material. While plaintiffs can bring a 10b-5 case based on a return drop such that litigation risk affects even a disclosing firm, we assume there are incremental costs of litigation for a non-discloser when a peer firm reveals the adverse news due to a stronger case for the plaintiff.

Reputation costs are another example of an externality cost. Individuals identified in enforcement actions lose their jobs, face restrictions on future employment, and incur pecuniary costs including fines and valuation losses on shareholdings (Karpoff et al., 2008a). Individuals can also face criminal charges and serve jail time (Karpoff et al., 2008a; Schrand and Zechman, 2012). Karpoff et al. (2008b) document reputational penalties at the firm rather than individual level. They

[5] Dye (1990) refers to such costs as "externality" costs; the decision to disclose by one firm imposes a real externality on a non-disclosing firm.

characterize the direct legal costs as fairly minimal, but the loss in firm market value associated with the reputational effects of the misconduct as "huge."[6]

The assumption of externality costs allows us to interpret observed correlation in decisions to withhold adverse news within an industry as "tacit collusion", defined as a strategic choice in response to the (conjectured) behavior of other firms. In the absence of these costs, one might observe correlated withholding due simply to correlated fundamentals.[7]

*2.3 Predictions*

Two features distinguish our setting from the settings considered in existing models of voluntary disclosure and are important to evaluating whether withholding private signals of adverse industry-wide news is a sustainable equilibrium. First, the assumption that the private signals contain a common industry-wide component implies that it is possible that the private valuation signals received by *all* firms in the industry are less than the current stock price valuation. That is, the posterior distribution of the valuations of firms in the industry is lower as a result of the news (and this is common knowledge). A distribution shift is not possible in the threshold models in which the signals are uncorrelated. As such, it remains a possibility that firms will conjecture that no firms will have incentives to disclose; withholding is an optimal strategy for every firm in the industry. The greater the raw magnitude of the *industry-wide* component of the bad news, the greater the likelihood that every firm's valuation signal is higher than its pre-signal valuation and that this relation is common knowledge. Thus, our first prediction is:

<u>Prediction 1</u>: *The probability of correlated withholding of industry-wide adverse news is increasing in the magnitude of the industry-wide component of the news.*

The second distinguishing feature of our setting that is important to the evaluation of correlated withholding as a sustainable equilibrium is that the manager anticipates subsequent

[6] Per Karpoff et al. (2008b): "For every dollar of inflated value when a firm's books are cooked, firm value decreases by that dollar when the misrepresentation is revealed; in addition, firm value declines $0.36 more due to fines and class-action settlements and $2.71 due to lost reputation." The analysis focuses on the effects of earnings misstatements, so these dollar amounts cannot be translated as penalties for disclosure omissions, but the evidence nonetheless supports the assumption that capital markets impose reputational penalties on firms that do not communicate.

[7] Mimicking is another explanation for observing correlated behavior in corporate decisions (e.g., capital structure decisions in Leary and Roberts, 2010). In the context of choosing to make financial statements less transparent, we do not put much weight on the mimicking explanation. While firms may voluntarily mimic other firms' to improve transparency or may be compelled to follow other firms (Jung, 2011), mimicking opacity seems less plausible.

revised signals at the time she makes the initial disclosure decision. The benefits from withholding adverse news will continue for the duration of the period that the news is withheld. The longer the expected duration that the news is withheld, the greater is the expected value of having an artificially elevated stock price, the greater are the economic incentives to withhold, and the more likely it is that all firms will share these beliefs. Thus, our second prediction is:

<u>Prediction 2</u>: *The probability of correlated withholding of industry-wide adverse news is increasing in the expected duration that the industry-wide component of the news is withheld.*

The expected duration, in turn, depends on two factors. First, the duration depends on the uncertainty of the shock. The more uncertain the reduction in valuation, the greater the likelihood that subsequent revised valuation signals will reverse the bad news before it is revealed. The news may never be revealed if the industry-wide bad news is not realized (e.g., a decline in average customer credit quality reverses prior to significant credit losses). Second, the duration depends on whether and when the withheld news is likely to be made public by external parties (e.g., an analyst or the media) who discover it or by a public announcement of macro-economic data (e.g., GDP figures that are relevant to firm valuation). In summary, the more uncertain the shock and the less likely it is to be revealed by external sources, the greater the expected duration of the period that the firm can maintain an artificially elevated stock price, and the more likely firms in an industry are to exhibit correlated withholding, hoping that the news never materializes.

Our assumption that managers anticipate revised signals also affects our predictions about the impact of litigation risk on disclosure decisions relative to the existing literature. A common assertion in the voluntary disclosure literature is that firms disclose bad news (i.e., do not withhold adverse news) early to mitigate litigation risk (e.g., Skinner, 1994). The empirical evidence, however, is mixed. Although Skinner (1994) finds evidence consistent with this argument, Francis, Philbrick, and Schipper (1994) do not find evidence that preemptive disclosure mitigates litigation risk. Skinner (1997) then examines a broader sample and finds support consistent with Francis et al. (1994) but suggests that, while issuing forecasts may not deter litigation, it may reduce the costs. Although the empirical evidence is weak, this traditional view would suggest that the probability of correlated withholding of industry-wide adverse news is *decreasing* in the expected costs of litigation.

The anticipation of revised signals in future periods, however, affects this prediction. As the probability of a positive revision (i.e., less adverse news) increases, the probability that the shock will materialize or be discovered declines, and the expected litigation costs decline, *ceteris paribus*. Thus, expected litigation costs will be decreasing in the probability of a reversal. This reasoning suggests that the impact of unconditional litigation risk on withholding may be mitigated. Litigation risk may have little detectable impact on disclosure decisions related to adverse but uncertain news. However, Skinner (1994) suggests that litigation costs are increasing in the duration of withholding because a longer withholding period means a longer class period. This evidence suggests that expected litigation costs are increasing in the duration that the news is withheld, which is increasing in the uncertainty about future signals.

In summary, evidence on the impact of litigation risk on disclosure decisions is mixed and weak but would suggest a negative relation between withholding and litigation risk. The possibility of subsequent revised signals has an ambiguous impact on this directional prediction. Our third prediction is:

<u>Prediction 3</u>: *The probability of correlated withholding of industry-wide adverse news is decreasing in industry-level litigation risk.*

Finally, we make cross-sectional predictions based on the equity incentives of managers in the industry. Our setting assumes that managers benefit from maintaining an artificially elevated stock price that occurs if the news is withheld. This artificial elevation will occur only if all firms in the industry withhold the news. If any firm discloses, the news effectively becomes public. Adverse selection should result in full disclosure (or a lower partial disclosure threshold), and the benefits of the elevated stock price disappear. In industries in which more firms have equity incentives, it is more likely that firms will have individual incentives to withhold and to believe that other firms share these incentives, increasing the likelihood that withholding of industry-wide adverse news is a sustainable equilibrium. Thus, our fourth prediction is:

<u>Prediction 4</u>: *The probability of correlated withholding of industry-wide adverse news is increasing in industry-level benefits of maintaining an inflated stock price.*

### *2.4 Related literature on tacit collusion in product markets*

The analogy of strategic coordination of disclosure choice to "collusion" in product markets is useful,[8] but the translation is admittedly less than perfect. The main distinction between this setting and product market games relates to the payoffs. In product market games, the payoffs if both players cooperate (i.e., collusion), if both pursue the non-cooperative strategy (i.e., prisoners' dilemma), and if one player deviates from the cooperative strategy while the other does not are functionally related to the players' strategic decisions. In our setting, strategic interactions do not affect the level of the payoff to cooperation, nor do they affect the level of the externality costs, which a player that deviates avoids. The expected valuation implications of the industry shock could ultimately depend on how firms react to the shock through their operating, investing, or financing decisions, and each firm's reaction to the shock could be related to the anticipated reactions of other firms. The focus of our analysis is on the disclosure of the shock, not on tacit collusion in a firm's operating, financing, or investing decisions as a reaction to the shock.[9] This idea, however, is incorporated into our setting through the assumption that firms anticipate receiving new signals in subsequent periods and this is common knowledge. Conjectures about the new signals can embed the firm's expectations about other firms' strategic reactions to the shock. In the empirical analysis, we consider fundamental industry characteristics that we expect will affect players' conjectures about strategic reactions and the extent to which this is common knowledge. For example, we predict that firms in more homogeneous industries are more likely to have similar payoffs in the case of cooperation, and to conjecture that other firms have the information and believe each firm has similar beliefs (and so on). This prediction mirrors predictions in the product market literature, in which industry homogeneity is assumed to affect firms' conjectures about the payoffs to cooperation because of the form of the strategic reaction functions (Ivaldi et al., 2003).[10]

Our predictions about when a withholding equilibrium is sustainable are similar to predictions of "tacit collusion" in product markets on several dimensions. An individual firm's decision to withhold is based not only on its own signal but also on its beliefs about other firms'

[8] See relevant discussions in Levenstein and Suslow (2006), Ivaldi et al. (2003), Symeonidis (2003), and Pindyck and Rubinfeld (1989).

[9] Rajan (1994) is an example of a model of firms' choices about disclosures of industry-wide news in the context of a specific strategic decision. The disclosure choice is a bank's decision about recording credit reserves, which serves as a signal to other banks that make lending decisions about the overall level of credit quality in the industry. The analysis focuses on the implications of the recorded reserves for strategic lending and other decisions, which affect payoffs and business strategy, and the overall availability of credit.

[10] Another difference between our assumed payoff structure and those in models of price or quantity decisions is that tacit collusion in product market decisions typically describes the cooperative equilibrium that arises when firms can credibly retaliate against firms that deviate (Bertomeu and Liang, 2008). The threat of retaliation does not exist in the disclosure setting.

signals and higher order beliefs. Firms' conjectures can depend on soft factors like *i*) the availability of a focal point; *ii*) the frequency of interactions between the players; and *iii*) trust among the firms in the industry. The reliance on conjectures makes these equilibriums inherently unstable, as evidenced by the numerous empirical studies on when tacit collusion or cartel formation occurs in product markets and why cartels dissolve.[11]

## *3. Measuring annual report opacity and collusion episodes*

Section 3.1 describes the analysis to identify collusion episodes. Section 3.2 reports subsequent return performance for the identified episodes compared to non-collusion industry years. The subsequent performance analysis shows that the identified episodes can empirically distinguish industries that realize poor performance following the collusion year and hence provides some validation that the *ex ante* identified episodes are associated with yet unrealized bad news.

### *3.1 Detecting periods when industry-wide bad news is withheld*

We define an industry-year collusion episode as an unexplained increase in the within-industry average annual report opacity. The underlying assumption of this measure is that firms increase opacity when they are attempting to hide information. Our model of opacity controls for industry-wide shocks to (disclosed) fundamentals that could also generate an observed increase in intra-industry FOG scores. Thus, we interpret the measured increase in opacity as discretionary, and interpret the increase as intentional obfuscation.

We measure annual report opacity using the FOG index (Li, 2008). We determine the discretionary component using a model that controls for firm-level determinants of FOG as identified in Li (2008). The control variables include: the log of market value of equity, the market to book ratio, special items scaled by total assets, return volatility, the number of non-missing data items in Compustat as a measure of complexity, firm age, an indicator for Delaware incorporation, the log of the number of geographic segments plus one, and the log of the number of business segments plus one. (See variable definitions in Appendix A.) Including the control variables in the regression mitigates the concern that observed changes in the FOG score are due to changes in fundamentals.

We regress firms' 10-K FOG scores on the control variables by FF49 industry for rolling two year windows over the period 1994 through 2008. We require at least ten observations within

[11] See Levenstein and Suslow (2006) for a review.

the industry for each year included with all variables specified in the regression. The model includes a year indicator for the later of the two years (YEAR2). An industry-year is considered a "collusion episode" in year 2 if the coefficient estimate on the YEAR2 indicator, controlling for industry fundamentals, is positive and significant (p-value < 0.10). This analysis generates observations of collusion episodes for the years 1995 through 2008 ($COLLUDE_{AVG}$). The "AVG" subscript indicates that this variable measures whether there is a significant increase in the within-industry average FOG score.

We estimate the model separately for two samples of firms. The first sample includes all firms in each FF49 industry with available data for the year ("FULL" sample). The second sample excludes firms with more than one industry segment based on the Fama-French 49 industries ("1SEG" sample). We create the second sample because, ex ante, we expect the parameter estimates from the FOG model estimated with the 1SEG sample to provide better controls for the fundamental industry characteristics that affect FOG, thus providing a cleaner measure of the change in the discretionary component of FOG. However, this sample has fewer observations, reducing its power to detect significant collusion episodes. The relative power of these samples to detect collusion episodes is ambiguous.

Table 1 reports data on select parameter estimates and model statistics for estimates from a regression of FOG on the control variables, by FF49 industry and year. Industry-year averages are reported for the FULL sample (Panel A) and the 1SEG sample (Panel B). Of the 15 years over which we attempt to estimate the model between 1994 and 2008, we are able to estimate it for 12 years (6 years) per industry on average in the FULL (1SEG) sample. The grand average numbers of firms per industry-year are 63 and 17, respectively. These figures illustrate our concern about the smaller number of observations in the 1SEG sample. The grand average of the 49 industry intercepts is 19.5 in the FULL sample and 19.2 in the 1SEG sample. Although the raw magnitudes are similar, only 16.5% of the intercepts are significant in the FULL sample, versus 30.3% in the 1SEG sample. Thus, the industry dummy captures less of the variation in FOG after controlling for the fundamentals in the FULL sample than in the 1SEG sample, consistent with our expectations, which reduces its power to detect significant collusion episodes.

We create a second measure of collusion episodes that is a subset of the previous episodes. This measure recognizes that it is important that the key influential players in the industry show increases in FOG. We first identify the number of influential firms in each industry-year combination. We find the lowest number of firms (*n*) such that the concentration ratio for *n* is greater than or equal to 50%. For example, if the concentration ratio for the top four (six) firms is 40% (55%), then the number of influential firms in the industry is six. We estimate a prediction model for FOG based on prior year determinants of FOG for the respective industry and predict current-period "normal" FOG levels for the influential firms. The difference between actual FOG and predicted FOG provides an estimate of abnormal FOG levels. A positive abnormal FOG level suggests that the firm has a higher than predicted level of FOG based on fundamental determinants and is consistent with the firm having increased its disclosure opacity or obfuscation. Our second measure of collusion represents the subset of collusion episodes identified based on a significant within-industry average increase in FOG that also has at least 60% of influential firms with positive abnormal FOG levels ($COLLUDE_{INFL}$).

Table 2 provides a summary of the collusion episodes for both the $COLLUDE_{AVG}$ and $COLLUDE_{INFL}$ episodes across both the FULL and 1SEG samples. Based on the FULL (1SEG) sample, there are 17 (15) estimated collusion episodes as determined by significant increases in FOG in two-year rolling regressions ($COLLUDE_{AVG}$) and there are 12 (9) when we also require individual influential firms to have positive abnormal FOG ($COLLUDE_{INFL}$). Technology-related industries are over-represented. All results are robust to exclusion of industries 35 through 37 (Computers, computer software, and electronic equipment). The tech industry firms appear in 2002, after the burst of the tech bubble. They also show significant increases in average FOG in 2003, but 2003 is not identified as a collusion episode because we only identify the first instance of a significant increase in FOG. We also note time clustering in 2002 and 2003. We do not expect that required disclosures related to the Sarbanes-Oxley Act of 2002 (SOX) are the source of this clustering given that the specific provisions were not effective until much later and the general provisions should have affected all industries. However, as a precaution, our tests of the determinants of collusion episodes include observations only for years in which there is at least one episode.

*3.2 Returns following identified episodes*

We use a traditional portfolio methodology to investigate the relation between collusion episodes and subsequent returns. Our intent is to provide validation for the episodes we identified. If the collusion episodes indeed represent cases of hiding adverse industry-wide news, we expect to observe poor returns, on average, following collusion episodes compared to non-collusion industry years. However, this is an inherently weak test of tacit collusion as firms are more likely to withhold when they believe there is a positive probability that the bad news will not materialize (or will be better than originally anticipated). We do not make specific predictions about the timing of the subsequent poor performance, however, we do not expect the poor performance to be immediate. The duration of the withholding period is predicted to be positively related to the likelihood of withholding. If firms had rationally anticipated that the news would be made public quickly, withholding should not have been the chosen strategy.

We measure subsequent returns for firms in the colluding industry as the cumulative abnormal size-decile adjusted returns (CARs) in months 1-, 3-, 6-, 12-, 24- and 36 after month four (April) following the collusion episode year end. We also examine subsequent returns for firms that were not in a colluding industry. We exclude firms identified as colluders in year *y* from the set of non-colluders in years *y+1* through *y+3,* so that returns for the colluding firms do not confound the *y+1* and *y+2* returns for the non-colluding samples in the subsequent period. For example, if an industry is labeled a colluder in 2002, it is excluded from the non-colluder sample in 2003-2005. We examine return performance for all firms in the industry, including firms that were not included in the determination of the collusion episodes due to data availability and we do not require firms to have data for all periods subsequent to the collusion episode.

Table 3 Panels A and B report the average firm-level CARs for the collusion and non-collusion portfolios for the 15 collusion episodes that were identified using the rolling two-year model to identify average increases in FOG for the 1SEG sample. Panel A presents average CARs for portfolios that include only single segment firms and exclude new entrants to the industry, defined as firms with returns available as of month four following the collusion episode year (*y*) but with missing returns as of December year *y*. Assuming these firms were privy to the news when they chose to enter, they are likely different from the existing firms, reducing the likelihood that the industry shock will affect them similarly.[12] Panel B presents average CARs for single and multi-segment firms, which significantly increases the sample size in our tests, but at the cost of less specific returns related to the industry subject to the adverse news.

In the colluding portfolios in Panels A and B, there are 1,750 (2,246) observations for individual single segment (single and multi-segment) firms with data available to compute CARs in the first month following the collusion episodes aggregated across the 15 collusion episodes; this number declines to 1,295 (1,708) by month +36. In both panels, the CARs of firms in colluding industries turn negative (but not significant) by the end of the second year following the collusion

[12] We also create a set of firms that includes single segment new entrants to the industry. All results (untabulated) are qualitatively the same, suggesting that new entrants and existing firms were similarly affected by the industry-wide news.

episode. The CARs of the colluders are significantly greater than those of the non-colluders as of each of the subsequent measurement dates up to one year following the collusion episode (Ret12mon). By Ret24mon, however, the CARs of the colluders in the single segment firms are significantly lower than those of the non-colluders. Results for the 9 collusion episodes that required positive abnormal FOG by influential single segment firms ($COLLUDE_{INFL}^{1SEG}$) are similar (untabulated), including when we limit the portfolio firms to the influential single segment firms.

Table 3 Panels C and D report CARs for cross-sections based on the propensity of the industry to experience negative tail risk (positive skew). We identify industries with high tail risk based on expected return skew implied by option prices of firms in an industry. Appendix A describes the process used to identify industries with high tail risk (*TAILRISK* = 1). Appendix B marks the high tail risk industries with an asterisk. There are 17 high tail risk industries, or approximately 35% of the FF49 industries. We expect the differences in CARs between the colluders and non-colluders to be more pronounced in high tail risk industries because they are more likely to experience large negative shocks. This analysis is meant to reduce noise in the portfolios of non-colluders by eliminating from the high tail risk portfolio the industries that were unlikely to collude because they had no bad news to withhold.

Panel C reports CARs for single segment firms in high tail risk industries with collusion episodes that were identified using the rolling two-year model to identify average increases in FOG for the 1SEG sample, which is comparable to the portfolio analyzed in Panel A. Except in month +1, the differences between the colluders and non-colluders are insignificant, in contrast to the higher CARs for the colluders reported in Panel A. The CARs turn significantly negative for the colluders but not for the non-colluders in Ret24mon, and the difference in CARs between the two samples is significantly negative in Ret36mon. Panel D shows that the low tail risk industries drive the positive differences in CARs between the colluders and non-colluders in months 1, 3, 6, and 12 that were reported in Panel A. In fact, in the low tail risk industries, the colluders have significantly higher CARs than the non-colluders in Ret36mon.

In untabulated analysis, we adjust CARs for performance-related delistings. These missing observations could create a bias in our sample that would work against finding subsequent bad news in the collusion industries if the missing observations are the firms with the most egregious bad news. We replace observations missing due to performance-related delistings with the 5$^{th}$ percentile CAR for the same FF49 industry in year *y*. The replacement of missing values with the 5$^{th}$ percentile described above is made for both the colluder and non-colluder portfolios. The delisting codes (from CRSP) we use to identify performance-related delistings are 400 to 500 and 520 to 584 (Shumway, 1997). We assume delistings for other reasons (e.g., changes in exchanges) do not create a bias in our portfolio returns. In Panels A and B, the CARs are lower, by construction, for both the colluders and non-colluders, but the differences in CARs are similar in magnitude and significance.

Overall, there is weak evidence of poor subsequent returns for the colluder samples. Our analysis of average *realized* returns understates the amount of bad news that was hidden if the bad news is either partially mitigated or never realized. Some industries likely hid industry-wide adverse news about possible future outcomes hoping that the bad outcomes would never materialize, consistent with survey responses in Graham et al. (2005).

## *4. Analysis of determinants of collusion episodes*

### *4.1 Research design*

In this section, we report the analysis of the determinants of the collusion episodes identified in Section 3. We estimate the following logit model:[13]

$$\begin{aligned} COLLUDE_{iy} = \alpha_0 + \delta_1 HETERO_{iy} + \delta_2 TAILRISK_i + \delta_3 PUBLIC_i \\ + \delta_4 KS\_LIT_{iy} + \delta_5 \%CRSP_{iy} + \delta_6 HERF_i + \epsilon_{iy} \end{aligned} \quad (1)$$

The dependent variable equals one for the industry-year episodes identified and equals zero for the remaining non-missing industry-year observations. The regression is run separately for the 17 (15) episodes identified using the rolling two-year regressions and the two samples ($COLLUDE_{AVG}^{FULL}$ and $COLLUDE_{AVG}^{1SEG}$) and for the 12 (9) episodes identified restricting the sample to episodes with positive abnormal FOG for the influential firms ($COLLUDE_{INFL}^{FULL}$ and $COLLUDE_{INFL}^{1SEG}$). An industry-year is considered missing if data were not available to estimate the FOG model. As described previously, an industry-year is also considered missing for an industry *j* in years *y+1* through *y+3* if it was defined as a colluding industry in year *y*. This exclusion reduces noise in the "non-colluding" observations, since an industry that colludes in year *y* may continue to hide the news. The logit model includes only the years with at least one colluding episode.

We measure each of the explanatory variables, described below, at the industry-year or industry level. Whether withholding industry-wide adverse news is a sustainable equilibrium

[13] We also estimate the model using a probit specification with consistent results.

depends not only on each firm's own incentives and costs, but also on the firms' conjectures about other firms' disclosure decisions. If firms conjecture that it is in the best interest of all firms in the industry to withhold, withholding is sustainable. If not, at least one firm is expected to deviate and disclose the news. In the presence of externality costs, no firm will be willing to withhold the industry-wide news, fearing that other firms will disclose it first. Because the sustainability of a withholding equilibrium depends on conjectures about the costs and benefits for *all* firms in the industry, we measure the explanatory variables at the industry or industry-year level. Table 4 presents summary descriptive statistics for the explanatory variables, which we describe next. Appendix B provides summaries of these variables by industry.

Our first prediction is that correlated withholding of industry-wide news is increasing in the magnitude of the industry-wide component of the news. Our proxy for the industry-wide component of the shock is intra-industry heterogeneity (*HETERO)*, measured by industry-year as the annual average of the standard deviation of residuals from monthly within-industry estimations of a standard market model (see Appendix A). This proxy captures the average idiosyncratic component of returns relative to the industry-wide component. Industries with greater cross-sectional variation in idiosyncratic news are less likely to sustain a large common industry-wide shock and this will be common knowledge. While this reasoning suggests that heterogeneity makes correlated withholding decisions *less* likely, heterogeneity also decreases the probability that traders will infer industry-wide news arrival from the disclosure of any individual firm. Trader uncertainty about news arrival remains. More firms will withhold information, hiding in the pool of firms that traders expect did not receive a signal consistent with the predictions of the threshold models. In summary, our prediction about the association between heterogeneity in industry operations and the likelihood of correlated withholding is directionally ambiguous.[14]

Our second prediction is that correlated withholding of industry-wide news is increasing in the expected duration that the industry news is withheld, which depends on two factors: the uncertainty of the shock and the likelihood of revelation by outside information sources. Our proxy

[14] This prediction is related to a case considered in DS. In their main analysis, the timing of the arrival of private signals is correlated, but the signal values are uncorrelated, and firms do not learn whether other firms have received a signal. Our prediction about the impact of heterogeneity is related to their comparative static analysis related to the number of firms ($n$) in the industry (Theorem 2a). Two forces are at play. As $n$ increases, traders infer less from the disclosure by any one firm about the probability that a non-discloser has received a private signal, but adverse selection motivates more firms to disclose. The second effect always dominates and the threshold is decreasing in $n$.

for uncertainty is industry-tail risk (*TAILRISK),* described previously, which is an industry-level indicator variable for the propensity of an industry to experience negative tail risk (positive skew).

Our proxy for the duration of the period until the news is revealed by external sources is an industry-level measure of the availability of public information (*PUBLIC).* For each of the FF49 industries, we separately estimate a standard single factor market model and a factor model that includes seven macro-economic risk factors: short-term interest rates, default spreads, term spreads, a foreign exchange factor, a producer price index, and the Fama-French SMB and HML factors. The seven factors represent observable public signals. The incremental power of these observable factors to explain returns for firms in the industry provides a proxy for the degree to which news about the industry is likely to be public. The variable *PUBLIC* is the incremental adjusted-$R^2$ from adding the observable risk factors to the single factor market model (see Appendix A for details). Table 4 reports that the average (median) incremental adjusted-$R^2$ is 5.5% (3.6%). Appendix B (final column) reports the *PUBLIC* variable by FF49 industry. Eight industries stand out as having substantially higher incremental $R^2$s after adding the observable macro-economic factors: coal, precious metals, tobacco products, defense, shipbuilding/railroad equipment, petroleum and natural gas, utilities, and non-metallic/industrial metal mining. Eight industries with minimal incremental $R^2$s from adding the observable macro-economic factors are: wholesale, business services, medical equipment, consumer goods, personal services, entertainment, electrical equipment, and machinery. This proxy is intended to capture the availability of public information about firms in an industry.

Our third prediction is that correlated withholding of industry-wide news is decreasing in the expected costs of litigation. Our measure for litigation risk uses Model 3 from Table 7 of Kim and Skinner (2012), which is intended to measure the *ex ante* probability of litigation, but we use a different sample. We estimate "*KS_LIT*" at the firm level for fiscal years between 1996 and 2008 and create an industry-year average. (See Appendix A for a description of their model and our sample.)

Our fourth prediction is that correlated withholding of industry-wide news is increasing in the benefits of maintaining an inflated stock price. Our proxy for the industry-year equity incentives to withhold is the number of firms in the industry that have data on CRSP divided by the number of firms that have data on Compustat (*%CRSP*). For CRSP, we count the firms in each FF49 industry/year with greater than 200 non-missing daily return observations. For Compustat, we count the firms in each FF49 industry/year in the fundamentals annual file with an available SIC code and non-missing annual firm-level sales and total assets data. Firms in industries with a higher percent of firms on CRSP are more likely to conjecture that other firms in the industry will have similar equity incentives to withhold adverse news in their filings. Appendix B presents the average annual *%CRSP* measures by FF49 industry. Industries with low percentages are: Shipbuilding, Railroad Equipment (73%), Rubber and Plastic Products (76%), Utilities (76%), and Apparel (78%), ignoring the "Almost Nothing" industry. Industries with high percentages are: Banking (215%), Precious Metals (305%), and Trading (344%). The average (median) is 121.4% (103.8%) and the inter-quartile range is 88.9% to 125.8% (not reported).

In summary, we have negative predictions for the association between the likelihood of correlated withholding and the availability of public information (*PUBLIC*) and litigation risk (*KS_LIT*). We have positive predictions for the association between correlated withholding and industry uncertainty (*TAILRISK*) and the degree of equity incentives to withhold bad news (*%CRSP*). The direction of the predicted association with the level of industry heterogeneity (*HETERO*) is ambiguous.

The model also includes one control variable, the industry-year revenue-based herfindahl index as a measure of industry concentration (*HERF*).[15] Industry concentration is expected to be associated with a number of industry characteristics that could be associated with incentives or opportunities for coordination. Greater industry concentration indicates that there are a smaller number of larger and perhaps more individually influential firms in an industry, which may enable easier coordination, perhaps because of explicit collusion opportunities, or perhaps because it is more likely that a focal point is available.[16] Greater concentration may also be associated with firms' conjectures that other firms have received a similar signal and believe each firm has (and so on), and that the signal affects all other firms in the same direction, again increasing the likelihood of collusion. In addition, greater industry concentration can also be associated with high entry barriers, in which case firms have less incentive to disclose adverse news to deter new entrants. Thus, the payoffs to withholding are greater than the payoffs to disclosure, *ceteris paribus*, increasing the sustainability of a withholding equilibrium.

---

[15] We also compute an asset based herfindahl index and the concentration of the top four, six, or eight firms in each industry and year based on market shares of total sales and assets. See Appendix A for details. We report the herfindahl index as it gives some weight to "medium-sized" firms, however, results are consistent with concentration ratios.

[16] Ivaldi et al. (2003) note this explanation for a relation between industry concentration and collusion even in product markets but claim that there is little evidence. Instead, concentration ratios are commonly predicted to be associated with tacit collusion in product markets because concentration determines the profitability of a collusive price (or quantity) strategy relative to the non-collusive strategy, and the short-term profits from deviating, both of which affect firms' equilibrium beliefs about whether the collusive product-market strategy is sustainable.

Table 4 Panel B reports the correlation matrices for the explanatory variables separately for the observations that will be included in the FULL sample logit analysis and in the 1SEG sample logit analysis. The correlations with industry size are also reported although industry size is not included in the model. In both samples, *PUBLIC* is negatively correlated with intra-industry heterogeneity (*HETERO*). The negative correlation between *HETERO* and *PUBLIC* could indicate that *PUBLIC* is capturing intra-industry homogeneity. *PUBLIC* is an industry-level proxy estimated over the entire sample period, while *HETERO* is based on monthly observations and generates an industry-year measure. Nonetheless, because of the correlation, we estimate the logit model with these variables entered individually as a robustness test when interpreting the results.

*4.2 Results*

Table 5 reports the logit model results. Panel A presents results when we define collusion episodes based on significant increases in average industry FOG using rolling two-year regressions ($\text{COLLUDE}_{\text{AVG}}$). Panel B presents results when we use the subset of collusion episodes in which the influential firms have significant positive abnormal FOG ($\text{COLLUDE}_{\text{INFL}}$). Columns (2) and (4) exclude *HETERO* given its negative correlation with *PUBLIC*. In Panel A, there are 17 (14) industry-year collusion episodes and 120 (112) non-collusion industry-year observations for the FULL (1SEG) sample. One collusion episode (from 1995) is lost because the litigation risk proxy is available starting in 1996. In Panel B, there are 12 (9) industry-year collusion episodes and 125 (119) non-collusion industry-year observations for the FULL (1SEG) sample.

The first significant result is a negative association between the availability of observable information as measured by *PUBLIC* and collusion episodes. This result, which holds consistently across the samples and the methods of identifying collusion episodes, is consistent with the prediction that more public information shortens the expected duration of the surplus to withholding, thus decreasing the likelihood that firms believe a withholding equilibrium is sustainable. In short, more observable public information deters collusion to withhold adverse news.

The second significant result is that litigation risk is positively associated with collusion episodes in the FULL sample in both panels. The positive association contradicts the prediction that greater litigation costs provide incentives for firms to preemptively disclose bad news (Skinner, 1994). However, it is important to recall that the litigation cost that firms will incur is conditional on a suit being brought, which depends on the probability of subsequent realization of the shock. Our proxy for litigation costs (*KS_LIT*) is based on a model of suits that were brought against firms in an industry. As such, *KS_LIT* could reflect greater uncertainty about the news. Firms are more likely to believe that future valuation signals may increase (i.e., the bad news will improve or never materialize), and that this is common knowledge, which increases the sustainability of a withholding equilibrium.

The final significant result is that greater equity incentives as measured by the percentage of CRSP firms relative to Compustat firms in the industry (*%CRSP*) is consistently positively associated with the likelihood of a collusion episode in the FULL samples. This result is consistent with our prediction that correlated withholding of industry-wide news is increasing in the benefits of maintaining an inflated stock price. A second interpretation of the result, however, is that the positive association comes primarily from the low end of the *%CRSP* distribution. Low observations of *%CRSP* represent industries in which the number of Compustat firms relative to CRSP firms is high, which means that firms file annual reports despite not having publicly traded equity. Thus, the positive association between *%CRSP* and collusion episodes could imply that low-*%CRSP* industries are subject to significant mandated disclosure requirements, such that firms in these industries are less likely to believe that a withholding equilibrium is sustainable. Additional analysis in Section 4.3 attempts to disentangle these explanations.

The association between concentration (*HERF*) and collusion is positive, significant in the 1SEG sample in Panel A. There are several explanations for a positive association. Higher concentration could reflect a smaller number of influential firms such that firms are more likely to conjecture that other firms will have the same industry-wide news and that it is common knowledge. Higher concentration could also be correlated with softer factors that can sustain a cooperative withholding equilibrium, such as whether there is a focal point. Finally, higher concentration could imply greater barriers to entry, which decrease firms' incentives to disclose adverse news (i.e., increase firms' incentives to withhold bad news).

Two determinants of collusion episodes that we investigate, but that are not significant, are intra-industry heterogeneity in operations (*HETERO*) and high tail risk (*TAILRISK)*. The predicted association between *HETERO* and collusion was ambiguous. On one hand, industry news is less likely to affect all firms in heterogeneous industries in the same direction and have the effect be common knowledge, which makes collusion less sustainable. On the other hand, intra-industry heterogeneity makes it less likely that investors will infer industry-wide news from the disclosure by one firm, decreasing the costs of withholding and making a collusive equilibrium more sustainable.

The panel data used to estimate the model includes only years in which there is at least one colluding industry. If we include all non-colluding year observations in the analysis, the results for the FULL samples are similar to those presented in Table 5 (untabulated). For the 1SEG sample in Panel A, the marginal probabilities of *%CRSP* and *KS_LIT* are statistically significant, as for the FULL sample. For the 1SEG sample in Panel B, the marginal probability of *KS_LIT* is statistically significant as for the FULL sample. The significance of the coefficient estimate on *PUBLIC* diminishes to 15%.

*4.3 Additional analyses*

Our first additional analysis is meant to determine whether correlated fundamentals explain our findings. A possible explanation for our findings is that the explanatory variables in the logit regression are correlated with changes in industry fundamentals that are omitted from the FOG model. However, if correlated fundamentals explain our findings, we expect the explanatory variables in the logit model to explain industry-wide significant *decreases* in FOG (i.e., increases in transparency) as well as the industry-wide *increases* in FOG that we use as an indication of collusion. We conduct a falsification test to determine if the explanatory variables explain significant industry-wide decreases in FOG, which we call transparency episodes, expecting that they should not. The FULL (1SEG) sample have 12 (10) significant industry-year decreases or transparency episodes and 213 (123) non-transparency industry-year observations. Table 6 reports the results of estimating the regression from Table 5 on the transparency episodes. The results for the collusion episodes are reported in columns (1) and (3) for convenient comparison.

For both panels in Table 6, industry heterogeneity (*HETERO*) is positively associated with the transparency episodes, in contrast to the lack of association with the collusion episodes. The association between transparency episodes and the availability of public information (*PUBLIC*) switches sign relative to the collusion episodes but remains insignificant. In Panel A for the FULL sample, the associations between litigation (*KS_LIT*) and equity incentives (*%CRSP*) that were observed for collusion episodes are not significant determinants of the transparency episodes. Finally, in Panel B for the 1SEG sample, industry concentration (*HERF*) loses significance. In summary, the variables that explain collusion episodes do not explain transparency episodes.

Our second additional analysis is meant to distinguish two explanations for the positive association between *%CRSP* and collusion episodes reported in Table 5. Our cross-sectional prediction is that firms in industries with high *%CRSP* are more likely to collude to hide adverse news because of greater equity incentives. A second explanation for a positive association between collusion episodes and *%CRSP*, however, is that firms in industries with low *%CRSP* are *less* likely to collude because of existing mandatory disclosure regulation. This explanation is similar to our prediction about the impact of the availability of observable public information on collusion. If it is likely that the information will be made public shortly, in this case due to mandatory disclosure requirements, then firms have incentives to disclose to avoid the externality costs.

In an attempt to distinguish these two explanations, we analyze whether the positive relation between *%CRSP* and collusion episodes comes from the low end or the high end of the distribution of *%CRSP* (or both). Table 7 reports the results. We use two model specifications given empirical problems associated with using interaction terms in logit models combined with our small sample sizes. In the first test specification, we expand the model in Table 5 to include an interaction term that distinguishes the low end of the *%CRSP* distribution. We create an indicator variable (*LOW*) equal to 1 for industry-year observations that are less than or equal to the median industry-year level of *%CRSP*. Industry years with *LOW* = 1 (*LOW* = 0) are considered to have low (high) equity incentives.[17] The interaction term equals *%CRSP*LOW*.

Table 7 Panel A reports the results for collusion episodes defined based on a significant increase in average industry collusion using two-year rolling regressions ($COLLUDE_{AVG}^{FULL}$) and Panel B reports results for the subset of collusion episodes for which influential firms have abnormal FOG ($COLLUDE_{INFL}^{FULL}$). In both panels, results are presented only for the FULL sample due to small sample sizes in the 1SEG sample. It is worth noting that the marginal probabilities for the other explanatory variables in the analysis in both panels are similar to those in Table 5 in terms of sign, magnitude, and significance (columns (1) and (4)). The marginal effect of *%CRSP* declines but remains positive and significant. The marginal effect of the interaction term (*%CRSP*LOW*) is not significant.

Because the interpretation of interaction terms in logit models can be unreliable (Ai and Norton, 2003), we also estimate the model from Table 5 for separate samples of industries with low equity incentives (*LOW* = 1) and high equity incentives (*LOW* = 0) based on the industry-year median. For industry years with high equity incentives in the $COLLUDE_{AVG}^{FULL}$ sample (Panel A, column (3)), the marginal effect of *%CRSP* increases in magnitude and significance (p-value = 0.06). This result suggests that the explanatory power for the collusion episodes derives from the high end

[17] The results are the same if we define low (high) equity incentives as *%CRSP*<1 (*%CRSP*≥1).

of the *%CRSP* distribution, which represents managers who are more likely to benefit from keeping the adverse news hidden from equity markets, and who believe other firms in the industry share those benefits and beliefs.

### *5. Conclusion*

This study empirically examines whether managers withhold adverse news when it is industry-wide. Whether withholding industry-wide adverse news is a sustainable equilibrium depends on firms' conjectures about other firms' disclosure decisions. If firms conjecture that it is in the best interest of all firms in the industry to withhold, withholding is sustainable. If not, at least one firm is expected to deviate and disclose the news, imposing costs on the non-disclosers. In this scenario, no firm will be willing to withhold the industry-wide news, fearing that other firms will disclose it first. Ultimately, because the sustainability of a withholding equilibrium depends on the relative magnitudes of the costs and benefits for each firm and conjectures about these costs and benefits for all firms in the industry, whether we expect to see correlated withholding of industry-wide bad news in practice is an empirical question.

We first document a relatively small number of "collusion episodes." Of 686 possible industry-year combinations, we find between 9 and 17 collusion episodes (depending on the sample and method of identification) in which an increase in intra-industry disclosure opacity suggests tacit collusion. Two years after the identified collusion episodes, we see evidence of significantly lower adjusted returns for the colluders than for the non-colluders. These results are consistent with the identified increases in opacity representing episodes in which the firms were indeed hiding bad news.

The collusion episodes are less likely in industries in which the news is likely to be public or become public quickly. Collusion is more likely in industries with a greater proportion of firms having equity incentives, in which the benefits of maintaining an elevated stock price are more likely to exceed the costs of withholding, and this would be common knowledge. Finally, withholding is more likely in industries with greater litigation risk. We do not find evidence suggesting that within industry heterogeneity in operations is associated with collusion episodes, potentially because of offsetting predicted associations.

## References


Acharya, Viral V., Peter DeMarzo and Ilan Kremer, 2011, Endogenous information flows and the clustering of announcements, *American Economic Review* 101: 2955-2979.

Ai, Chunrong and Edward C. Norton, 2003, Interaction terms in logit and probit models, *Economics Letters* 80: 123-129.

Bertomeu, Jeremy and Pierre Jinghong Liang, 2008, Dynamic voluntary disclosure with endogenous proprietary costs, Working paper.

Chen, Shuping, Dawn Matsumoto, and Shiva Rajgopal, 2011, Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance, *Journal of Accounting and Economics* 51 (102): 134-150.

Dunn, Kimberly and Siva Nathan, 2005, Analyst industry diversification and earnings forecast accuracy, *Journal of Investing* 14: 7-14.

Dye, Ronald A., 1985, Disclosure of nonproprietary information, *Journal of Accounting Research* 23: 123 -145.

Dye, Ronald A. and Sri S. Sridhar, 1995, Industry-wide disclosure dynamics, *Journal of Accounting Research* 33, (1): 157-174.

Easterbrook Frank H. and Daniel R. Fischel, 1984, Mandatory disclosure and the protection of investors, *Virginia Law Review* 70: 669-715.

Francis, Jennifer, Donna Philbrick, and Katherine Schipper, 1994, Shareholder litigation and corporate disclosures, *Journal of Accounting Research* 32: 137-164.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal, 2005, The economic implications of corporate financial reporting, *Journal of Accounting and Economics* 40 (1-3): 3-73.

Grossman, S.J., 1981, The informational role of warranties and private disclosure about product quality, *Journal of Law and Economics* 24: 461-484.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole, 2003, The economics of tacit collusion, Final report for DG competition, European commission.

Jorgensen, Bjorn N. and Michael T. Kirschenheiter, 2012, Interactive discretionary disclosure, *Contemporary Accounting Research* 29, 382-397.

Jung, Michael J., 2011, Investor overlap and diffusion of disclosure practices, Working paper, NYU Stern School of Business.

Jung and Kwon, 1988, Disclosure when the market is unsure of information endowment of managers, *Journal of Accounting Research* 26: 146-153.

Karpoff, Jonathan M., D. Scott Lee, Gerald S. Martin, 2008a, The consequences to managers for financial misrepresentation, *Journal of Financial Economics* 88: 193–215.

Karpoff, Jonathan M., D. Scott Lee, Gerald S. Martin, 2008b, The cost to firms of cooking the books, *Journal of Financial and Quantitative Analysis* 43: 581-611.

Kim, Irene and Douglas J. Skinner, 2012, Measuring securities litigation risk, *Journal of Accounting and Economics* 53 (1-2): 290-310.

Kothari, S.P., Susan Shu, and Peter D. Wysocki, 2009, Do managers withhold bad news?, *Journal of Accounting Research* 47: 1-36.

Li, Feng, 2008, Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics* 45: 221-247.

Leary, Mark T. and Michael R. Roberts, 2010, Do peer firms affect corporate financial policy?, Working paper, University of Pennsylvania Wharton School.

Levenstein, Margaret C. and Valerie Y. Suslow, *2006,* What determines cartel success?, *Journal of Economic Literature* 44 (1): 43-95.

Milgrom, Paul, 1981, Good news and bad news: representation theorems and applications. *Bell Journal of Economics* 17: 18-32.

O'Brien, Patricia C.,1990, Forecast accuracy of individual analysts in nine industries, *Journal of Accounting Research* 28 (2): 286-304.

Pindyck, Robert S. and Daniel L. Rubinfeld, 1989, *Microeconomics*, MacMillan Publishing Company, New York.

Rajan, Raghuram G., 1994, Why bank credit policies fluctuate: A theory and some evidence, *The Quarterly Journal of Economics* 109 (2): 399-441.

Rogers, Jonathan L. and Andrew Van Buskirk, 2009, Shareholder litigation and changes in disclosure behavior, *Journal of Accounting and Economics* 47 (1/2): 136-156.

Schrand, Catherine M. and Sarah L.C. Zechman, 2012, Executive overconfidence and the slippery slope to financial misreporting, *Journal of Accounting and Economics* 53: 311-329.

Shin, Hyun S., 2003, Disclosures and asset returns, *Econometrica* 71 (2003):105 - 133.

Shin, Hyun S., 2006, Disclosure risk and price drift, *Journal of Accounting Research* 44: 351 - 379.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52: 327-340.

Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news?, *Journal of Accounting Research* 32, 38-60.

Skinner, Douglas J., 1997, Earnings disclosure and stockholder lawsuits, *Journal of Accounting and Economics,* 23: 249-282.

Sletten, Ewa, 2012, The effect of stock price on discretionary disclosure, *Review of Accounting Studies* 17: 96-133.

Symeonidis, George, 2003, In which industries is collusion more likely?, Evidence from the UK, *Journal of Industrial Economics* 51 (1): 45-74.

Tse, Senyo and Jennifer Wu Tucker, 2010, Within-industry timing of earnings warnings: do managers herd?, *Review of Accounting Studies* 15: 879-914.

**Table 1: Summary of FOG prediction model estimates**

Averages of select parameter estimates and model statistics for industry-year estimates of the FOG model from 1994-2008 for the FULL sample (Panel A) and the sample of single segment firms (Panel B). The control variables included in the model are *MVE, MTB, Special items, Return volatility, Non-missing items, Firm age, Delaware, GEO Segments, and BUS Segments*. See Appendix A for variable definitions. Significance is determined by a 2-tailed p-value<0.10.

| | | Panel A: FULL sample | | | | Panel B: 1SEG sample | | | |
|---|---|---|---|---|---|---|---|---|---|
| **FF49 industry** | | Annual # regs | Avg annual "N" | Avg IND intercept | % sig intercepts | Annual # regs | Avg annual "N" | Avg IND intercept | % sig intercepts |
| 1 | Agric | 8 | 10 | 18.80 | 75.00% | - | 3 | . | . |
| 2 | Food | 15 | 38 | 14.85 | 26.67% | 5 | 12 | 16.63 | 20.00% |
| 3 | Soda | - | 4 | . | . | - | 2 | . | . |
| 4 | Beer | 1 | 8 | 36.25 | 100.00% | - | 3 | . | . |
| 5 | Smoke | - | 4 | . | . | - | 1 | . | . |
| 6 | Toys | 14 | 28 | 17.07 | 35.71% | 4 | 9 | 27.15 | 50.00% |
| 7 | Fun | 15 | 47 | 20.00 | 6.67% | 7 | 18 | 16.54 | 42.86% |
| 8 | Books | 14 | 25 | 24.49 | 7.14% | 3 | 6 | 24.99 | 100.00% |
| 9 | Hshld | 15 | 59 | 19.08 | 13.33% | 8 | 18 | 14.05 | 25.00% |
| 10 | Clths | 14 | 40 | 20.04 | 14.29% | 4 | 13 | 20.04 | 25.00% |
| 11 | Hlth | 15 | 63 | 19.96 | 13.33% | 6 | 19 | 23.17 | 33.33% |
| 12 | Medeq | 15 | 105 | 23.21 | 0.00% | 12 | 29 | 22.76 | 41.67% |
| 13 | Drugs | 15 | 202 | 21.22 | 0.00% | 15 | 49 | 21.14 | 13.33% |
| 14 | Chems | 15 | 59 | 20.12 | 0.00% | 5 | 13 | 21.70 | 0.00% |
| 15 | Rubbr | 15 | 30 | 20.77 | 26.67% | 4 | 8 | 16.16 | 50.00% |
| 16 | Txtls | 5 | 10 | 18.28 | 80.00% | 3 | 5 | 19.92 | 100.00% |
| 17 | Bldmt | 15 | 60 | 20.04 | 13.33% | 4 | 10 | 22.40 | 25.00% |
| 18 | Constr | 15 | 42 | 17.68 | 6.67% | 7 | 11 | 19.74 | 14.29% |
| 19 | Steel | 15 | 46 | 19.67 | 0.00% | 5 | 12 | 17.05 | 20.00% |
| 20 | Fabpr | 8 | 13 | 14.01 | 25.00% | 2 | 7 | 21.20 | 50.00% |
| 21 | Mach | 15 | 109 | 17.63 | 6.67% | 10 | 27 | 6.34 | 20.00% |
| 22 | Elceq | 15 | 44 | 17.51 | 6.67% | 4 | 7 | 16.55 | 50.00% |
| 23 | Autos | 15 | 45 | 18.32 | 0.00% | 4 | 11 | 23.82 | 25.00% |
| 24 | Aero | 13 | 13 | 16.63 | 69.23% | - | 4 | . | . |
| 25 | Ships | - | 6 | . | . | - | 2 | . | . |
| 26 | Guns | - | 6 | . | . | - | 1 | . | . |
| 27 | Gold | 6 | 10 | 19.43 | 16.67% | 3 | 6 | 15.97 | 33.33% |
| 28 | Mines | 6 | 10 | 20.21 | 50.00% | - | 3 | . | . |
| 29 | Coal | - | 4 | . | . | - | 2 | . | . |
| 30 | Oil | 15 | 114 | 18.70 | 0.00% | 15 | 36 | 17.70 | 26.67% |
| 31 | Util | 15 | 91 | 18.74 | 0.00% | 6 | 16 | 20.73 | 33.33% |
| 32 | Telcm | 15 | 90 | 20.36 | 0.00% | 13 | 23 | 21.31 | 46.15% |
| 33 | Persv | 14 | 36 | 17.98 | 14.29% | 3 | 8 | 22.20 | 0.00% |
| 34 | Bussv | 15 | 172 | 21.68 | 0.00% | 15 | 44 | 22.84 | 13.33% |
| 35 | Hardw | 15 | 86 | 20.67 | 0.00% | 12 | 28 | 2.11 | 41.67% |
| 36 | Softw | 15 | 270 | 19.02 | 0.00% | 15 | 64 | 20.54 | 0.00% |
| 37 | Chips | 15 | 180 | 21.13 | 0.00% | 12 | 42 | 12.72 | 41.67% |
| 38 | Labeq | 15 | 71 | 16.40 | 13.33% | 4 | 16 | 18.83 | 0.00% |
| 39 | Paper | 15 | 37 | 21.82 | 13.33% | 4 | 8 | 32.28 | 25.00% |
| 40 | Boxes | 8 | 10 | 0.93 | 50.00% | - | 3 | . | . |
| 41 | Trans | 15 | 81 | 19.46 | 0.00% | 8 | 20 | 19.92 | 25.00% |
| 42 | Whlsl | 15 | 121 | 19.08 | 0.00% | 8 | 31 | 25.11 | 12.50% |
| 43 | Rtail | 15 | 115 | 19.41 | 0.00% | 15 | 33 | 22.21 | 40.00% |
| 44 | Meals | 14 | 36 | 21.95 | 0.00% | 5 | 10 | 14.30 | 40.00% |
| 45 | Banks | 15 | 39 | 21.85 | 6.67% | 4 | 11 | 9.51 | 50.00% |
| 46 | Insur | 15 | 139 | 20.93 | 0.00% | 15 | 35 | 26.04 | 6.67% |
| 47 | Rlest | 14 | 24 | 22.16 | 7.14% | 3 | 7 | 13.66 | 33.33% |
| 48 | Fin | 15 | 200 | 21.06 | 0.00% | 14 | 53 | 21.70 | 7.14% |
| 49 | Other | 14 | 27 | 18.50 | 28.57% | 3 | 12 | 16.38 | 0.00% |
| Grand avg | | 12 | 63 | 19.48 | 16.51% | 6 | 17 | 19.16 | 30.29% |

**Table 2: Summary of collusion episodes**

This table details the industry-years labeled as colluders for the full sample and the single segment firm sample. Panel A provides details for the full sample of firms with available data and Panel B provides details for the sample of single segment firms. The collusion episodes determined by a significant industry-level increase in FOG in two-year rolling regressions ($COLLUDE_{AVG}$) are reported in columns (1), (2), (5) and (6). The collusion episodes determined by a significant industry-level increase in FOG combined with positive abnormal FOG for influential firms ($COLLUDE_{INFL}$) are reported in columns (3), (4), (7), and (8).

| | **Panel A: FULL Sample** | | | | | **Panel B: 1SEG Sample** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **$COLLUDE_{AVG}$** | | **$COLLUDE_{INFL}$** | | | **$COLLUDE_{AVG}$** | | **$COLLUDE_{INFL}$** | |
| Year | # yrs | Collusion episodes (1) | Fama-French 49 industry (2) | Collusion episodes (3) | Fama-French 49 industry (4) | # yrs | Collusion episodes (5) | Fama-French 49 industry (6) | Collusion episodes (7) | Fama-French 49 industry (8) |
| 1995 | 29 | 0 | | 0 | | 23 | 1 | 42 (whlsl) | 0 | |
| 1996 | 41 | 0 | | 0 | | 36 | 0 | | 0 | |
| 1997 | 42 | 0 | | 0 | | 38 | 0 | | 0 | |
| 1998 | 43 | 1 | 22 (elceq) | 1 | 22 (elceq) | 37 | 3 | 7 (fun)<br>15 (rubbr)<br>22 (elceq) | 3 | 7 (fun)<br>15 (rubbr)<br>22 (elceq) |
| 1999 | 42 | 0 | | 0 | | 33 | 0 | | 0 | |
| 2000 | 42 | 0 | | 0 | | 32 | 0 | | 0 | |
| 2001 | 39 | 0 | | 0 | | 32 | 0 | | 0 | |
| 2002 | 40 | 7 | 24 (aero)<br>32 (telcm)<br>34 (bussv)<br>35 (hardw)<br>36 (softw)<br>37 (chips)<br>46 (insur) | 5 | 32 (telcm)<br>34 (bussv)<br>35 (hardw)<br>37 (chips)<br>46 (insur) | 36 | 6 | 32 (telcm)<br>34 (bussv)<br>35 (hardw)<br>36 (softw)<br>37 (chips)<br>49 (other) | 4 | 34 (bussv)<br>36 (softw)<br>37 (chips)<br>49 (other) |
| 2003 | 31 | 8 | 8 (books)<br>11 (hlth)<br>13 (drugs)<br>14 (chems)<br>20 (fabpr)<br>38 (labeq)<br>42 (whlsl)<br>48 (fin) | 5 | 11 (hlth)<br>20 (fabpr)<br>38 (labeq)<br>42 (whlsl)<br>48 (fin) | 26 | 4 | 13 (drugs)<br>14 (chems)<br>21 (mach)<br>48 (fin) | 1 | 21 (mach) |
| 2004 | 23 | 1 | 9 (hshld) | 1 | 9 (hshld) | 22 | 0 | | 0 | |
| 2005 | 25 | 0 | | 0 | | 26 | 0 | | 0 | |
| 2006 | 28 | 0 | | 0 | | 29 | 0 | | 0 | |
| 2007 | 29 | 0 | | 0 | | 27 | 1 | 17 (bldmt) | 1 | 17 (bldmt) |
| 2008 | 23 | 0 | | 0 | | 22 | 0 | | 0 | |
| Total | 477 | 17 | | 12 | | 419 | 15 | | 9 | |

**Table 3: Returns following the collusion episodes**

This table presents average cumulative abnormal size-decile adjusted returns (CARs) in portfolios of firms in industries defined as colluders and non-colluders. Collusion episodes are determined by a significant industry-level increase in FOG in two-year rolling regressions for the 1SEG sample ($\mathrm{COLLUDE}_{\mathrm{AVG}}^{\mathrm{1SEG}}$). Returns are presented for months 1-, 3-, 6-, 12-, 24- and 36 after month four (April) of the collusion episode year *y*. Panels A and B present average CARs for single segment firms and single and multi-segment firms, respectively, with data as of December of the collusion year y. Panels C and D present average CARs for single segment firms in high *TAILRISK* and low *TAILRISK* industries, respectively. *, **, and *** equal 10%, 5%, and 1% significance at the 2-tailed level.

| | | Ret1mon | Ret3mon | Ret6mon | Ret12mon | Ret24mon | Ret36mon |
|---|---|---|---|---|---|---|---|
| **Panel A: Single segment firms** | | | | | | | |
| Colluders: | Return | 0.0367*** | 0.0583*** | 0.0961*** | 0.0691*** | -0.0205 | 0.0318 |
| | N | 1,750 | 1,716 | 1,667 | 1,596 | 1,452 | 1,295 |
| Non-colluders: | Return | 0.0043*** | -0.0015 | -0.0042* | 0.0036 | 0.0214*** | 0.0259*** |
| | N | 28,523 | 28,020 | 27,285 | 25,929 | 23,437 | 20,578 |
| **Difference in returns** | | **0.0324*** ** | **0.0598*** ** | **0.1003*** ** | **0.0655*** ** | **-0.0419** ** | **0.0059** |
| t-value | | 6.72 | 7.62 | 8.77 | 4.02 | 2.13 | 0.23 |
| **Panel B: Single+multi-segment firms** | | | | | | | |
| Colluders: | Return | 0.0319*** | 0.0529*** | 0.0825*** | 0.0660*** | -0.0148 | 0.0301 |
| | N | 2,246 | 2,204 | 2,142 | 2,060 | 1,893 | 1,708 |
| Non-colluders: | Return | 0.0033*** | -0.0002 | 0.0004 | 0.0012 | 0.0101*** | 0.0116** |
| | N | 44,141 | 43,403 | 42,350 | 40,365 | 36,720 | 32,269 |
| **Difference in returns** | | **0.0286*** ** | **0.0531*** ** | **0.0821*** ** | **0.0648*** ** | **-0.0249** | **0.0185** |
| t-value | | 6.98 | 8.07 | 8.55 | 4.73 | 1.53 | 0.87 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Panel C: High *TAILRISK* Single segment firms** | | | | | | | |
| Colluders: | Return | 0.0625** | 0.0627 | 0.0439 | -0.0260 | -0.2979* | -0.3654*** |
| | N | 27 | 26 | 25 | 23 | 19 | 18 |
| Non-colluders: | Return | 0.0007 | -0.0041 | -0.0289*** | -0.0188 | -0.0042 | -0.0329 |
| | N | 1,977 | 1,936 | 1,873 | 1,760 | 1,549 | 1,285 |
| **Difference in returns** | | **0.0618*** | **0.0668** | **0.0728** | **-0.0072** | **-0.2937** | **-0.3325*** ** |
| t-value | | 1.92 | 1.23 | 0.93 | 0.04 | 1.58 | 3.11 |
| **Panel D: Low *TAILRISK* Single segment firms** | | | | | | | |
| Colluders: | Return | 0.0589*** | 0.0989*** | 0.1640*** | 0.1549*** | 0.0194* | 0.1256*** |
| | N | 982 | 952 | 919 | 870 | 790 | 695 |
| Non-colluders: | Return | 0.0044*** | -0.0013 | -0.0021 | 0.0049 | 0.0226*** | 0.0282*** |
| | N | 28,128 | 27,649 | 26,938 | 25,642 | 23,239 | 20,208 |
| **Difference in returns** | | **0.0545*** ** | **0.1002*** ** | **0.1661*** ** | **0.1500*** ** | **-0.0032** | **0.0974*** ** |
| t-value | | 7.65 | 8.83 | 9.87 | 6.18 | 0.11 | 2.60 |

**Table 4: Descriptive statistics for industry-level characteristics**

This table provides descriptive statistics on industry-level characteristics that are the explanatory variables in the logit model. Industry size is the average number of firms in the industry with non-missing sales and assets data on Compustat. *HETERO* is the average monthly standard deviation of the idiosyncratic component of returns within an industry in year *y*. *TAILRISK* is an indicator variable that equals one for industries that have high negative expected return skewness computed based on the methods in Van Buskirk (2011), and zero otherwise (see Appendix A). *PUBLIC* proxies for the availability of public information and is the difference between the adjusted $R^2$ values from estimation of a standard market model and a factor model for each firm *i* within an industry, estimated using monthly return and factor data over the period 1994-2008. The factor model includes seven macro-economic factors described in Appendix A. *KS_LIT* is an industry-year measure of litigation risk (see Appendix A). *%CRSP* proxies for industry-level equity incentives and is the average annual number of firms with CRSP data relative to the number of firms with Compustat data (see Appendix B). *HERF* is the revenue-based herfindahl index for the 50 largest firms in the industry. Panel B presents the correlations between these variables by sample. Statistical significance in the collusion matrices at 10%, 5%, and 1% are denoted by *, **, and ***, respectively.

***Panel A: Average annual measures (477 industry-year observations in the FULL sample)***

| | Mean | Median | Std. Dev. | Minimum | Maximum |
|---|---|---|---|---|---|
| Industry size | 150 | 105 | 124.79 | 12 | 803 |
| H1 Heterogeneity (*HETERO)* | 0.1610 | 0.1552 | 0.0506 | 0.0520 | 0.3373 |
| H2 Uncertainty (*TAILRISK)* | 0.3774 | 0.0000 | 0.4852 | 0 | 1 |
| H2 Public information (*PUBLIC*) | 0.0401 | 0.0317 | 0.0292 | 0.0068 | 0.1861 |
| H3 Litigation risk (*KS_LIT*) | 0.0460 | 0.0428 | 0.0176 | 0.0162 | 0.1562 |
| H4 Equity incentives (*%CRSP*) | 1.0940 | 0.9861 | 0.5444 | 0.2273 | 4.8992 |
| Control: Industry concentration *(HERF)* | 0.1012 | 0.0732 | 0.1757 | 0.0225 | 0.4895 |

***Panel B: Correlation matrix***

***(FULL sample)***

| | *Industry size* | *HETERO* | *TAILRISK* | *PUBLIC* | *KS_LIT* | *%CRSP* |
|---|---|---|---|---|---|---|
| *HETERO* | 0.2229*** | | | | | |
| *TAILRISK* | 0.0066 | 0.0141 | | | | |
| *PUBLIC* | -0.0553 | -0.2644*** | -0.2740*** | | | |
| *KS_LIT* | 0.3437*** | 0.2340*** | 0.1366*** | -0.0803 | | |
| *%CRSP* | 0.0931** | -0.2150*** | 0.0608 | 0.3045*** | -0.0772 | |
| *HERF* | -0.2534*** | 0.0305 | 0.0353 | 0.1774*** | -0.0418 | 0.0737 |

***(1SEG sample)***

| | *Industry size* | *HETERO* | *TAILRISK* | *PUBLIC* | *KS_LIT* | *%CRSP* |
|---|---|---|---|---|---|---|
| *HETERO* | 0.2141*** | | | | | |
| *TAILRISK* | -0.0765 | -0.0385 | | | | |
| *PUBLIC* | 0.0667 | -0.2624*** | -0.2084*** | | | |
| *KS_LIT* | 0.3179*** | 0.2528*** | 0.0778 | -0.0416 | | |
| *%CRSP* | 0.0931* | -0.2463*** | 0.0525 | 0.2883*** | -0.0870* | |
| *HERF* | -0.1312*** | 0.1241** | 0.1516*** | 0.0486 | 0.0483 | 0.1069** |

**Table 5: Characteristics of industries that have collusion episodes**

This table presents estimated marginal effects from a logit regression model of collusion episode occurrence on industry and industry-year level correlates. The unit of observation is at the industry-year level. Panel A presents results for collusion episodes determined by a significant industry-level increase in FOG in two-year rolling regressions ($COLLUDE_{AVG}$). Panel B presents results for the subset of collusion episodes for which influential firms have a significant positive abnormal FOG ($COLLUDE_{INFL}$). Results are presented for the FULL sample (Columns 1 and 2) and 1SEG samples (Columns 3 and 4). See Table 4 for definitions of the independent variables. P-values of the $\chi^2$ test for coefficient significance are presented in parentheses. Statistical significance (two-sided) at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.

**Panel A: Collusion episodes identified by significant increase in industry-average FOG ($COLLUDE_{AVG}$)**

- **FULL sample: 17 collusion episodes; 120 non-collusion industry-year observations**
- **1SEG sample: 14 collusion episodes; 112 non-collusion industry-year observations**

| | Predicted Sign | FULL sample | | 1SEG sample | |
|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) |
| Intercept | | -0.4232*** | -0.3575*** | -0.3794*** | -0.2804*** |
| | | (0.00) | (0.00) | (0.00) | (0.00) |
| H1 Heterogeneity (*HETERO)* | ? | 0.4198 | | 0.7079 | |
| | | (0.51) | | (0.13) | |
| H2 Uncertainty (*TAILRISK)* | + | -0.0561 | -0.0597 | -0.0628 | -0.0699 |
| | | (0.26) | (0.22) | (0.16) | (0.14) |
| H2 Public information (*PUBLIC*) | - | -1.9829 | -2.3427** | -1.2754 | -1.9819* |
| | | (0.12) | (0.05) | (0.20) | (0.07) |
| H3 Litigation risk (*KS_LIT*) | - | 2.5881** | 3.1251*** | 1.3179 | 2.2467*** |
| | | (0.05) | (0.01) | (0.18) | (0.01) |
| H4 Equity incentives (*%CRSP*) | + | 0.0942** | 0.0878** | 0.0489 | 0.0340 |
| | | (0.02) | (0.03) | (0.15) | (0.32) |
| Industry concentration (*HERF*) | ? | 0.4291 | 0.4585 | 0.6414* | 0.7749** |
| | | (0.19) | (0.16) | (0.06) | (0.03) |
| Pseudo $R^2$ | | 20.84% | 20.43% | 25.16% | 22.46% |
| Wald $\chi^2$ test statistic | | 13.71 | 13.73 | 11.89 | 11.91 |
| Wald $\chi^2$ p-value | | 0.03 | 0.02 | 0.06 | 0.04 |

**Panel B: Collusion episodes restricted to positive abnormal FOG for influential firms ($COLLUDE_{INFL}$)**

- **FULL sample: 12 collusion episodes; 125 non-collusion industry-year observations**
- **1SEG sample: 9 collusion episodes; 119 non-collusion industry-year observations**

| | Predicted Sign | FULL sample | | 1SEG sample | |
|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) |
| Intercept | | -0.1617 | -0.1882** | -0.0787 | -0.0563 |
| | | (0.16) | (0.04) | (0.40) | (0.40) |
| H1 Heterogeneity (*HETERO)* | ? | -0.1354 | | 0.0948 | |
| | | (0.74) | | (0.66) | |
| H2 Uncertainty (*TAILRISK)* | + | -0.0169 | -0.0158 | -0.0422 | -0.0391 |
| | | (0.58) | (0.61) | (0.23) | (0.24) |
| H2 Public information (*PUBLIC*) | - | -2.1770** | -2.0554*** | -1.0615* | -1.1568* |
| | | (0.02) | (0.01) | (0.09) | (0.07) |
| H3 Litigation risk (*KS_LIT*) | - | 1.7023* | 1.5531** | 0.4294 | 0.5129 |
| | | (0.06) | (0.03) | (0.31) | (0.23) |
| H4 Equity incentives (*%CRSP*) | + | 0.0576** | 0.0612** | -0.0048 | -0.0092 |
| | | (0.04) | (0.02) | (0.86) | (0.69) |
| Industry concentration (*HERF*) | ? | 0.0808 | 0.0727 | 0.3405 | 0.3346 |
| | | (0.71) | (0.74) | (0.17) | (0.18) |
| Pseudo $R^2$ | | 22.89% | 22.76% | 33.98% | 33.59% |
| Wald $\chi^2$ test statistic | | 10.23 | 10.35 | 4.34 | 4.10 |
| Wald $\chi^2$ p-value | | 0.12 | 0.07 | 0.63 | 0.53 |

**Table 6: Falsification test: Characteristics of industries that have transparency episodes**

This table presents estimated marginal effects from a logit regression model of industry years that exhibit significant decreases (rather than increases) in FOG, which we call transparency episodes, and industry and industry-year level correlates. The unit of observation is at the industry-year level. The table presents results for collusion and transparency episodes determined by a significant industry-level increase in FOG in two-year rolling regressions. Results are presented for the FULL sample (Columns 1 and 2) and 1SEG samples (Columns 3 and 4). Columns (1) and (3) report results for models (1) and (3), respectively, from Table 5 for the FULL and 1SEG sample collusion episodes for convenient comparison. Columns (2) and (4) provide the results for the transparency episodes. The unit of observation is at the industry-year level. See Table 4 for definitions of the independent variables. P-values of the $\chi^2$ test for coefficient significance are presented in brackets. Statistical significance (two-sided) at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.

| | Predicted Sign | **Panel A: FULL sample** | | **Panel B: 1SEG sample** | |
|---|---|---|---|---|---|
| | | Collusion (1) | Transparency (2) | Collusion (3) | Transparency (4) |
| Intercept | | -0.4232*** | -0.2230*** | -0.3794*** | -0.2571** |
| | | (0.00) | (0.00) | (0.00) | (0.01) |
| H1 Heterogeneity (*HETERO)* | ? | 0.4198 | 0.6527*** | 0.7079 | 0.8606** |
| | | (0.51) | (0.00) | (0.13) | (0.03) |
| H2 Uncertainty (*TAILRISK)* | + | -0.0561 | 0.0079 | -0.0628 | -0.0305 |
| | | (0.26) | (0.70) | (0.16) | (0.45) |
| H2 Public information (*PUBLIC*) | - | -1.9829 | 0.2656 | -1.2754 | 0.0232 |
| | | (0.12) | (0.46) | (0.20) | (0.89) |
| H3 Litigation risk (*KS_LIT*) | - | 2.5881** | -0.4745 | 1.3179 | -0.9623 |
| | | (0.05) | (0.41) | (0.18) | (0.41) |
| H4 Equity incentives (*%CRSP*) | + | 0.0942** | 0.0308 | 0.0489 | 0.0226 |
| | | (0.02) | (0.11) | (0.15) | (0.62) |
| Industry concentration (*HERF*) | ? | 0.4291 | -0.2176 | 0.6414* | -0.3210 |
| | | (0.19) | (0.19) | (0.06) | (0.46) |
| Pseudo $R^2$ | | 20.84% | 13.88% | 25.16% | 7.68% |
| Wald $\chi^2$ test statistic | | 13.71 | 11.81 | 11.89 | 6.65 |
| Wald $\chi^2$ p-value | | 0.03 | 0.07 | 0.06 | 0.35 |

**Table 7: Further analysis of *%CRSP* as a determinant of collusion episodes**

This table presents estimated marginal effects from logit regression models of collusion episode occurrence on industry and industry-year level correlates allowing for a separate relation between firms with low and high equity incentives as measured by *%CRSP*. The unit of observation is at the industry-year level. Results are presented for the FULL sample. Panels A and B present results for collusion episodes determined by a significant industry-level increase in FOG in two-year rolling regressions ($COLLUDE_{AVG}$) and by a significant industry-level increase in FOG combined with positive abnormal FOG for influential firms ($COLLUDE_{INFL}$), respectively. Columns (1) and (4) report results for the first model from Table 5 with an additional interaction term of *%CRSP* and *LOW*, which is an indicator variable for industry-year observations that have a *%CRSP* less than or equal to the median industry-year level. Columns (2), (3), (5) and (6) report results for the subsample of observations with low equity incentives (LOW = 1) and high equity incentives (LOW = 0), respectively. P-values of the $\chi^2$ test for coefficient significance are presented in brackets. Statistical significance (two-sided) at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.

| | | **Panel A: Collusion episodes identified by significant increase in industry-average FOG ($COLLUDE_{AVG}$)** [18] | | | **Panel B: Collusion episodes restricted to positive abnormal FOG for influential firms ($COLLUDE_{INFL}$)** [19] | | |
|---|---|---|---|---|---|---|---|
| Independent variable | Predicted Sign | (1) | (2) LOW %CRSP | (3) HIGH %CRSP | (4) | (5) LOW %CRSP | (6) HIGH %CRSP |
| Intercept | | -0.3908*** | -0.4939** | -0.4716* | -0.1520 | -0.1364 | -0.0565 |
| | | (0.01) | (0.04) | (0.09) | (0.18) | (0.48) | (0.60) |
| H1 Heterogeneity (*HETERO)* | ? | 0.4014 | 0.1348 | 0.8905 | -0.1474 | -0.7693 | 0.0802 |
| | | (0.53) | (0.86) | (0.42) | (0.72) | (0.24) | (0.69) |
| H2 Uncertainty (*TAILRISK)* | + | -0.0500 | -0.0049 | -0.1539* | -0.0150 | -0.0223 | -0.0093 |
| | | (0.32) | (0.92) | (0.10) | (0.63) | (0.61) | (0.61) |
| H2 Public information (*PUBLIC*) | - | -1.9410 | 1.0868 | -4.7864** | -2.1541** | -0.6101 | 0.8418 |
| | | (0.12) | (0.48) | (0.02) | (0.02) | (0.65) | (0.51) |
| H3 Litigation risk (*KS_LIT*) | - | 2.5742* | 2.0742 | 1.8319 | 1.7368* | 2.2060** | 0.2996 |
| | | (0.06) | (0.12) | (0.35) | (0.06) | (0.05) | (0.58) |
| H4 Equity incentives (*%CRSP*) | + | 0.0820* | 0.1908 | 0.1441* | 0.0524* | 0.0636 | 0.0253 |
| | | (0.08) | (0.29) | (0.06) | (0.09) | (0.69) | (0.53) |
| *%CRSP*LOW* | | -0.0313 | | | -0.0135 | | |
| | | (0.60) | | | (0.72) | | |
| Industry concentration (*HERF*) | ? | 0.3971 | 0.3235 | 1.1185* | 0.0827 | 0.1664 | 0.0009 |
| | | (0.21) | (0.48) | (0.10) | (0.70) | (0.69) | (0.99) |
| Pseudo $R^2$ | | 21.11% | 21.02% | 30.79% | 23.05% | 10.67% | 49.30% |
| Wald $\chi^2$ test statistic | | 13.96 | 6.04 | 8.13 | 10.26 | 4.50 | 0.70 |
| Wald $\chi^2$ p-value | | 0.05 | 0.42 | 0.23 | 0.17 | 0.61 | 0.99 |

[18] The FULL sample includes 17 collusion episodes and 120 non-collusion industry-year observations. The LOW (HIGH) %CRSP sample has 7 (10) collusion episodes; 74 (46) non-collusion industry-year observations.

[19] The FULL sample includes 12 collusion episodes and 125 non-collusion industry-year observations. The LOW (HIGH) %CRSP sample has 5 (7) collusion episodes; 76 (49) non-collusion industry-year observations.

*Appendix A: Variable definitions*
This appendix provides detailed definitions of variables that are included in the analysis.

**Control variables used in the FOG model**

| | |
|---|---|
| *MVE*: | Log of market value of equity (CSHO * PRCC_F), winsorized at 1%. |
| *MTB:* | Market to book ratio (*MVE* + LT / AT), winsorized at 1%. |
| *Special items:* | Special items scaled by total assets (SPI/AT), winsorized at 1%. |
| *Return volatility:* | Standard deviation of monthly stock returns for the 12 months ending in the third month of the current year (i.e., from current month three back to lag nine), winsorized at 1%. This variable is set to missing if there are less than 11 months of return data. |
| *Non-missing items:* | Number of non-missing items in Compustat, winsorized at 1%, to proxy for complexity. The total number of items possible is based on all data items for the balance sheet, income statement, and cash flow statement per the WRDS Compustat web interface. |
| *Firm age:* | Number of years a firm has been listed on CRSP. |
| *Delaware:* | An indicator = 1 if the firm was incorporated in Delaware (INCORP= "DE"), 0 otherwise. |
| *GEO Segments:* | Log of the number of geographic segments plus one. Segment information is obtained from the Compustat Segment file and then collapsed into Fama-French 49 industry codes. The variable is set to one if missing in the datafile. |
| *BUS Segments:* | Log of the number of business segments plus one, obtained from the Compustat Segment file (but collapsed based on FF49 industries). The variable is set to one if missing in the datafile. |

**Explanatory variables used in the logit model of the collusion episodes**

<u>*PUBLIC*: Availability of public information</u>
For each of the FF49 industries, we estimate a standard market model (A1) and a factor model (A2) using monthly data between January 1994 and December 2008:

$$r_{im} = \alpha_j + \beta_j^{mkt} r_{mkt,m} + \varepsilon_{im} \tag{A1}$$

$$r_{im} = \alpha_j + \beta_j^{mkt} r_{mkt,m} + \sum_{z=1}^{7} \delta_j^z FACTOR_m^z + \xi_{im} \tag{A2}$$

where $r_{im}$ is the monthly return for firm *i* in industry *j* (*j* = 1 to 49); $r_{mkt,m}$ is the monthly return on the CRSP equally-weighted market index; and $FACTOR_m^z$ is the monthly value of one of seven macro-economic risk factors identified in prior research and described below. *PUBLIC* is the difference between the adjusted $R^2$ values of the two models for each FF49 industry. The greater the difference in adjusted $R^2$s for an industry, the greater is the availability of public information.

The seven factors in Equation (A2) are:

(1) Short-term interest rates = the 3-month treasury bill rate from CRSP.
(2) Default premium = Moody's Seasoned Baa Corporate Bond Yield (available from http://www.federalreserve.gov/releases/h15/data.htm) minus the 10-year constant maturity government bond yield (from http://research.stlouisfed.org/fred2/).
(3) Term premium = the yield on 10-year constant maturity government bonds minus the yield on one-year constant maturity government bonds (from http://research.stlouisfed.org/fred2/).
(4) Foreign Exchange Rates = weighted average of the foreign exchange value of the U.S. dollar against a subset of the broad index currencies that circulate widely outside the country of issue (from http://research.stlouisfed.org/fred2/).
(5) Producer price index = the total finished goods producer price index from the Bureau of Labor Statistics.
(6) Small minus Big (SMB) = the Fama-French monthly benchmark factor for the performance of small stocks relative to big stocks.
(7) High – low book-to-market (HML) = the Fama-French monthly benchmark factor for the performance of value stocks relative to growth stocks.

SMB and HML are from Kenneth French's website:
(http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

*HETERO: Intra-industry heterogeneity*

We estimate a single factor market model using monthly returns for each FF49 industry for each calendar month between January 1994 and December 2008. *HETERO* represents the industry-year average of the standard deviations across all months in each calendar year.

*TAILRISK: Industry tail risk*

Using option data from OptionMetrics, we calculate the Van Buskirk (2011) measure of expected volatility skew before each I/B/E/S earnings announcement that meets his data requirements. This measure is calculated as the average implied volatility for out-of-the-money puts less the average implied volatility for at-the-money calls. We compute skew for 85,293 firm-quarters between 1995 and 2010. We calculate the firm average skew by year to convert this measure to firm-year observations so that firms with more populated data do not influence the measurement. The average (and median) firm-year have volatility skew of approximately 0.03 (higher values indicate more negative skewness).

On average, the industries have 32.8 firms in any given year (median = 20.2). The minimum (maximum) is 1.7 (131.2) per year for industry 1 (industry 36). The inter-quartile range is 5.9 to 50.9. For each industry, we obtained the median and maximum measure of skew over the 13 years for which we have observations. Industries with a maximum greater than the median industry maximum (0.061) and with a median greater than the global industry median of 0.03 are designated as having high skew (*TAILRISK* = 1).

An industry level proxy is preferable to an industry-year level proxy when using a market-based measure to estimate the extent to which an industry is subject to large but uncertain shocks. While this construct may vary over time for an industry, using an annual measure would require assumptions about when the market anticipates the yet unrealized shock, which managers are potentially attempting to hide.

*%CRSP: Stock price incentives to collude*

For each FF49 industry and year, we count the firms with greater than 200 non-missing daily return observations on CRSP and divide by the number of firms in the Compustat fundamentals annual file with an available SIC code and non-missing annual firm-level sales and total assets data.

*KS_LIT: Litigation risk*

Our proxy for litigation risk is based on Model 3 from Table 7 of Kim and Skinner (KS, 2012). This model is intended to capture the factors that make a firm more vulnerable to securities litigation prior to the revelation of "triggering events" like major price declines. In this sense, the model provides an *ex ante* probability of litigation risk. Specifically, we estimate the following model:

$$SUED = b_0 + b_1(FPS_t) + b_2(LNASSETS_{t-1}) + b_3(SALES\_GROWTH_{t-1}) + b_4(RETURN_{t-1}) + b_5(RETURN\_SKEWNESS_{t-1}) + b_6(RETURN\_STD\_DEV_{t-1}) + b_7(TURNOVER_{t-1}) + e$$

We estimate *KS_LIT* for fiscal years between 1996 and 2008 for a sample of 48,844 firm-years (including 2,667 lawsuits). The firm-year observations are averaged to create the industry-year observations. Our sample of sued firms is generated from a database provided by Woodruff-Sawyer, a San Francisco-based insurance brokerage firm. As explained in Rogers and Van Buskirk (2009), Woodruff-Sawyer aggregates data from a number of sources including Securities Class Action Clearinghouse (SCAC), which is the source of the sued firm sample in KS. Estimating the model on this sample allows us to relax some of KS's sample requirements that are not necessary for our study (e.g., they require return data to be available for three years prior to the beginning of each fiscal year). As expected, the sign and significance of each of the independent variables is similar to that reported by KS (see KS for specific variable definitions).

*HERF*: Industry concentration

We compute a revenue-based herfindahl index for each FF49 industry for each year as:

$$HERF\_S_{jy} = \left( \sum_{i=1}^{m} \left( Sales_i \Big/ \sum_{i=1}^{m} Sales_i \right)^2 \right)$$

where $Sales_i$ is revenues for each firm *i* in industry *j* in year *y* and *m* equals 50 if there are at least 50 firms in the industry and equals the number of firms in industry *j* in year *y* if the number is less than 50. We similarly compute a herfindahl index based on market share of total assets.

We also compute industry-year level concentration ratios as the sum of the market shares of sales or total assets for the *n* largest firms in an industry:

$$C_n = \sum_{i=1}^{n} (Market\ Share)_i$$

We compute the concentration ratio for the top four, six and eight firms. Concentration ratios are commonly used in the cartel literature, but there is no consensus on the appropriate *n* to include.

**Appendix B: Summary of explanatory variables across the Fama-French 49 industries**

An asterisk after the industry name indicates that the industry is designated as having high tail risk (*TAILRISK*=1). The remaining columns report the average annual number of firms in each FF49 industry over the period 1994-2008 with available data on Compustat and CRSP, the average annual percent of firms on CRSP relative to Compustat (*%CRSP*), the *PUBLIC* score, and the average industry-year HERF_S.

| | | Industry | | Average annual # of firms | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Compustat | CRSP | %CRSP | PUBLIC | HERF_S |
| 1 | Agric | Agriculture | | 18 | 17 | 94.44% | 0.064 | 0.232 |
| 2 | Food | Food Products | | 90 | 82 | 91.11% | 0.027 | 0.065 |
| 3 | Soda | Candy & Soda | | 11 | 20 | 181.82% | 0.068 | 0.306 |
| 4 | Beer | Beer & Liquor | | 13 | 24 | 184.62% | 0.052 | 0.329 |
| 5 | Smoke | Tobacco Products | | 6 | 8 | 133.33% | 0.174 | 0.822 |
| 6 | Toys | Recreation | * | 53 | 53 | 100.00% | 0.021 | 0.156 |
| 7 | Fun | Entertainment | * | 102 | 87 | 85.29% | 0.017 | 0.143 |
| 8 | Books | Printing and Publishing | | 46 | 59 | 128.26% | 0.025 | 0.064 |
| 9 | Hshld | Consumer Goods | * | 111 | 96 | 86.49% | 0.016 | 0.127 |
| 10 | Clths | Apparel | * | 79 | 62 | 78.48% | 0.036 | 0.070 |
| 11 | Hlth | Healthcare | | 112 | 120 | 107.14% | 0.020 | 0.089 |
| 12 | Medeq | Medical Equipment | | 187 | 187 | 100.00% | 0.015 | 0.074 |
| 13 | Drugs | Pharmaceutical Products | * | 317 | 324 | 102.21% | 0.050 | 0.092 |
| 14 | Chems | Chemicals | | 92 | 103 | 111.96% | 0.033 | 0.059 |
| 15 | Rubbr | Rubber and Plastic Products | | 58 | 44 | 75.86% | 0.032 | 0.070 |
| 16 | Txtls | Textiles | | 33 | 29 | 87.88% | 0.058 | 0.155 |
| 17 | Bldmt | Construction Materials | | 103 | 88 | 85.44% | 0.022 | 0.071 |
| 18 | Constr | Construction | * | 70 | 71 | 101.43% | 0.023 | 0.061 |
| 19 | Steel | Steel Works Etc. | | 76 | 82 | 107.89% | 0.055 | 0.068 |
| 20 | Fabpr | Fabricated Products | | 22 | 20 | 90.91% | 0.052 | 0.141 |
| 21 | Mach | Machinery | | 184 | 177 | 96.20% | 0.019 | 0.055 |
| 22 | Elceq | Electrical Equipment | * | 80 | 136 | 170.00% | 0.019 | 0.230 |
| 23 | Autos | Automobiles and Trucks | * | 80 | 90 | 112.50% | 0.047 | 0.220 |
| 24 | Aero | Aircraft | | 22 | 26 | 118.18% | 0.068 | 0.256 |
| 25 | Ships | Shipbuilding, Railroad Equip. | | 11 | 8 | 72.73% | 0.138 | 0.425 |
| 26 | Guns | Defense | * | 10 | 9 | 90.00% | 0.153 | 0.747 |
| 27 | Gold | Precious Metals | | 21 | 64 | 304.76% | 0.186 | 0.501 |
| 28 | Mines | Non-Metallic/Industrial Metal Mining | | 18 | 29 | 161.11% | 0.093 | 0.192 |
| 29 | Coal | Coal | * | 7 | 12 | 171.43% | 0.254 | 0.251 |
| 30 | Oil | Petroleum and Natural Gas | | 199 | 243 | 122.11% | 0.126 | 0.121 |
| 31 | Util | Utilities | | 224 | 171 | 76.34% | 0.097 | 0.026 |
| 32 | Telcm | Communication | * | 193 | 238 | 123.32% | 0.025 | 0.067 |
| 33 | Persv | Personal Services | * | 69 | 68 | 98.55% | 0.017 | 0.066 |
| 34 | Bussv | Business Services | | 334 | 387 | 115.87% | 0.007 | 0.037 |
| 35 | Hardw | Computers | | 163 | 137 | 84.05% | 0.027 | 0.136 |
| 36 | Softw | Computer Software | | 524 | 447 | 85.31% | 0.031 | 0.150 |
| 37 | Chips | Electronic Equipment | | 338 | 350 | 103.55% | 0.037 | 0.066 |
| 38 | Labeq | Measuring & Control Equip. | | 124 | 122 | 98.39% | 0.022 | 0.093 |
| 39 | Paper | Business Supplies | | 66 | 67 | 101.52% | 0.035 | 0.074 |
| 40 | Boxes | Shipping Containers | | 16 | 17 | 106.25% | 0.066 | 0.140 |
| 41 | Trans | Transportation | * | 139 | 152 | 109.35% | 0.027 | 0.045 |
| 42 | Whlsl | Wholesale | | 234 | 251 | 107.26% | 0.007 | 0.057 |
| 43 | Rtail | Retail | * | 313 | 288 | 92.01% | 0.028 | 0.066 |
| 44 | Meals | Restaurants, Hotels, Motels | * | 115 | 131 | 113.91% | 0.036 | 0.075 |
| 45 | Banks | Banking | | 337 | 723 | 214.54% | 0.069 | 0.056 |
| 46 | Insur | Insurance | * | 207 | 200 | 96.62% | 0.060 | 0.045 |
| 47 | Rlest | Real Estate | | 56 | 49 | 87.50% | 0.045 | 0.146 |
| 48 | Fin | Trading | * | 315 | 1085 | 344.44% | 0.035 | 0.110 |
| 49 | Other | Almost Nothing | | 79 | 25 | 31.65% | 0.047 | 0.261 |