NETWORKS FINANCIAL INSTITUTE
At Indiana State University

Working Paper

2011-WP-09
May 2011

Predicting Fraud by Investment Managers

William C. Gerken and Stephen G. Dimmock

Abstract: We test the predictability of investment fraud using a panel of mandatory disclosures filed with the U.S. Securities and Exchange Commission (SEC). We show that past regulatory and legal violations, conflicts of interest, and monitoring have significant power to predict fraud. Avoiding the 5% of firms with the highest ex ante predicted fraud risk would allow an investor to avoid 29% of fraud cases and over 40% of the total dollar losses from fraud. We examine the ability of investors to implement fraud prediction models based on the disclosure filings, and suggest changes in SEC data access policies that could benefit investors.

About the Authors: The research by Dr. **William Gerken** focuses on investment fraud, financial institutions, corporate governance. He received the Commonfund best paper award at the European Finance Association meeting and best paper in Networks Financial Institute's 2010-11 National Competition for Research Papers on Financial Services Policy Reform. Professor Gerken is an Assistant Professor in the Department of Finance, College of Business, Auburn University and currently teaches Financial Institutions and Markets. Before coming to Auburn University, Dr. Gerken taught derivatives at Michigan State University, where he received the Excellence in Teaching award from the Department of Finance. Prior to entering academia, he worked for Citigroup in Atlanta. He is a CFA charterholder and a member of the CFA Society. **Stephen G. Dimmock** is an Assistant Professor of Finance in the Division of Banking and Finance, Nanyang Business School, at Nanyang Technological University in Singapore.

Keywords: Fraud, Investment Fraud, Operational Risk, SEC, Disclosure, Form ADV.

JEL Code: G2, G20, G28, K2, K22.

The authors are grateful for comments from: two anonymous referees, James Angel, Devraj Basu, Ranadeb Chaudhuri, Elroy Dimson, Will Goetzmann, Philip Hamill, Chuan-Yang Hwang, Dusan Isakov, Zoran Ivkovic, Jun-Koo Kang, Ayla Kayhan, Naveen Khanna, Clive Lennox, Dinah McNichols, Jay Patel, Josh Pollet, Melvyn Teo, Fan Yu, Lei Zhang, seminar participants at Auburn and Nanyang Technological Universities, Singapore Monetary Authority, 2010 Asian FMA, Current Topics in Securities Regulation Conference at Notre Dame, One-Day Conference on Professional Asset Management at Erasmus University, 2010 EFA, and Fourth Singapore International Conference on Finance, 2010 CRSP Forum, 2010 FMA, and the Emerging Scholars in Banking and Finance conference. They are especially grateful to Scott Weisbenner for numerous helpful discussions and acknowledge the financial support of Networks Financial Institute at Indiana State University and its award for best paper in Financial Services Regulatory Reform for 2010-11.

Disclosure lies at the heart of a well-functioning market, serving as both a catalyst to attract investment and a deterrent against fraud. For investment advisers in the United States, required disclosures are made through the Securities and Exchange Commission (SEC) using Form ADV, which requires investment advisers to provide information about their operations, conflicts of interest, disciplinary histories, and other material facts that could help current and potential clients make informed decisions.

This paper argues that a method for identifying investment advisers with high fraud risk would benefit not only market participants but also the policy makers and regulators charged with market oversight. It asks whether the information these firms are required to disclose is useful for predicting investment fraud, and if so, which disclosures can be used to predict fraud. It concludes that the disclosures in Form ADV can be used to predict fraud, but that there are significant barriers to implementing prediction models, particularly because of the way public access to this information is provided.

This analysis employs an annual panel of Form ADVs filed from August 2001 through July 2006.[1] The panel includes 13,853 investment advisers who advise more than 20 million clients and control more than $32 trillion in assets. These firms advise all U.S. mutual funds, nearly all institutional investment funds, and many hedge funds. This panel of historical Form ADV filings is not publicly available, and we are the first researchers to use these data. The analysis also includes a review of all SEC litigation actions and administrative proceedings from August 2001 through July 2010 to identify those cases in which investment advisers defrauded their clients. We then test whether the information disclosed in Form ADV can be used to predict investment fraud.

The findings show that Form ADV disclosures related to past regulatory violations,

[1] We are able to obtain Form ADV data for only the period 2001 through 2006. We form an annual panel using data as of August 1 of each year to maximize the number of years, as our sample of Form ADV filings ends July 31, 2006.

conflicts of interest, and monitoring are all significant in the fraud prediction models. Within the sample, an investor who avoided the 5% of firms with the highest predicted fraud risk would avoid 29% of fraud cases and over 40% of the dollar losses related to fraud. Out-of-sample tests show that the predictability of fraud is robust. If the Form ADV data were not useful for predicting fraud, this would suggest that either disclosure is so effective that it eliminates the predictability that would occur in the absence of disclosure or that the disclosed information is worthless. Our findings thus provide evidence that regulators require investment advisers to disclose relevant information.

These findings are subject to several limitations. First, only *detected* fraud cases are included in the prediction models. Although we conduct extensive out-of-sample tests of the predictive models, we cannot reject the possibility that the absence of undetected fraud cases biases the prediction models. Second, while we find that Form ADV disclosures such as conflicts of interest do predict fraud, the results cannot be interpreted to imply that conflicts of interest alone cause fraud or that their prohibition would result in a reduction in fraud. Prediction alone does not imply causality, as the predictive variables may be jointly determined with the decision to commit fraud. Third, prediction is not the sole purpose of disclosure: It is also intended to deter fraud by providing investors with insight into a firm's activities and conflicts of interest, and allowing investors to make informed investment decisions. The tests in this analysis do not address the deterrent effect of disclosure, and the results are limited in scope to whether disclosed information does predict fraud.

In the main predictive regressions, the dependent variable includes two types of fraud cases: those cases initiated in the year subsequent to the Form ADV disclosure, and cases continued from the previous year. It seems reasonable to assume that predicting fraud cases before they are initiated produces greater benefits because it reduces the total losses from fraud. Thus we are interested in testing whether it is relatively more difficult to predict the initiation of fraud. To test whether Form ADV disclosures can be used to predict the actual initiation of fraud, we estimate a multinomial probit regression, using separate equations for fraud cases initiated in the subsequent year and fraud cases that were initiated in prior years but continued into the subsequent year. The results show that it is possible to predict the initiation of fraud at least as accurately as the continuation of previously initiated, but as-of-yet undetected, fraud cases.

We next compare two types of predictive models, both of which take the perspective of an investor attempting to implement a fraud prediction model during the sample period. The first type of model predicts fraud using only the limited subset of information that would have been publicly accessible. Until 2010, the general public had access to only contemporaneous cross-sections of Form ADV filings; thus, the independent variables in these tests are taken from the contemporaneous cross-sections of Form ADV filings. These tests show what would have been possible during our sample given the actual data access policies in place. The second type of model predicts fraud using a panel of prior Form ADV filings. These tests show what would have been possible if the historical Form ADV filings have been available contemporaneously during the sample. The predictive models based on a panel of historical Form ADV filings are moderately better at predicting fraud out-of-sample. We discuss simple changes to data access policies that could improve investors' ability to predict fraud.

1. Related Research

To our knowledge, just two papers, Bollen and Pool (2010) and Zitzewitz (2006), test whether it is possible to detect fraud by investment advisers. Bollen and Pool (2010) build on earlier studies that test whether hedge funds manipulate reported returns [Bollen and Pool (2008); Bollen and Pool (2009); and Straumann (2008)] and find that suspicious return patterns can be used to predict fraud charges. Zitzewitz (2006) shows that daily fund flows provide information about late trading in mutual funds. Although these papers, like ours, develop methods to detect fraud, they analyze returns and fund flows rather than firms' disclosures of business practices and conflicts of interest. An advantage of using firms' disclosures is that we can actually predict fraud, whereas methods based on returns and flows can detect only past or ongoing fraud. A further advantage is that Form ADV disclosures are mandatory, whereas the disclosure of returns is optional for many investment advisers.

Brown, Goetzmann, Liang, and Schwarz (2008, 2009) examine operational risk using a cross-section of Form ADV filings from hedge fund advisers. The authors define "problem" funds as those managed by an adviser reporting prior legal or regulatory violations, either committed by the adviser itself or an affiliated firm. Brown et al. then test whether Form ADV data are associated with prior problems. Because historical Form ADV data are not publicly available, the authors create a measure of operational risk, the ω-score, based on the correlations between contemporaneous Form ADV data and historical hedge fund data. They then test whether the ω-score can predict hedge fund closure, flows, and returns. We also use Form ADV data, but our work differs from Brown et al. in several ways: First, we focus on fraud rather than their very broad definition of operational risk. (Indeed, of the 126 "problem" hedge fund advisers identified by Brown et al., we find that only six have prior incidents of fraud). Second, their measure of operational risk includes violations by affiliated firms, such as broker-dealers. These differences are empirically important; we replicate the ω-score of Brown et al. but find it has an insignificant relation with subsequent fraud. Third, we use the historical Form ADV filings to make ex ante predictions of fraud. Finally, we use a more comprehensive sample of investment advisers, which includes advisers to mutual and pension funds, in addition to hedge funds.

2. Data

2.1. Investment Fraud

This study combines two types of data: (1) investment fraud data and (2) disclosures made by investment advisers in their Form ADV filings. To obtain investment fraud data, we search all SEC administrative proceedings and litigation releases[2] that contain the terms "fraud" and "investment adviser" (or "investment advisor") filed from August 2001 through July 2010. From these documents, we identify all cases that involve violations of the anti-fraud provisions in the Investment Advisers Act. Even when a fraud case is initially detected by another agency, the SEC launches an administrative action, which we observe. The main dependent variable in our paper includes only those cases in which fraud directly harms the firm's investment clients. We do not include insider trading, short sale violations, crimes by the brokerage division of a firm, or other activities, unless these crimes cause direct losses to the firm's investment clients.

Many fraud cases span several years and involve multiple legal actions. Figure 1 shows the timeline of a fraud initiated in September 2002 by K.W. Brown & Co., an investment adviser that traded securities on behalf of clients and for its proprietary account. K.W. Brown & Co. purchased securities but delayed assigning them to specific accounts. Eventually, the firm would allocate profitable trades to its proprietary account and unprofitable trades to clients, resulting in losses of over $9 million to the firm's investors. The SEC uncovered problems in March 2003 during a routine examination and, in June 2003, notified the firm that problems had been identified. The firm continued defrauding clients for nine more months, until March 2004. In April 2005, criminal charges were filed against the firm and its key employees. Nearly two more years passed until the firm and its employees were convicted in December 2007. In January 2008, the SEC filed an administrative proceeding to bar Kevin W. Brown, his wife, and another employee from the securities industry. The firm was deregistered in June 2008.

Because this kind of extended legal scenario is common, and because our goal is to predict fraud rather than detect it, we aggregate all legal actions associated with a single underlying fraud into a single "case" and identify the periods in which fraud occurred. For example,

[2] See `http://www.sec.gov/litigation/admin.shtml` and `http://www.sec.gov/litigation/litreleases.shtml`.

we define the K.W. Brown & Co. fraud case as occurring from September 2002 until March 2004, and use the information from K.W. Brown & Co.'s August 2002 Form ADV filing to predict the initiation of fraud in September 2002. We also use information from K.W. Brown & Co.'s August 2003 Form ADV filing to predict the continuation of fraud into 2004. For the remaining years of the sample, we classify K.W. Brown & Co. as a clean firm. By predicting the occurrence of fraud rather than its detection, we avoid potential biases caused by a correlation between detection and time variation in the predictive variables.

We collect information on all investment fraud cases between August 2001 and July 2010, including cases committed by those firms that are not registered with the SEC, and thus are not required to file Form ADV (see the next subsection for more detail). To address the economic importance of fraud committed by registered versus non-registered investment advisers, Panel A of Table 1 summarizes fraud cases for both types of firms. Slightly over half of investment fraud cases are committed by registered firms, and these firms are responsible for the overwhelming majority of the dollar losses from fraud. Thus although the scope of our tests is limited to registered investment advisers, these firms are responsible for the most economically meaningful fraud cases.

Panel B summarizes firm-wide fraud, committed with the knowledge of a firm's executive officers, as well as fraud by rogue employees who evade their firms' internal controls. The vast majority of fraud cases are firm-wide. Panel B also summarizes the dollar losses and the duration of the fraud cases. Because fraud often involves the falsification of records, it is difficult to determine exact losses. Thus, some loss amounts are simply not available; and when amounts are provided, they are generally a lower bound, reflecting only the proven losses. Fraud duration is defined as the period extending from the initiation of the fraud until the SEC's first relevant legal filing. The median fraud case persists for five and a half years before detection. The maximum durations, summarized in the last column of Panel B, reflect the fact that the sample includes cases that were initiated prior to 2001.

To test whether past fraud is a predictor of future fraud, we collect additional information on fraud by investment advisers by searching SEC administrative proceedings and litigation releases filed from September 1995[3] through July 2001, and create two variables. The first, Past Fraud, is an indicator variable equal to one if a prior administrative proceeding or litigation release indicates the firm has committed fraud. The second, Past Affiliated Fraud, is an indicator variable equal to one if a prior administrative proceeding or litigation release indicates an affiliated firm has committed fraud (affiliation implies the firms are under common control, such as common ownership or executives). To be consistent with the main dependent variable, for both Past Fraud and Past Affiliated Fraud we include only cases that harmed investment advisory clients. We match Past Affiliated Fraud to investment advisers using the affiliated firm identifiers disclosed in Schedule D of Form ADV. To prevent a look-ahead bias in the predictive regressions, these variables include only cases in which the fraud has ceased and was publicly revealed prior to the date of the other predictive variables. For example, in the K.W. Brown & Co. case summarized in Figure 1, Past Fraud would equal one only after April 2005 when the fraud had ceased and the first relevant SEC legal filing was publicly accessible.

2.2. Form ADV Data

The Investment Advisers Act,[4] which expressly defines and prohibits investment adviser fraud, requires all investment advisers with more than $25 million in assets under management and with 15 or more U.S. clients to register with the SEC. The Act defines an investment adviser as any entity that receives compensation for managing securities portfolios for clients or that provides advice regarding individual securities.[5] Registered investment advisers must

[3]Online access to administrative proceedings and litigation releases begins in September 1995.

[4]The other major law governing investment managers is the Investment Company Act, which covers companies targeting retail investors such as mutual funds. The Investment Company Act provides additional investor protection and requires additional disclosure filings.

[5]Section 203(b)(3) of the Investment Advisers Act exempts firms with fewer than 15 U.S. clients during the preceding 12 months and that do not advise funds registered under the Investment Company Act or

file Form ADV to disclose past regulatory violations and potential conflicts of interest.

Form ADV contains 12 items and four schedules. Items 1 through 6 contain descriptive information about a firm and its operations. Items 7 and 8 require disclosure of certain conflicts of interest. Item 9 requires disclosure regarding the custody of clients' assets. Item 10 requires disclosure of control persons. Item 11 requires disclosure of past legal and regulatory violations. Item 12 reports information about small businesses. Schedules A, B, and C identify the direct and indirect owners of the firm. Schedule D requires disclosure of affiliations with other financial firms.

The SEC provides a public link on its website to an Investment Adviser Registration Depository, which includes the most recent Form ADV filings from all registered investment advisers.[6] Until recently, the latest filings could be accessed only one at a time, and past filings were unavailable. Beginning in January 2010, the SEC began to provide downloadable files of historical Form ADV data.[7] Downloadable files from July 2006 through November 2009 contain summaries of the schedules rather than Form ADV's line-item data. Downloadable files from December 2009 until the present day contain the line-item data, but not the schedule data.

The SEC provided us with a database of all Form ADV filings from August 2001 through July 2006, including initial filings, amendments, schedules, and the filings of now-defunct firms. These data are not publicly accessible, and to our knowledge no other researchers have examined them. To create an annual panel for the predictive regressions, we select each firm's most recent filing as of August 1 of each year.[8] This annual panel includes 53,994 firm-year

"hold themselves out to the public" as investment advisers. Some hedge funds use this exemption to avoid registration. An SEC ruling in 2004 required hedge fund advisers to register by February 1, 2006, but a U.S. District Court reversed this ruling in June 2006. Despite these exemptions, many hedge fund advisers did register prior to 2006, either voluntarily or because they also advised other investment portfolios or had more than 15 clients.

[6]See `http://www.sec.gov/IARD`.

[7]See `http://www.sec.gov/foia/docs/invafoia.htm`.

[8]Firms must file Form ADV at least once per year, but often file more frequently; the median number of

observations representing 13,853 unique investment management firms. We match the SEC investment fraud documentation and Form ADV data using the firms' full legal names.[9]

2.2.1. Form ADV Variables

Table 2 summarizes a cross-section of the investment advisers' characteristics and disclosures, using information from each firm's first Form ADV filing during the sample. Panel A shows that the median firm is wholly employee-owned. Employee Ownership, calculated as in Dimmock, Gerken, and Marietta-Westberg (2011), is included because external owners may deter fraud by monitoring employees. The Average Account Size is $55 million, but this variable is highly skewed; the mean is only $1.4 million. Percent Client Agents is the percentage of the firm's clients who are agents (e.g., pension fund managers) rather than the direct beneficiaries of the invested funds. On average, 23.2% of a firm's clients are agents. This additional layer of agency is potentially related to fraud in two ways: Agents have weaker incentives to monitor investment advisers but may also have greater expertise and financial sophistication. Assets Under Management (AUM) varies greatly across firms: The median AUM is $90 million, but the mean is greater than $2.2 billion.

Panel B of Table 2 tabulates many of the variables disclosed in Form ADV (see Appendix A for detailed definitions). Column one shows summary statistics for the full sample. Column two shows summary statistics for firms in which no fraud is committed from the date of their first Form ADV filing through July 2007 (clean firms). Column three shows summary statistics for firms in which fraud is committed during the sample period (fraud firms). The third column also reports the univariate significance of the difference between the clean and fraud samples, using Fisher's exact test.

---

filings per firm-year is 11. We choose August 1 in order to maximize the number of annual observations since our set of Form ADV filings ends July 31, 2006.

[9] Of the 271 fraud cases committed by registered investment advisers, we are unable to match 13 because these firms ceased filing Form ADV before the sample began.

Item 11 of Form ADV requires each investment adviser to disclose its disciplinary history, as well as that of its (non-clerical) employees, its affiliated firms, and the employees of affiliated firms. The 24 questions in Item 11 are divided into three categories: criminal, regulatory, and civil judicial. From these questions, we create two indicator variables. Past Regulatory equals one if the firm discloses past regulatory violations, indicating sanctions by the SEC, the Commodity Futures Trading Commission, or a self-regulatory organization such as the Financial Industry Regulatory Authority (FINRA). The second variable combines the remaining two categories; Past Civil or Criminal equals one if the firm discloses unfavorable civil judicial decisions related to investment advising, or if the firm discloses criminal convictions. Fraud firms are significantly more likely to report both types of violations.

The disclosure information in Item 11 covers a wide range of regulatory and legal offenses, and the offences are often quite minor, such as failing to follow protocols for record storage. Minor violations seem to be the norm rather than the exception, and should be interpreted as such: Less than 2.5% of firms that report past violations have a prior instance of fraud. But because Form ADV does not distinguish whether the investment adviser or its affiliate(s) committed the reported violations, there is a strong positive correlation between prior violations and the presence of affiliated firms. To avoid a spurious correlation between the prior violations of affiliated firms and investment adviser fraud, our dependent variables do not include fraud committed by affiliated firms.

Items 7 and 8 of Form ADV require firms to disclose conflicts of interest. From this information we create three variables. Referral Fees equals one if the firm compensates other parties for client referrals. Interest in Transactions equals one if the firm trades directly with its clients or has a direct financial interest in securities recommended to its clients; these practices create potential conflicts and provide a mechanism for fraud. Soft Dollars equals one if the firm directs clients' trades to a brokerage with relatively high commissions and, in return, the broker supplies the adviser with research or other benefits. Since clients pay the costs while the investment adviser realizes the benefits, soft dollars create a potential conflict of interest.

The next four variables are intended to measure monitoring. Broker in Firm equals one if the firm employs registered representatives of a broker-dealer. Trading through an affiliated broker-dealer removes one form of external oversight and provides a mechanism for fraud. Investment Company Act equals one if the firm manages money on behalf of a fund registered under the Investment Company Act, such as a mutual fund. The Act increases regulation and disallows certain conflicts of interest but also indicates the firm's investors are relatively unsophisticated. Custody equals one if the firm has possession, or the authority to obtain possession, of its clients' assets. Custody facilitates fraud by removing external oversight. However, SEC Rule 206(4)-2 requires audits of investment advisers with such custody, including at least one unannounced visit per year, which may reduce fraud. Dedicated CCO equals one if the firm's chief compliance officer (CCO) does not have another formal job title. All registered investment firms must designate a CCO who is responsible for ensuring compliance with SEC regulation, but often the CCO has other potentially conflicting roles within the firm.

Hedge Fund Clients equals one if over 75% of the firm's clients are hedge fund clients. Slightly over 13% of the firms in our sample primarily advise hedge funds, but only 6.2% of fraud firms advise hedge funds. We include this variable for two reasons: First, hedge funds are relatively opaque, which could facilitate fraud. Second, prior to 2006 some hedge fund advisers were not required to file Form ADV, which could create a sample selection bias if non-reporting is associated with fraud.

3. Predicting Fraud

In this section, we test whether the Form ADV data can be used to predict investment fraud. The purpose of these tests is prediction and, as noted previously, we make no claims regarding causality. Many of the independent variables are endogenous (e.g., a firm's executives may deliberately choose an organizational structure that enables fraud), but because our goal is prediction rather than establishing causality, the potential endogeneity of the independent variables does not change the interpretation.

A major caveat in interpreting our findings is that we observe only detected fraud. Three factors affect observed fraud: the *un*observable true rate of fraud, the probability of detection given a fixed level of monitoring, and the allocation of monitoring resources. Ideally, the regressions will predict the true rate of fraud. However, if certain predictive variables are correlated with either monitoring or detection, this could affect the interpretation of the results. Further, the predictive variables could be correlated with monitoring and detection for two reasons. First, any predictive variable that decreases the probability of detection will increase the incentive to commit fraud. In general, this problem biases against significant results because predictive variables that are associated with a higher rate of fraud will also be associated with a lower detection rate. Second, if the difficulty of detecting fraud affects the allocation of monitoring resources, this may, or may not, outweigh the added difficulty of detecting fraud. These difficulties could cause the empirically observed relations to differ from the true relation between firms' disclosures and the unobservable true rate of fraud.

We address the issue of undetected fraud in two ways. First, even though the panel of independent variables ends in 2006, we search for detected fraud cases through July 2010. For each case, we identify when the fraud occurred. The occurrence of fraud in a given year is the dependent variable in the predictive regressions, even if the fraud remains undetected for years. Second, we test the relation between the fraud prediction variables and the duration of detected fraud cases. One can reasonably assume that fraud cases that have a low probability of detection will also have a longer duration. The results in Appendix B show that none of the independent variables are statistically significant. While these regressions have low power because of the small number of observations, the findings are suggestive of the possibility that the detection rate does not drive the results. Unfortunately, a direct test of the relation between these variables and fraud detection is not possible. Certain types of fraud may go undetected, a possibility that could bias the results.

3.1. Prediction Models

Panel A of Table 3 shows the results of probit regressions that predict investment fraud using Form ADV disclosures. In column one, the sample is a cross-section of firms. The independent variables are taken from each firm's first Form ADV filing during the sample period; the dependent variable equals one if the firm commits fraud *at any time* between the first filing and July 2007. This specification includes indicator variables for the year in which the firm's first Form ADV was filed. The z-scores are based on robust standard errors.

In the remaining columns, the sample is an unbalanced panel of firm-year observations. The dependent variable equals one if a fraud occurs during the subsequent 12 months. In columns two and three, the sample includes all firm-year observations. In column four, the sample excludes firms with prior fraud cases identified in SEC administrative proceedings and litigation releases. In the last column, the sample also excludes firms that disclose the more minor legal or regulatory violations in Item 11, committed either by the firm itself or an affiliated firm. The z-scores, reported below the coefficients, are based on standard errors clustered by firm and year. The chi-square tests at the bottom of each column show the significance of the overall model.

Past Fraud is insignificant in both the cross-sectional and panel regression. There are few firm-year observations with prior fraud, and so the regressions have low power with respect to this variable. Past Affiliated Fraud does not predict subsequent fraud in the cross-sectional regression, but in one of the panel regression specifications, the coefficient is marginally significant and negative. Unlike the other predictive variables, Past Fraud and Past Affiliated Fraud are not disclosed in Form ADV.

Past Regulatory and Past Civil or Criminal are both significant positive predictors of subsequent fraud, even in the sample that excludes firms with prior fraud. The simplest explanation is that past problems, although frequently minor, indicate poor internal controls or unethical management. But two additional explanations exist: Past violations could increase the rate of detected fraud due to the increased probability of an SEC examination. Also, because each firm must disclose both its own prior violations and those of its affiliated firms, prior violations are strongly correlated with the size and scope of an investment firm's affiliated businesses (i.e., financial conglomerates are more likely to report prior violations). These affiliations could increase conflicts of interest and provide the means to commit fraud.

The next three variables measure several potential conflicts of interest between investment advisers and their clients. Referral Fees has a significant positive relation with subsequent fraud. Fraud firms could be relatively willing to pay referral fees because fraud increases the marginal profit per dollar managed. Interest in Transaction also has a significant positive relation with subsequent fraud. When investment managers take the opposite side of a transaction from their clients, this creates an obvious conflict of interest and also provides a mechanism for fraud. Soft Dollars does not significantly predict fraud.

We include several variables to measure the monitoring of investment advisers. Broker in Firm has a significant positive relation with subsequent fraud. Trading through an in-house brokerage removes external oversight and creates a mechanism for committing certain types of fraud. Investment Company Act has a significant positive relation with subsequent fraud. The Act increases regulatory oversight of these firms, which may increase the probability fraud is detected. Alternatively, the true rate of fraud may be higher because these firms exploit their clients' lack of financial sophistication. The next three variables, Custody, Dedicated CCO, and Majority Employee Owned, are not significant in the panel regressions, although Dedicated CCO is significant in the cross-sectional regression. Note that while some variables are insignificant in these regressions, it should not be inferred that there are no benefits from their disclosure, as disclosure may deter fraud.

The next three variables (Logarithm of Average Account Size, Percent Client Agents, and Hedge Fund Clients) also measure monitoring but are based on client characteristics. Although all clients have an incentive to monitor their investments, large investors have a stronger incentive and possibly a greater ability to do so. The results for the Logarithm of Average Account Size show that larger investors are associated with fewer subsequent fraud cases. This could be a selection effect, meaning that large investors select honest managers. Alternatively, because of financial sophistication or economies of scale in monitoring, large investors may deter fraud because of a higher probability of detection. Both arguments suggest that large investors are associated with a lower rate of fraud rather than a lower detection rate.

Percent Client Agents, the second variable measuring client characteristics, has a significant positive relation with subsequent fraud. After conditioning on average account size, firms whose clients include a high proportion of agents are more likely to commit fraud. Although agents may have reputational concerns and greater financial sophistication, they do not bear the full cost of fraud, which reduces their incentive to monitor and suggests that they can be swayed through gifts or kickbacks. The reduced incentives of agents appear to outweigh their potentially higher sophistication.

Hedge Fund Clients is an indicator for firms that primarily manage hedge funds. The results do not provide evidence of a relation between hedge fund management and fraud. Hedge funds are relatively non-transparent, however, and so the detected fraud cases may understate the true amount of fraud that occurs within hedge funds. Moreover, not all hedge funds were required to file Form ADV during the early part of the sample, which could create a sample selection bias. Nonetheless, in annual cross-sectional regressions (Table 5) we find that the coefficient on hedge fund management is not significantly different in the later years of the sample, which suggests that sample selection is not a problem.

3.2. The Economic Interpretation of the Prediction Models

The probit regressions in Panel A of Table 3 show that the Form ADV variables have a statistically significant relation with subsequent fraud. The key question of interest, however, is not the statistical significance of the individual predictive variables but: Could the overall model enable an investor to avoid fraud? To address this question, we take the predicted values from the regressions in Panel A of Table 3 and examine the tradeoff between correctly predicted fraud cases and the false positive rate. False positives, which occur when the model incorrectly predicts that a clean firm will commit fraud in the subsequent year, can be interpreted as the opportunity cost to investors of erroneously limiting their investment opportunity set. Although failing to predict fraud is likely more costly than mistakenly avoiding an honest investment adviser, we do not take a strong position on cost asymmetry and instead illustrate the possible tradeoffs.

Figure 2 shows a receiver operating characteristic (ROC) curve for the prediction model in the second column of Panel A of Table 3. The points on the ROC curve are generated non-parametrically by taking each observation's predicted value from the probit model as a cut-point, and then computing both the proportion of fraud firm-years correctly predicted and the false positives. Random prediction of fraud would result in a straight 45-degree line. Initially the curve rises steeply, showing that a considerable number of fraud firm-years can be avoided at a low false positive rate.

The ROC curve in Figure 2 shows the full range of all possible tradeoffs between the prediction of fraud and false positives. Following a similar format as Dechow, Ge, Larson, and Sloan (2010), in Panel B of Table 3 we provide greater detail for one possible tradeoff between the prediction of fraud and the false positive rate. Specifically, Panel B shows the proportion of fraud firm-years that could be predicted within sample at a false positive rate of 5%. The columns in Panel B correspond to the columns in Panel A. For example, the model in the second column correctly predicts 150 of 517 fraud firm-years (29.0%) at a false positive rate of 5% (we incorrectly predict fraud in 2,673 clean firm-years that are associated with 885 distinct firms). The last row of Panel B shows the percentage of total dollar losses that could have been avoided at a false positive rate of 5%. The dollar losses from fraud are winsorized at 99^{th} percentile because of outliers; for multiyear fraud cases, the losses are evenly distributed across years. The model in the second column correctly predicts 41.3% of the total dollar losses from fraud at a false positive rate of 5%, which indicates that the model predicts economically meaningful fraud cases and not merely small cases.

The results in Panel B are similar for all models, except for the specification reported in the last column, in which the sample does not include firms that report prior legal or regulatory violations, either by the firm or its affiliates. For this sample, both the percentage of fraud firm-years predicted and the percentage of the total dollar losses to fraud avoided are substantially lower. By comparison, the results for the sample that excludes firms with prior publicly revealed fraud, shown in the fourth column, are very similar to the full sample. Thus, the difference in the last column is not due to some firms committing fraud numerous times. Rather, it indicates that fraud is relatively easy to predict among firms with past regulatory and legal violations.

3.3. Out-of-Sample Prediction of Fraud

A key concern for any prediction model is out-of-sample validity. In this subsection, we test whether the predictions made within sample, reported in Panel B of Table 3, are robust out-of-sample. We do this in two ways: Panel C summarizes the out-of-sample predictive performance of each model, using Form ADV filings as of August 2006 to predict fraud that occurred from August 2007 through July 2010. Panel D shows the results from K-fold cross-validation tests, which are explained in the next subsection.

The prediction models reported in Panel A of Table 3 are estimated on a sample that includes only firm-years prior to August 1, 2007. To conduct an out-of-sample test of these prediction models, we search the SEC litigation releases and administrative actions and identify cases of investment fraud that occurred between August 1, 2007 and July 31, 2010. Each firm in the sample on August 1, 2006 receives a predicted value from the within-sample regressions reported in Panel A of Table 3. We then test whether these predicted values are able to accurately classify the out-of-sample fraud risk of the firms.

Panel C of Table 3 shows the proportion of fraud cases correctly predicted at a false positive rate of 5%. The proportion of fraud cases predicted out-of-sample is usually higher than within sample, although given the small number of observations this difference is not statistically significant. Also, although we use the within-sample cutoff values to classify the firms out-of-sample, the false positive rate does not increase.

3.3.1. K-Fold Cross-Validation Tests

As a further robustness test of the prediction models in Panel A of Table 3, we perform K-fold cross-validation tests over the period August 2001 through August 2007. The idea behind these tests is simple. Each model is estimated on a randomly selected subsample of firms, and the coefficient estimates from this subsample are used to classify the firms in the hold-out sample. Specifically, each firm in the sample is randomly assigned to one of 10 groups (note that we randomly assign *firms*, and not *firm-years*, to avoid overstating the results due to non-independence). We then estimate the prediction model 10 times, excluding each randomly formed group once. Each observation in the excluded group is assigned a predicted value, using the coefficients estimated from the observations in the other nine groups. The cutoff scores for fraud prediction are calculated within sample and used to classify the observations in the hold-out sample. We repeat this process 20 times, for a total of 200 hold-out samples.

The results, shown in Panel D of Table 3, indicate that the predictive power of the models is only slightly lower in the hold-out samples. For example, the specification in the second column correctly predicts 150 fraud firm-years within sample, compared to an average of 143.3 fraud firm-years in the hold-out samples of the K-fold tests. The minimum number of fraud firm-years predicted across the 20 repetitions of the K-fold test is 135 and the maximum is 149, which suggests the model is quite stable.

The results of the out-of-sample and K-fold cross-validation tests support the robustness of the fraud predictions in Panel B. Note that these are robustness tests of the models' overall predictions, and do not provide evidence as to the robustness of the coefficients on individual variables. Overall, the results from the four panels of Table 3 show that the information investment advisers are required to disclose is relevant and useful for predicting fraud.

3.4. Initiation versus Continuance of Fraud

The model in column four of Table 3 tests whether fraud is predictable after excluding from the sample all firms that have a previously disclosed fraud case in the SEC administrative proceedings or litigation releases, but the dependent variable does not distinguish between the year in which the fraud is initiated and later years in which the fraud continues. In this section, we explore a related but distinct question: Can Form ADV data be used to predict the initiation of a new fraud? This question is important because predicting fraud at initiation minimizes the harm, and because initiating a new fraud and continuing a preexisting fraud are economically different decisions. For example, Dechow, Sloan, and Sweeney (1996) and Dechow, Ge, Larson, and Sloan (2010) show that accounting fraud is often initiated in response to firm performance. Thus, certain predictive variables might measure a time-varying factor that triggers the initiation of fraud, whereas other variables might measure a time-invariant propensity toward fraud.

Panel A of Table 4 shows the results of a multinomial probit regression. In the first column, the dependent variable equals one for firms that initiate a new fraud in the subsequent year. In the second column, the dependent variable equals one for firms that continue a preexisting fraud into the subsequent year. The excluded category is clean firm-years. The third column shows p-values from chi-square tests of the hypothesis that the coefficients are equal in both equations. All significance tests are based on standard errors clustered by firm. When interpreting the results, keep in mind that there is a potential sample selection effect. Easily detected fraud cases are less likely to persist to become continued fraud cases. As a result, easily detected fraud cases are likely overrepresented in the sample of initiated fraud cases, which could overstate the estimates of the predictability of fraud.

The last row in Panel A of Table 4 shows the p-value from a chi-square test of the joint hypothesis that the coefficients are equal in both equations. The test does not reject this hypothesis. Although the coefficients for Referral Fees and Broker in Firm are significantly higher in the initiation of fraud equation, the significance does not persist after adjusting for multiple comparisons. The fact that the coefficient estimates are not significantly different likely reflects the fact that the Form ADV variables are quite stable over time, and suggests that the predictive variables primarily measure a time-invariant component of the propensity for fraud.

Panel B of Table 4 shows the proportion of fraud firm-years that could be predicted within sample. At a false positive rate of 5%, the model predicts 37.9% of initiated fraud cases, compared with 25.8% of continued fraud firm-years. A chi-square test of classification accuracy shows that the model is significantly better (at the 5% level) at predicting newly initiated fraud cases. Thus, although the coefficient estimates are not significantly different between the two equations, the predictive accuracy is significantly higher for the initiation of fraud.

4. Annual Cross-Sectional Regressions

As noted earlier, the models presented in Table 3 use observations from the entire sample period. Although this approach allows for relatively powerful tests of the relation between the disclosed information and fraud it may obscure time effects, which could arise in several ways. First, there could be time effects in the actual rate of fraud due to changes in the legal or operating environment. Second, there could be time effects due to the detection rate. Recall that Table 1 shows that the median fraud in the sample persists for five and a half years. This suggests that the dependent variable for the 2006 cross-section likely includes less than half of the fraud cases that actually occurred in that year. By contrast, the 2001 cross-section likely includes a much higher proportion of the fraud cases that actually occurred in that year.

To examine whether there are time effects in the prediction of fraud, Panel A of Table 5 shows annual, cross-sectional probit regressions that predict investment fraud that occurs during the subsequent 12 months. For example, the model in column one uses Form ADV data available on August 1, 2001 to predict fraud that occurs from August 2001 through July 2002. Because fraud cases can persist for multiple years, these annual regressions are not independent, and aggregating coefficients across years could lead to faulty conclusions.

We test whether the coefficient estimates are significantly different across years with Wald tests. These tests are adjusted for non-independence, which could occur because the same firm can appear in multiple years. The coefficients for Dedicated CCO are significantly different across years at the 1% level, and the coefficients for Investment Company Act are significantly different at the 10% level. Both variables significantly predict fraud in the early years of the sample, but not the later years. This change is possibly related to the mutual fund late trading scandal that occurred in the early years of the sample. The firms involved in the late trading scandal managed funds that were registered under the Investment Company Act, and were mostly large financial conglomerates, which are more likely to have a dedicated CCO. Custody and Majority Employee Owned are also significantly different across years.

Note that because there are fewer observations in these annual cross-sectional regressions, the Wald tests have low power to reject the hypothesis that the coefficients are equal across years. For example, Referral Fees is significant in Table 3, but in Table 5 Referral Fees is significant in only two years. We cannot reject that that the coefficients are jointly equal to zero, nor can we reject that the annual coefficients are jointly equal to the full sample coefficient.

Panel B of Table 5 shows the ability of the cross-sectional regressions to predict fraud within sample at a 5% false positive rate for each year. The regressions predict a higher proportion of fraud in the first three years of the sample. The August 2005 cross-section has the worst performance, predicting only 18.6% of fraud cases. The results in Panel B, however, show that the annual cross-sectional regressions have significant power to predict fraud in each year of the sample, which suggests the results in Table 3 are not driven by a single period.

5. Data Access and Implementation

The predictive regressions in Table 3 use information from the full sample period, and so do not directly address how well an investor could have predicted fraud as of August 1 of each year during the sample (e.g., the fraud predictions in 2003 are based on coefficients estimated using data from 2001 through 2006). In this section, for each year in the sample we test how well investors could have implemented predictive models using only Form ADV data that had previously been publicly accessible. During the sample period, the SEC did not provide public access to historical Form ADV filings; investors could access only a contemporaneous cross-section. For this reason we compare two types of predictive models. In the first, we estimate predictive models that use only the contemporaneous cross-section of Form ADV filings. These tests mimic the predictions an investor could have made during the sample period, given the actual data access policies in place. In the second, we estimate predictive models that use data from an annual panel of historical Form ADV filings. These tests mimic the predictions an investor could have made if historical Form ADV filings had been publicly accessible.

Table 6 shows the results of fraud prediction models that use only the contemporaneously accessible cross-section of Form ADV filings as of August 1 of each year. To illustrate, in the column labeled Aug '05, the independent variables are taken from each firm's most recent Form ADV filing prior to August 1, 2005. The dependent variable equals one for all firms with an observable prior fraud case (i.e., a fraud case that occurred between September 28, 1995 and July 31, 2005, and which was identified in an SEC administrative action or litigation release filed before July 31, 2005). We use the coefficient estimates from this regression to make out-of-sample predictions of the fraud cases that occur between August 1, 2005 and July 31, 2006.

For comparison purposes, Table 7 shows the results of fraud prediction models that use an annual panel of historical Form ADV filings. Like Table 6, these regressions use only information that existed at the time of the prediction. But unlike Table 6, they use information that was not contemporaneously accessible by the public. To illustrate, in the column labeled Aug '05, the independent variables are taken from each firm's Form ADV filings as of August 1 in 2001, 2002, 2003, and 2004. For each August 1 firm-year observation, the dependent variable equals one if the firm commits fraud during the subsequent 12 months, and the fraud is publicly revealed before August 1, 2005. We combine the coefficient estimates from this model with each firm's Form ADV data as of August 1, 2005 to make out-of-sample predictions of the fraud cases that occur between August 1, 2005 and August 1, 2006.

The models presented in Tables 6 and 7 differ in several ways. Most obviously, the panel models in Table 7 use more data to estimate fraud risk than do the cross-sectional models in Table 6. More important, the models in Table 6 are backward looking: they show the relation between contemporaneous variables and *past* fraud. If the contemporaneous Form ADV filings are all that is publicly accessible, then investors can only estimate fraud risk from backward-looking regressions [e.g., Brown, Goetzmann, Liang, and Schwarz (2008)]. If historical filings are accessible, investors can estimate forward-looking prediction models and then estimate fraud risk by combining the estimated coefficients with the contemporaneous disclosures of the firms. This is a conceptually important distinction. The backward-looking regressions only include the subsample of firms that survived the legal consequences of committing fraud. The forward-looking models in Table 7 have one disadvantage, however: These models require at least two years of data to estimate, and so it is not possible to estimate this model for the year beginning August 1, 2001.

For the regressions reported in Tables 6 and 7 our main interest is not the coefficient estimates, but rather the comparison of the models' ability to correctly predict fraud. We compare the prediction of fraud cases occurring between August 1, 2002 and August 1, 2007. During this period there were a total of 413 fraud firm-years. At a false positive rate of 5%, the cross-sectional regressions, shown in Table 6, predict 24.7% of the fraud cases (the fraud firm-years predicted between August 2002 and August 2007 sum to 102). The panel regressions using all prior years, shown in Table 7, predict 31.2% of the fraud firm-years (a sum of 129 fraud firm-years). A chi-square test of classification accuracy shows that the panel regressions in Table 7 predict a significantly larger number of fraud cases (p-value< 0.01). Although the absolute difference in predictive accuracy is only 6.5 percentage points (an improvement of 26.3% relative to the model in Table 6), these tests provide evidence that public access to historical Form ADV filings could benefit investors. Moreover, the marginal cost to the SEC of allowing public access to historical filings would be quite low.

To implement a fraud prediction model, such as those tested in this paper, an investor would have had to collect manually a large number of Form ADV filings, convert the filings into a database, and estimate prediction models. For most investors, the cost of doing so may well have exceeded the perceived benefits. This problem is exacerbated by the fact that investors are atomistic: Even if the aggregate benefit of processing the disclosed information is greater than the cost to a single investor, the benefit to any single investor may be insufficient. As shown by Becker (1968), the socially optimal level of a crime occurs when the marginal benefit from a further reduction in the crime is equal to the marginal cost of increased enforcement. Allowing public access to historical Form ADV filings would reduce the marginal cost of increased enforcement by facilitating investors' use of these data. This, in turn, should reduce the marginal benefit to an investment adviser of committing fraud due to an increase in the probability of detection. Thus improved public access to these disclosure data should reduce the occurrence of fraud.

6. Conclusion

This paper finds that required disclosures related to past regulatory and legal violations, conflicts of interest, and monitoring are significant predictors of investment fraud. We stress, however, that these are tests of prediction and do not imply a causal relation between the disclosed information and fraud. If during the period August 2001 through August 2007 investors had avoided the 5% of firms with the highest ex ante predicted fraud risk, they could have avoided more than $4 billion in losses from fraud. Based on the SEC's estimate of 9.01 hours to fill out Form ADV and an assumed cost of $1,000 per hour, during this same period, the direct costs of disclosure were at most $500 million. Thus, even ignoring the deterrent effect of Form ADV, this simple, back-of-the-envelope calculation suggests that the benefits of Form ADV substantially outweigh the costs. However, the investing public had a potentially limited ability to develop and use predictive models based on Form ADV data because the SEC did not provide access to historical data. As a result, the realized benefits of disclosure during the period may have been lower. The results suggest that improving public access to comprehensive historical disclosures could increase the benefits they were meant to provide.

Table 1: Summary of Investment Fraud

This table summarizes cases of investment fraud committed by investment advisors between August 2001 and July 2010 as reported in SEC administrative actions and litigation releases. Registered denotes firms that file a Form ADV with the SEC. Firm-wide fraud is committed by high level executives, or at the very least, with the firms' implicit acceptance. Rogue employee fraud is committed by individuals who evade their firms' internal control systems and the firms do not knowingly benefit.

Panel A: Registered vs. Non-Registered Advisors

	Total	Firm-Wide	Rogue Employee	Investor Losses ($ billion)
Non-Registered	251	244	7	4.5
Registered	258	217	41	32.4
Total	509	461	48	36.9

Panel B: Fraud Characteristics

		Amount ($ million)				Duration (years)		
	Obs.	Mean	Median	Max	Missing Obs.	Mean	Median	Max
Firm-Wide	217	196.3	6.0	18,000.0	56	6.0	5.6	23.9
Rogue Employee	41	25.4	3.0	300.0	8	5.4	5.4	12.2
Total	258	167.2	5.1	18,000.0	64	5.9	5.5	23.9

Table 2: Summary of Investment Advisory Firms

This table summarizes information from each firm's first Form ADV filing during the period August 2001 through July 2006. There are 13,853 unique firms in the sample. Employee Ownership is the aggregate employee ownership of the firm. Percent Client Agents is the percentage of clients that are agents for the owners of the assets. Past Fraud equals one if the firm has been publicly accused of fraud. Past Affiliated Fraud equals one if the firm's affiliates have been accused of fraud. Past Regulatory equals one if the firm reports past regulatory violations. Past Civil or Criminal equals one if the firm reports past civil or criminal violations. Referral Fees equals one if the firm compensates any party for client referrals. Interest in Transactions equals one if the firm: recommends securities in which it has an ownership interest, serves as an underwriter, or has any other sales interest. Soft Dollars equals one if the firm receives benefits other than execution from a broker-dealer in connection with clients' trades. Broker in Firm equals one if the firm employs registered representatives of a broker-dealer. Investment Company Act equals one if the firm is registered under the Investment Company Act of 1940. Custody equals one if the firm has custody of clients' cash or securities. Dedicated CCO equals one if the chief compliance officer has no other job title. Hedge Fund Clients equals one if more than 75% of the firm's clients are hedge funds. The column Clean (Fraud) summarizes firms in which a fraud is not committed from first filing through July 2007 (is committed). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels based on Fisher's exact test.

Panel A: Firm Characteristics

	Mean	SD	25^{th}	50^{th}	75^{th}
Employee Ownership	68.2%	44.2	0.0	100.0	100.0
Avg. Acct. Size ($ thousand)	55,361	328,522	339	1,442	21,667
Percent Client Agents	23.2%	32.6	0.0	8.3	30.0
Assets Under Mgmt. ($ million)	2,213	16,433	37	90	400
Firm Age (years)	5.1	7.7	0.4	1.1	8.1

Panel B: Firm Disclosures

	All	Clean	Fraud
Past Fraud	0.2%	0.2	1.6***
Past Affiliated Fraud	1.6%	1.6	2.6
Past Regulatory	12.1%	11.9	32.6***
Past Civil or Criminal	3.3%	3.1	12.5***
Referral Fees	40.0%	39.7	59.8***
Interest in Transaction	30.4%	30.1	52.2***
Soft Dollars	55.7%	55.6	63.0**
Broker in Firm	40.8%	40.4	66.3***
Investment Company Act	9.8%	9.6	29.0***
Custody	23.9%	23.7	33.7***
Dedicated CCO	10.7%	10.7	12.4
Hedge Fund Clients	13.4%	13.5	6.2***

Table 3: Predicting Fraud

The full sample consists of 53,994 firm-year observations. The independent variables are taken from each firm's Form ADV filing as of August 1 each year from 2001 through 2006. In the first column, the sample includes only the first Form ADV filed during the sample period. In the second and third columns, the full sample is included. In the fourth column, the sample excludes firms with a previously disclosed fraud. In the fifth column, the sample excludes all firms that disclose any type of prior legal or regulatory violation, either by the firm itself or an affliated firm, in Item 11 of Form ADV. Column one of Panel A shows the results of a cross-sectional probit regression predicting fraud. The dependent variable equals one if the firm commits fraud in any subsequent year through July 2007. Standard errors are robust. Columns two through five show the results of pooled probit regressions predicting fraud. The dependent variable equals one if the firm commits fraud in the subsequent year. Standard errors are clustered by firm and year. In the interest of brevity the constants are not reported. Z-scores are reported in square brackets. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The columns in Panels B, C, and D correspond to the columns in Panel A. Panel B shows the proportion of fraud that could be predicted within sample. Panel C shows the out-of-sample performance of each model, using Form ADV filings in August 2006 to predict fraud cases that occur between August 2007 through July 2010. Panel D shows the results from K-fold cross-validation tests.

Panel A: Predictors of Fraud					
	Cross Section	Full Sample	Full Sample	No Prior	No Violations
Past Fraud	0.329		0.272		
	[0.98]		[1.46]		
Past Affiliated Fraud	-0.224		-0.184	-0.196*	
	[1.14]		[1.54]	[1.68]	
Past Regulatory	0.175**	0.284***	0.282***	0.285***	
	[2.25]	[4.20]	[4.16]	[4.15]	
Past Civil or Criminal	0.223*	0.191**	0.200**	0.209**	
	[1.93]	[2.13]	[2.29]	[2.32]	
Referral Fees	0.135**	0.100*	0.099*	0.099*	0.139**
	[2.09]	[1.79]	[1.79]	[1.78]	[2.40]
Interest in Transaction	0.138*	0.197***	0.198***	0.196***	0.184**
	[1.93]	[2.89]	[2.91]	[2.86]	[2.24]
Soft Dollars	-0.029	-0.051	-0.046	-0.040	-0.073
	[0.46]	[0.89]	[0.81]	[0.71]	[1.10]
Broker in Firm	0.237***	0.118**	0.120**	0.120**	0.096
	[3.89]	[2.01]	[2.05]	[2.02]	[1.55]
Investment Company Act	0.103	0.263***	0.269***	0.278***	0.273***
	[1.39]	[3.29]	[3.36]	[3.58]	[2.83]
Custody	0.309***	0.094	0.097	0.088	0.028
	[3.79]	[1.43]	[1.50]	[1.36]	[0.33]
Dedicated CCO	0.288***	-0.088	-0.085	-0.085	-0.056
	[2.67]	[0.86]	[0.82]	[0.82]	[0.53]
Majority Employee Owned	0.045	0.009	0.001	0.008	0.033
	[0.65]	[0.11]	[0.02]	[0.10]	[0.37]
Log (Avg. Acct. Size)	-0.043***	-0.072***	-0.070***	-0.065***	-0.028
	[3.45]	[4.25]	[4.12]	[3.68]	[1.12]
Percent Client Agents	0.001	0.003***	0.003***	0.003***	0.003***
	[1.40]	[3.91]	[3.88]	[3.77]	[2.91]
Hedge Fund Clients	-0.035	0.031	0.031	0.020	0.030
	[0.27]	[0.27]	[0.27]	[0.18]	[0.22]
Log (AUM)	0.036***	0.060***	0.059***	0.054***	0.020
	[3.76]	[4.10]	[3.98]	[3.57]	[0.93]
Log (Firm Age)	0.014	0.002	0.002	0.002	0.008
	[1.18]	[0.20]	[0.19]	[0.20]	[0.66]
Model Chi-Square	175.2***	181.5***	198.9***	176.9***	63.2***
Observations	13,853	53,994	53,994	53,750	45,920

Panel B: Within Sample Predictions					
# Fraud	193	517	517	501	310
Fraud Predicted	59	150	152	140	44
	30.6%	29.0	29.4	27.9	14.2
# Clean Firms	13,660	53,477	53,477	53,249	45,610
Clean Firm False Positives	683	2,673	2,673	2,662	2,280
	5.0%	5.0	5.0	5.0	5.0
Prop. of $ Losses Avoided	37.4%	41.3	43.0	40.5	7.9

Panel C: Out-of-Sample Predictions (August 2007 - July 2010)					
	Cross Section	Full Sample	Full Sample	No Prior	No Violations
# Fraud	27	27	27	25	18
Fraud Predicted	9	10	9	7	1
	33.3%	37.0	33.3	28.0	5.6
Fraud Not Predicted	18	17	18	18	17
	66.7%	63.0	66.7	72.0	94.4
# Clean Firms	10,356	10,356	10,356	10,293	8,912
Clean Firms Not Accused	9,839	9,839	9,839	9,779	8,467
	95.0%	95.0	95.0	95.0	95.0
Clean Firm False Positives	517	517	517	514	445
	5.0%	5.0	5.0	5.0	5.0

Panel D: K-Fold Cross-Validation Hold Out Sample Predictions (August 2001 - July 2007)					
Avg # Fraud Predicted	53.6	143.3	142.4	129.7	35.0
Avg % Fraud Predicted	27.8%	27.7	27.5	25.9	11.3
Stdev Fraud Predicted (#)	1.39	3.64	3.75	4.32	2.66
Min # Fraud Predicted	51	135	133	120	32
Max # Fraud Predicted	56	149	148	137	42
Avg # False Positives	678.4	2,669.2	2,669.2	2,658.2	2,275.8
Avg % False Positives	5.0%	5.0	5.0	5.0	5.0
Stdev False Positives	0.68	0.95	0.95	0.99	0.91
Min # False Positives	677	2,668	2,668	2,656	2,274
Max # False Positives	679	2,671	2,671	2,660	2,277

Table 4: Initiation vs. Continuation

The sample consists of 53,994 firm-year observations. The independent variables are taken from each firm's Form ADV filings as of August 1 each year from 2001 through 2006. Panel A shows the results of a multinomial probit regression predicting fraud. In the first column, the dependent variable equals one for firms that initiate a new fraud in the subsequent year. In the second column, the dependent variable equals one for firms that continue a preexisting fraud in the subsequent year. The excluded category is clean firms. The third column shows p-values from chi-square tests of the null hypothesis that the estimated coefficients are equal in both equations. In the interest of brevity the constants are not reported. All significance tests are based on standard errors clustered by firm. Z-scores are reported in square brackets. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The columns in Panel B correspond to the columns in Panel A. Panel B shows the proportion of fraud that could be predicted within sample.

Panel A: Predicting Initiation versus Continuance of Fraud			
	Initiate	Continue	p-value Difference
Past Fraud	0.489	0.551	
	[0.86]	[1.24]	0.917
Past Affiliated Fraud	-0.684	-0.364	
	[1.30]	[1.13]	0.508
Past Regulatory	0.673***	0.714***	
	[2.67]	[3.63]	0.874
Past Civil or Criminal	0.633**	0.407*	
	[2.11]	[1.81]	0.459
Referral Fees	0.897***	0.117	
	[3.45]	[0.70]	0.002***
Interest in Transaction	0.446	0.522**	
	[1.63]	[2.48]	0.786
Soft Dollars	-0.346	-0.077	
	[1.43]	[0.41]	0.246
Broker in Firm	0.774***	0.270	
	[3.03]	[1.52]	0.037**
Investment Company Act	0.545**	0.627***	
	[1.98]	[2.87]	0.782
Custody	-0.003	0.295	
	[0.01]	[1.48]	0.282
Dedicated CCO	-0.328	-0.172	
	[1.15]	[0.93]	0.583
Majority Employee Owned	0.057	-0.039	
	[0.23]	[0.20]	0.721
Log (Avg. Acct. Size)	-0.217***	-0.173***	
	[3.83]	[4.07]	0.459
Percent Client Agents	0.008**	0.009***	
	[2.14]	[3.37]	0.964
Hedge Fund Clients	0.245	-0.008	
	[0.44]	[0.02]	0.670
Log (AUM)	0.198***	0.146***	
	[4.37]	[4.14]	0.273
Log (Firm Age)	-0.062	0.031	
	[1.27]	[0.68]	0.119
Overall Model p-value Difference			0.143

Panel B: Within Sample Predictions			
# Fraud	87	430	
Fraud Predicted	33	111	
	37.9%	25.8	
# Clean Firms	53,907	53,564	
Clean Firm False Positives	2,673	2,673	
	5.0%	5.0	

Table 5: Annual Cross-Sectional Regressions

The sample consists of 53,994 firm-year observations. Each column contains an annual cross-sectional regression, in which the independent variables are taken from each firm's Form ADV filing as of August 1 each year from 2001 through 2006. Panel A shows the results of annual cross-sectional probit regressions predicting fraud. The dependent variable equals one if the firm commits fraud in the subsequent year. Each column reports coefficients from an annual cross-section. In the interest of brevity we do not report coefficients for the constants. Standard errors are robust. Z-scores are reported in square brackets. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The columns in Panel B correspond to the columns in Panel A. Panel B shows the proportion of fraud that could be predicted within sample.

Panel A: Predictors of Fraud						
	Aug '01	Aug '02	Aug '03	Aug '04	Aug '05	Aug '06
Past Fraud	0.280	0.172	-0.337	0.451	0.625**	0.553*
	[0.82]	[0.52]	[0.78]	[1.49]	[2.25]	[1.81]
Past Affiliated Fraud	-0.321	-0.241	-0.194	-0.010	0.201	-0.051
	[1.23]	[1.15]	[1.11]	[0.05]	[0.94]	[0.19]
Past Regulatory	0.187*	0.235**	0.350***	0.371***	0.248**	0.050
	[1.89]	[2.42]	[3.79]	[3.44]	[2.01]	[0.37]
Past Civil or Criminal	0.239*	0.212*	0.092	0.328**	-0.095	0.342*
	[1.73]	[1.66]	[0.71]	[2.41]	[0.47]	[1.76]
Referral Fees	0.041	0.021	0.071	0.130	0.187*	0.255**
	[0.45]	[0.26]	[0.85]	[1.51]	[1.90]	[2.25]
Interest in Transaction	0.265***	0.252***	0.235**	0.138	0.124	-0.029
	[2.78]	[2.63]	[2.48]	[1.23]	[1.03]	[0.23]
Soft Dollars	-0.037	-0.067	-0.015	-0.020	-0.075	-0.190
	[0.40]	[0.78]	[0.17]	[0.21]	[0.74]	[1.53]
Broker in Firm	0.202**	0.127	0.071	0.043	0.006	0.147
	[2.33]	[1.51]	[0.84]	[0.45]	[0.05]	[1.23]
Investment Company Act	0.245**	0.325***	0.306***	0.264**	-0.120	-0.081
	[2.43]	[3.25]	[2.89]	[2.26]	[0.73]	[0.42]
Custody	0.006	0.084	0.166*	0.061	0.246**	0.339***
	[0.06]	[0.89]	[1.76]	[0.58]	[2.37]	[2.65]
Dedicated CCO	0.247	0.348**	0.420***	-0.083	0.057	-0.126
	[1.53]	[2.54]	[3.41]	[0.59]	[0.58]	[1.08]
Majority Employee Owned	-0.089	-0.110	0.026	0.220**	0.252**	-0.022
	[0.88]	[1.14]	[0.28]	[2.26]	[2.23]	[0.17]
Log (Avg. Acct. Size)	-0.100***	-0.089***	-0.076***	-0.056**	-0.034	-0.061
	[3.90]	[4.18]	[3.29]	[2.14]	[1.02]	[1.61]
Percent Client Agents	0.004***	0.003**	0.004***	0.002*	0.003*	0.001
	[3.03]	[2.57]	[2.98]	[1.74]	[1.72]	[0.50]
Hedge Fund Clients	0.006	0.098	0.150	0.127	-0.091	-0.070
	[0.02]	[0.46]	[0.79]	[0.61]	[0.44]	[0.28]
Log (AUM)	0.091***	0.080***	0.062***	0.041*	0.028	0.046
	[4.24]	[4.49]	[3.21]	[1.94]	[1.08]	[1.62]
Log (Firm Age)	0.019	0.007	-0.003	0.009	0.017	0.000
	[1.00]	[0.24]	[0.13]	[0.36]	[0.72]	[0.00]
Model Chi-Square	138.0***	138.7***	135.3***	78.1***	44.6***	63.9***
Observations	7,352	7,747	8,562	9,088	10,862	10,383
Panel B: Within Sample Predictions						
# Fraud	104	116	115	83	59	40
Fraud Predicted	39	45	37	22	11	10
	37.5%	38.8	32.2	26.5	18.6	25.0
Clean Firms	7,248	7,631	8,447	9,005	10,803	10,343
Clean Firm False Positives	362	381	422	450	540	517
	5.0%	5.0	5.0	5.0	5.0	5.0

Table 6: Point-in-Time Tests Using Publicly Accessible Data

Each column uses only the contemporaneously accessible cross-section of Form ADV filings as of August 1 of each year. Panel A shows the estimates from cross-sectional probit regressions. The dependent variable equals one for firms which have a publicly observed prior history of fraud (fraud occurring and detected between January 1, 1996 and August 1 of the year in which the independent variables are observed). The independent variables reflect the publicly accessible data as of August 1 of the year in the column. In the interest of brevity we do not report coefficients for the constants. Standard errors are robust. Z-scores are reported in square brackets. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The columns in Panel B correspond to the columns in Panel A. Panel B reports an out-of-sample test, in which the model is used to predict fraud that occurs in the subsequent year.

	Aug '01	Aug '02	Aug '03	Aug '04	Aug '05	Aug '06
Panel A: Point-in-Time Cross-Sections						
Past Affiliated Fraud	0.757***	0.774***	0.111	-0.006	-0.067	-0.188
	[2.70]	[3.14]	[0.39]	[0.03]	[0.30]	[0.87]
Past Regulatory	0.705***	0.976***	0.984***	0.820***	0.953***	1.166***
	[3.99]	[4.66]	[6.56]	[5.24]	[7.40]	[8.69]
Past Civil or Criminal	0.642***	0.445**	0.594***	0.620***	0.409**	0.464***
	[2.85]	[2.28]	[3.44]	[3.59]	[2.53]	[3.50]
Referral Fees	0.437**	0.347*	0.333**	0.234	0.137	0.127
	[2.17]	[1.91]	[2.14]	[1.64]	[1.15]	[1.07]
Interest in Transaction	0.366	0.447**	0.019	-0.251	0.063	0.081
	[1.56]	[2.10]	[0.11]	[1.39]	[0.43]	[0.65]
Soft Dollars	-0.401**	-0.311*	-0.210	-0.342**	-0.102	0.017
	[2.14]	[1.91]	[1.32]	[2.26]	[0.85]	[0.14]
Broker in Firm	0.093	0.045	-0.090	-0.209*	-0.157	-0.133
	[0.54]	[0.25]	[0.69]	[1.71]	[1.26]	[1.05]
Investment Company Act	0.206	0.155	-0.088	0.284*	0.118	0.143
	[0.88]	[0.71]	[0.43]	[1.65]	[0.66]	[0.87]
Custody	-0.092	-0.019	-0.039	0.165	-0.067	-0.007
	[0.48]	[0.11]	[0.23]	[1.11]	[0.55]	[0.06]
Dedicated CCO	-0.321	-0.178	-0.035	0.174	0.089	0.098
	[0.86]	[0.78]	[0.15]	[1.05]	[0.78]	[0.92]
Majority Employee Owned	-0.114	0.102	0.011	-0.056	0.083	0.148
	[0.51]	[0.55]	[0.06]	[0.38]	[0.67]	[1.16]
Log (Avg. Acct. Size)	-0.161***	-0.114***	-0.044	-0.062*	-0.069**	-0.067**
	[3.52]	[3.09]	[1.09]	[1.78]	[2.10]	[2.13]
Percent Client Agents	0.000	-0.001	-0.001	0.002	0.001	0.000
	[0.12]	[0.22]	[0.26]	[0.86]	[0.28]	[0.06]
Hedge Fund Clients	0.384	0.074	-0.079	-0.122		-0.421
	[0.93]	[0.20]	[0.20]	[0.32]		[1.08]
Log (AUM)	0.121***	0.095***	0.037	0.057*	0.057**	0.048*
	[3.39]	[2.86]	[1.20]	[1.92]	[2.07]	[1.85]
Log (Firm Age)	0.056*	0.064	0.183***	0.160***	0.124**	0.197***
	[1.70]	[1.30]	[3.29]	[2.79]	[2.20]	[3.99]
Model Chi-Square	224.8***	116.2***	115.1***	150.1***	139.4***	155.7***
Observations	7,352	7,747	8,562	9,088	9,110	10,383
Panel B: Out-of-Sample Predictions						
# Fraud	104	116	115	83	59	40
Fraud Predicted	28	33	27	20	12	10
	26.9%	28.4	23.5	24.1	20.3	25.0
# Clean Firms	7,248	7,631	8,447	9,005	10,803	10,343
Clean Firm False Positives	360	374	419	456	470	552
	5.0%	4.9	5.0	5.1	4.4	5.3

Table 7: Predictions Using a Panel of All Prior Years

Each column represents the predictions an investor could have made at a specified point-in-time had historical Form ADV been publicly available. For each column, the sample consists of a panel of all previously available annual Form ADV filings as of August 1 of each year. (E.g. In Aug 2002, the independent variables are taken from the Aug 2001 cross-section of Form ADV filings. In Aug 2003, the independent variables are taken from the Aug 2002 and Aug 2001 samples of Form ADV filings.) Panel A shows the results of fraud prediction models that use all prior Form ADV filings to predict fraud For each firm-year observation the dependent variable equals one if the firm commits a fraud during the subsequent 12 months. In the interest of brevity we do not report coefficients for the constants. Standard errors are clustered by firm and year. Z-scores are reported in square brackets. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Panel B shows the proportion of fraud that could be predicted in year T+1, based on a prediction model formed at time T using all data up to time T-1.

Panel A: Panel of All Prior Years					
Past Fraud	0.280	0.226	0.102	0.178	0.254
	[0.82]	[0.94]	[0.37]	[0.79]	[1.21]
Past Affiliated Fraud	-0.321	-0.271*	-0.237*	-0.219*	-0.187
	[1.23]	[1.83]	[1.91]	[1.89]	[1.52]
Past Regulatory	0.187*	0.213***	0.261***	0.288***	0.293***
	[1.89]	[3.53]	[3.42]	[3.82]	[4.15]
Past Civil or Criminal	0.239*	0.220***	0.176*	0.211**	0.181*
	[1.73]	[2.61]	[1.80]	[2.25]	[1.96]
Referral Fees	0.041	0.031	0.045	0.061	0.084
	[0.45]	[0.67]	[0.83]	[1.15]	[1.53]
Interest in Transaction	0.265***	0.257***	0.250***	0.227***	0.216***
	[2.78]	[4.92]	[4.27]	[3.42]	[3.29]
Soft Dollars	-0.037	-0.053	-0.041	-0.034	-0.033
	[0.40]	[1.04]	[0.74]	[0.61]	[0.61]
Broker in Firm	0.202**	0.162***	0.127**	0.117*	0.111*
	[2.33]	[2.86]	[2.04]	[1.93]	[1.89]
Investment Company Act	0.245**	0.289***	0.301***	0.304***	0.276***
	[2.43]	[4.17]	[4.60]	[4.68]	[3.59]
Custody	0.006	0.048	0.094	0.075	0.088
	[0.06]	[0.72]	[1.20]	[1.07]	[1.37]
Dedicated CCO	0.247	0.305***	0.346***	0.199	0.027
	[1.53]	[3.10]	[3.46]	[1.29]	[0.22]
Majority Employee Owned	-0.089	-0.099*	-0.053	-0.005	0.016
	[0.88]	[1.90]	[0.71]	[0.06]	[0.19]
Log (Avg. Acct. Size)	-0.100***	-0.094***	-0.087***	-0.079***	-0.072***
	[3.90]	[7.75]	[5.85]	[4.68]	[4.01]
Percent Client Agents	0.004***	0.004***	0.004***	0.003***	0.003***
	[3.03]	[4.18]	[4.29]	[3.90]	[3.94]
Hedge Fund Clients	0.006	0.072	0.114	0.107	0.053
	[0.02]	[0.52]	[0.87]	[0.83]	[0.42]
Log (AUM)	0.091***	0.084***	0.076***	0.067***	0.061***
	[4.24]	[7.72]	[5.46]	[4.39]	[3.89]
Log (Firm Age)	0.019	0.015	0.008	0.008	0.007
	[1.00]	[1.44]	[0.68]	[0.71]	[0.69]
Model Chi-Square	138.0***	543.8***	304.8***	218.7***	205.5***
Observations	7,352	15,099	23,661	32,749	43,611

Panel B: Out-of-Sample Predictions (Year T+1)					
# Fraud	116	115	83	59	40
Fraud Predicted	48	38	21	12	10
	41.4%	33.0	25.3	20.3	25.0
# Clean Firms	7,631	8,447	9,005	10,803	10,343
Clean Firm False Positives	386	402	496	603	517
	5.1%	4.8	5.5	5.6	5.0

Figure 1: One Fraud Case's Timeline

This figure shows the timeline of one particular fraud, committed by K.W. Brown & Company, from initiation to the end of all legal actions. Beginning in September 2002 the firm began defrauding clients through self-dealing. The firm traded securities for its own proprietary account as well as on behalf of clients. The firm engaged in ex post allocation of trades; securities were purchased but not allocated to a specific account. At a later date, profitable trades were retroactively allocated to the firm's proprietary account and unprofitable trades were allocated to clients. This resulted in over $4.5 million in illegal gains for the firm, and more than $9 million in client losses.

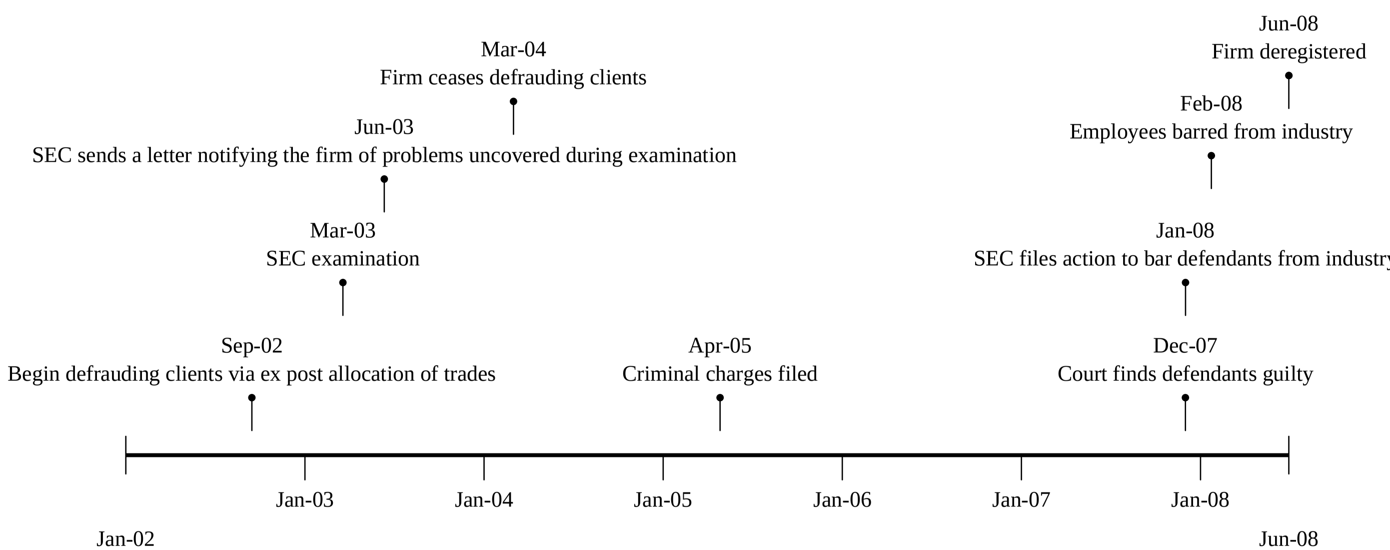
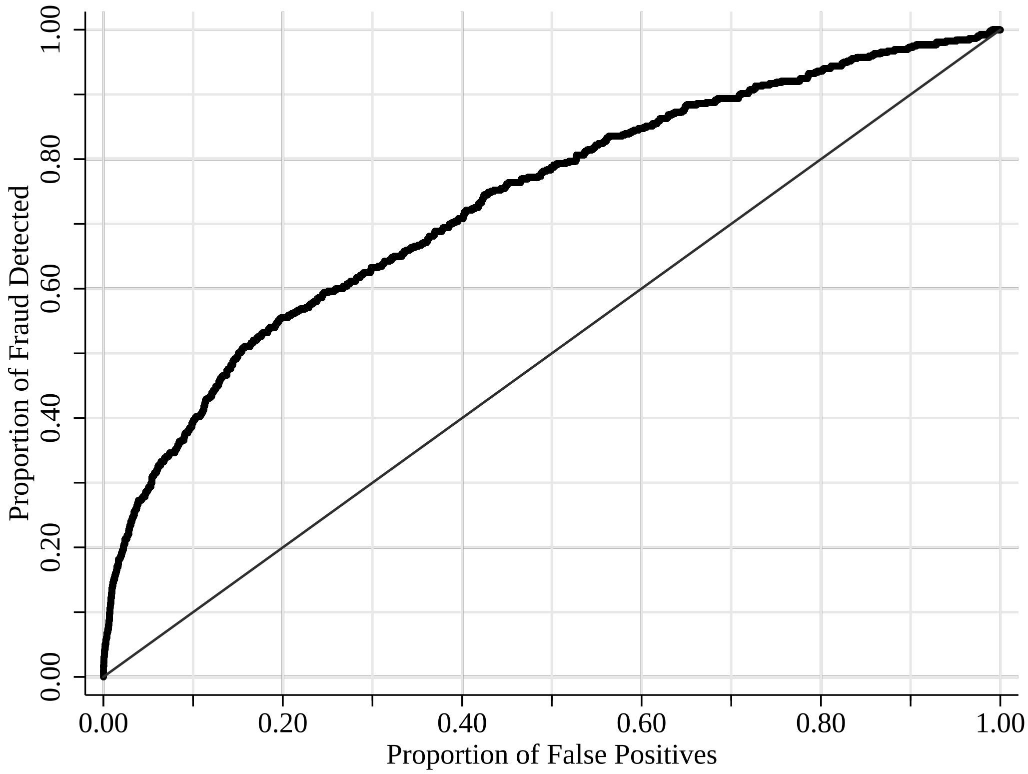


Figure 2: Proportion of Fraud Predicted for All False Positive Rates

This figure shows the receiver operating characteristic (ROC) curve for the probit regression results from the second column of Table 4. The ROC curve shows the relation between the proportion of fraud detected and the proportion of false positives for all possible classification cut-points. The ROC curve is generated by taking each observation's estimated fraud probability, computing the sensitivity and false positives using that point as a cut-point, and then plotting the results.



Appendix A: Variable Definitions

Variable	Definition	Data Source
Past Fraud	The firm committed a publicly observed fraud	SEC administrative proceeding or litigation release was filed for firm prior to August 1 of firm-year observation.
Past Affiliated Fraud	An affiliate of the firm committed a publicly observed fraud	SEC administrative proceeding or litigation release was filed for affiliated firm prior to August 1 of firm-year observation and Schedule D Section 7.A reports fraud firm as affiliate.
Past Regulatory	Filed a regulatory disclosure reporting page (DRP)	One of more of: Item 11c1-3, 11d1-5, 11e-4
Past Civil or Criminal	Filed a criminal or civil DRP	One of more of: Item 11a1-2, 11b1-2, 11h1a, 11h1b, 11h1c, 11h2
Referral Fees	Do you or any related person, directly or indirectly, compensate any person for client referrals?	Item 8f
Interest in Transaction	Do you or any related person: buy (or sell) securities from advisory clients; recommend securities in which you have an ownership interest or serve as underwriter, general or managing partner or have any other sales interest	One of more of: Item 8a1, 8a3, 8b2, 8b3
Soft Dollars	Do you or any related person receive research or benefits other than execution from a broker-dealer or a third party in connection with client securities transactions?	Item 8e
Broker in Firm	Employs registered representatives of a broker-dealer	Item 5b2>0
Investment Company Act	Investment adviser (or sub-adviser) to an investment company registered under the Investment Company Act	Item 2a4
Custody	Do you or any related person have custody of any advisory clients' cash or securities?	One of more of: Item 9a1-2, 9b1-2
Dedicated CCO	CCO has no other stated role within firm	CCO on Schedule A has no other "Title or Status"
Majority Employee Owned	Over 50% aggregate employee ownership	Imputed using Dimmock, Gerken, and Marietta-Westberg (2011) method
Log (Avg. Acct. Size)	Logarithm of assets under management per client	Log (Item 5f2c/(Item 5f2f+1)+1)
Percent Client Agents	Percent of banking, mutual, pension, charitable, corporate, and government clients	Sum of items: 5d3, 5d4, 5d5, 5d7, 5d8, 5d9 imputed using Dimmock, Gerken, and Marietta-Westberg (2011) method
Hedge Fund Clients	Primarily hedge fund clients	Item 5d6 $\geq$ 75%
Log (AUM)	Logarithm of assets under management	Log (Item 5f2c+1)
Log (Firm Age)	Logarithm of firm age in years	Log (years since date registered with the SEC)

Appendix B: Length of Fraud

This table presents Tobit regression estimates. The dependent variable is the logarithm of the length of the fraud in years. The full sample includes one observation per fraud with sufficient data to calculate duration.

	Full Sample	Full Sample	No Prior	No Violations
Past Fraud		0.076		
		[0.26]		
Past Affiliated Fraud		-0.087	-0.186	
		[0.36]	[0.73]	
Past Regulatory	0.026	0.029	0.031	
	[0.19]	[0.21]	[0.23]	
Past Civil or Criminal	0.065	0.071	0.087	
	[0.38]	[0.40]	[0.47]	
Referral Fees	-0.155	-0.155	-0.161	-0.179
	[1.24]	[1.24]	[1.27]	[1.08]
Interest in Transaction	-0.087	-0.085	-0.064	-0.155
	[0.74]	[0.72]	[0.53]	[1.03]
Soft Dollars	0.064	0.065	0.058	0.086
	[0.55]	[0.56]	[0.48]	[0.56]
Broker in Firm	-0.112	-0.112	-0.117	-0.108
	[0.90]	[0.90]	[0.93]	[0.72]
Investment Company Act	0.036	0.039	0.021	0.008
	[0.23]	[0.25]	[0.13]	[0.04]
Custody	0.147	0.148	0.147	0.231
	[1.26]	[1.26]	[1.24]	[1.42]
Dedicated CCO	0.060	0.067	0.089	0.061
	[0.47]	[0.52]	[0.66]	[0.33]
Majority Employee Owned	0.171	0.172	0.170	0.081
	[1.30]	[1.30]	[1.28]	[0.49]
Log (Avg. Acct. Size)	0.007	0.008	0.007	-0.033
	[0.22]	[0.26]	[0.21]	[0.63]
Percent Client Agents	-0.001	-0.001	-0.001	-0.001
	[0.45]	[0.46]	[0.42]	[0.29]
Hedge Fund Clients	-0.289	-0.294	-0.295	-0.244
	[1.14]	[1.16]	[1.15]	[0.76]
Log (AUM)	-0.002	-0.003	-0.002	0.028
	[0.09]	[0.12]	[0.06]	[0.65]
Log (Firm Age)	0.070**	0.070**	0.070**	0.095**
	[2.26]	[2.26]	[2.24]	[2.49]
	1.214***	1.208***	1.202***	1.281***
	[5.29]	[5.24]	[5.14]	[4.79]
R^2	0.042	0.043	0.043	0.062
Observations	182	182	176	116

References

Becker, Gary S., 1968, Crime and Punishment: An Economic Approach, *The Journal of Political Economy* 76, 169–217.

Bollen, Nicolas P.B., and Veronika K. Pool, 2008, Conditional Return Smoothing in the Hedge Fund Industry, *Journal of Financial and Quantitative Analysis* 43, 267–298.

Bollen, Nicolas P.B., and Veronika K. Pool, 2009, Do Hedge Fund Managers Misreport Returns? Evidence from the Pooled Distribution., *Journal of Finance* 64, 2257–2288.

Bollen, Nicolas P.B., and Veronika K. Pool, 2010, Predicting Hedge Fund Fraud with Performance Flags, . Working Paper, Vanderbilt University.

Brown, Stephen J., William N. Goetzmann, Bing Liang, and Christopher Schwarz, 2008, Mandatory Disclosure and Operational Risk: Evidence from Hedge Fund Registration, *Journal of Finance* 63, 2785–2815.

Brown, Stephen J., William N. Goetzmann, Bing Liang, and Christopher Schwarz, 2009, Estimating Operational Risk for Hedge Funds: The ω-Score, *Financial Analysts Journal* 65, 43–53.

Dechow, Patricia, Richard Sloan, and Amy Sweeney, 1996, Causes and Consequences of Earnings Manipulation: An Analysis of Firms Subject to Enforcement Actions by the SEC, *Contemporary Accounting Research* pp. 1–36.

Dechow, Patricia M., Weili Ge, Chad R. Larson, and Richard G Sloan, 2010, Predicting Material Accounting Misstatements, *Contemporary Accounting Research* forthcoming.

Dimmock, Stephen G., William C. Gerken, and Jennifer Marietta-Westberg, 2011, What Determines the Allocation of Managerial Ownership within Firms?, . Working Paper, Nanyang Technological University.

Straumann, Daniel, 2008, Measuring the quality of hedge fund data, *RiskMetrics Journal* Winter, 65–93.

Zitzewitz, Eric, 2006, How Widespread Was Late Trading in Mutual Funds?, *American Economic Review* 96, 284–289.