10-25-2010

Are Investors' Gains and Losses from Securities Fraud Equal Over Time? Theory and Evidence

Alicia J. Davis
University of Michigan Law School, alidavis@umich.edu

ARE INVESTORS' GAINS AND LOSSES FROM SECURITIES FRAUD EQUAL OVER TIME? THEORY AND EVIDENCE

Alicia J. Davis*
University of Michigan Law School
625 S. State Street
Ann Arbor, MI 48108
(734) 763-2221
alidavis@umich.edu

Draft dated October 25, 2010

1. Introduction

The conventional wisdom among securities regulation scholars has long been that diversified investors that trade frequently are hedged against the risk of secondary market securities fraud. The view is not merely that securities fraud risk is idiosyncratic and thus diversifiable like many other business risks (Davis Evans 2007).[1] This is widely believed to be true. The argument also is not simply that, in the aggregate, investor gains[2] and losses from fraud affecting only secondary market prices are equal. This is a mathematical certainty and unassailable.[3] Rather, the specific claim is that, for any individual[4] actively trading diversified investor, gains and losses from secondary market fraud are approximately equal over time. This view appears to have taken hold following the publication of a well-known article by Frank Easterbrook and Daniel Fischel (1985, 641), in which they state:

> Diversified investors act substantially as if risk neutral…An investor with a diversified portfolio will be the hidden gainer in a transaction [tainted by fraud] as often as he will be a loser. Every losing buyer … is matched with a gaining seller. Over the long run, any reasonably diversified investor will be a buyer half the time and a seller half the time. Such an investor perceives little good in a legal rule that forces his winning self to compensate his losing self over and over.[5]

---

[1] One contrary example is Booth (2007, 13), who asserts, "[t]he risk of simple securities fraud is like any other ordinary business risk."

[2] One typically does not think of an innocent investor as one that benefits from securities fraud. However, an investor that buys a stock on the open market before fraud causes the stock price to be artificially inflated, but then sells such stock at a time when the price is artificially inflated receives an improper gain on the transaction equal to the amount of the inflation. Thus, this investor unknowingly "benefits" from the commission of securities fraud. The price also could be deflated artificially in the case of "good news" fraud. Investors that buy while positive facts are withheld and hold until after the truth is revealed also are innocent beneficiaries of the fraud. See Booth (2007) for a brief discussion of "good news" fraud.

[3] Of course, as is widely acknowledged, in the presence of managers engaging in illegal insider trading, innocent outside shareholders will lose in the aggregate.

[4] Here, "individual" means "single" and not "natural person."

[5] Ultimately, Easterbrook and Fischel caution that completely eliminating compensation for securities fraud victims altogether could lead to allocative efficiency losses. Easterbrook and Fischel (1985, 641) state, "[T]he optimal damages in [secondary market fraud] cases are [not] zero just because most gains and losses net out. There will be the usual net harms of the costs of guarding against and litigating about the wrong, and there will be an allocative efficiency loss if transactions of a particular sort create uncompensated risk. The larger the transfer among investors, the more they will spend guarding against the problem." They, however, do go on to argue that any compensation should be significantly less than "the gross transfer of wealth" and that "there should be a presumption in favor of netting out gains before computing the [damages] award" (641-42).

Similarly, Janet Cooper Alexander (1996, 1502), in an influential article calling for an end to securities litigation and fraud victim compensation, asserts:[6]

> [M]any investors may not really need compensation from litigation, because they have diversified against the risk of securities violations…Institutional investors, in particular, hold diversified portfolios and trade frequently. The chance of being on the losing or winning side of a transaction when the stock price is distorted by a securities violation can be assumed to be random. The more trades investors make, the more likely that, in the aggregate, their gains from trading while material facts are withheld will equal their losses.

Compensation critics often take the argument one step further by asserting that being harmed by securities fraud is a choice; investors can protect themselves from the risk of fraud simply by diversifying. The following statement by Richard Booth (2007, 10-11) is representative of this view:

> [B]oth diversified and undiversified investors may suffer some harm from securities fraud. An undiversified investor who picks a single stock can lose her entire investment. But it does not follow that an undiversified investor should have a remedy if she voluntarily assumes the unnecessary risk that goes with failure to diversify. Again, *it is irrational for a passive investor not to diversify*. Securities law should protect only reasonable investors. A diversified investor is equally likely to be on the winning side of a trade as on the losing side. While a diversified investor with many different portfolio stocks may still suffer harm from an individual trade, the harm is likely to be small. *More important, gains and losses will net out over time. Such investors **are fully protected** from simple securities fraud through diversification.* They need no remedy (citations omitted, emphasis supplied).

---

[6] For a similar argument made contemporaneously, see Langevoort (1996, 646) ("[A]ctive traders with large, diversified portfolios have roughly the same chance of being winners as losers from securities fraud, and over time these gains and losses will tend to net out toward zero…").

2

The prevailing view on the netting of gains and losses from secondary market securities fraud has significant policy implications. Adherents to this view argue that diversified investors do not need compensation, and undiversified investors do not deserve it, as the law should make no effort to provide a remedy to those who fail to act reasonably. Were this view adopted by policymakers, there would be no principled basis, aside perhaps for deterrence effects and corrective justice considerations, on which to maintain a compensatory scheme.

The conventional wisdom on the netting out of fraud-related gains and losses, however, is at odds with probability theory. The argument against compensation implicitly equates fraud-related trading with a coin toss game in which a player wins $1 on heads and loses $1 on tails. While it is true that, if we accept this analogy (which I argue we should not), we should expect investors' fraud-related gains and losses, on average, to be equal over time, probability theory holds that as the number of trials (in this context, purchases and sales of fraud-tainted stock) increases, the lower the probability of being break-even. Thus, one would expect, consistent with probability theory, investors that trade frequently to be less likely, not more likely, to break even from fraud than investors that engage in fewer trades. Moreover, though it is true that with increased trials, the likelihood of the proportion of gains and losses being roughly equal will increase, investors generally do not engage in enough trading activity to have a reasonable degree of certainty in reaching the expected proportion of equal gains and losses.

To date, there has been little empirical evidence brought to bear on the question of the netting out of fraud-related gains and losses. In 2005, 20 years after the conventional wisdom in this area appeared to take hold, Thakor et al. published the first (to my knowledge) study on the issue. The study was commissioned by the U.S. Chamber of Commerce, and, in a reflection of the importance of this issue to the public more broadly, received high-profile mainstream media attention (see Glater 2005, Lehn 2006).

In the piece, the authors find that, though individual and small institutional investors, as a group, suffered significant harm from securities fraud over a 10-year period, large institutional investors suffered an average (median) *net* loss of only $5 million ($0.25 million). Declaring the average and median net losses trivial, the authors conclude that, as a group, large, diversified investors suffer virtually no net harm

3

from secondary market securities fraud and are overcompensated when one factors in recoveries from class actions. Thus, the results of the Thakor study appear to confirm the conventional wisdom: Though there are a few outliers with substantial losses,[7] on average, institutional investors are break-even with respect to fraud-related gains and losses.

Though an important contribution to our understanding of the equality of gains and losses from secondary market fraud, the Thakor study suffers from a few significant limitations. First, the study's results are not representative of the fraud-related gains and losses any *single* investor would have experienced over the 10-year study period. The universe of institutional investors used in the Thakor study includes investment managers with $100 million or more under discretionary management. Each quarter, these investors must file Form 13F statements with the Securities and Exchange Commission disclosing their investment holdings. Form 13F's aggregate data for fund families; they do not contain individual fund level data. Fund families generally contain several funds that pursue a number of strategies and hold a broad swath of the market.[8] However, individuals invest in funds, not fund families, and track the performance of different funds within a fund family separately.[9] Therefore, to understand whether any single fund is likely to suffer a net loss from securities fraud, one must somehow isolate the trades of that fund – something that is not possible with currently available public data.

A second limitation of the study is that the fraud-related gain and loss calculations the authors perform are imprecise. The authors do not have information on the exact timing of trades by the fund families under study, so they use the quarterly changes in the investment holdings disclosures to infer quarterly stock trades made by the institutions. In three separate analyses, the authors assume that stock

[7] The authors find some instances of what they consider extreme outliers in the study. Of the 2,596 investors in the study, only one investor had a net gain of over $200 million, but 18 investors had net losses in excess of $200 million. Eight of these 18 investors had losses of over $500 million, and one had a net loss of over $1 billion.
[8] For example, the 13F filed by Janus reflects the holdings of the entire family of Janus funds, not just a single fund. According to Janus' website (ww4.janus.com, last viewed on October 24, 2010), Janus offers retail investors 24 equity funds (including seven global funds) and three asset allocation funds (with portfolios that contain a mix of stock and other securities) that follow a variety of investment strategies and invest in stocks of companies of various sizes. When aggregating all of the holdings of these 27 funds, a large portion of the market is likely included.
[9] According to a survey of institutional investors by Cox and Thomas (2005), investment managers sometimes place recoveries from securities class actions in their general fund, rather than allocating such recoveries to the specific fund portfolio that held the stock and suffered the loss. If this practice occurs frequently, analyzing net gains and losses for fund families also may be worthwhile.

4

holding changes occur either at the end of the quarter, the middle of the quarter or in proportion to observed trading volume in the relevant stock during the quarter. Accurate gain and loss calculations require data on the exact dates on which the funds in the sample traded. These data are not publicly available.

Finally, and most importantly, though the study appears to confirm that, on average, investors are break-even from fraud, it does not provide detailed evidence on the distribution of fraud-related gains and losses among different types of investors. The latter point is the question with which I am primarily concerned.

In this Article, I test the following hypotheses, which are informed by relevant probability theory: (1) investors of all types, on average, break even from fraud and (2) active traders with diversified portfolios (prototypical institutional investors), not just undiversified retail investors, are at material risk for suffering substantial fraud-related losses. To overcome data limitations of the type faced by the authors of the Thakor study described above, this study supplements observational data on actual fraud cases and stock prices with data derived from simulated investor trades. Using available data on investor behavior, including level of portfolio diversification and trading frequency, I create thirteen different investor prototypes with characteristics consistent with representative institutional or retail investors. In a simulation that calculates fraud-related gains and losses from randomly selected trades, I find, during a 10-year period over 2,000 iterations, that the institutional investor prototypes, which hold diversified portfolios and trade relatively frequently, suffer net losses that, on average, are trivial. This finding is consistent with Hypothesis 1: investors, on average, do not suffer significant net losses from fraud.

The results of this simulation provide preliminary evidence not only on the likely average experience of selected investor types, however, but also, in an attempt to test Hypothesis 2, on the potential range of outcomes. This study reveals that though the mean net loss for the institutional investor prototypes in the simulation is effectively zero, there is significant variance in the outcomes for these investors. In this simulation, a large number of diversified investors, not merely a few outliers, enjoy significant fraud-related net gains or incur substantial fraud-related net losses over time. My findings,

5

therefore, refute the claims of investor compensation opponents who argue that (1) virtually no diversified institutional investor should suffer net harm from fraud over the long term and (2) individual investors can protect themselves from fraud-related net harm if they invest through mutual funds or other intermediaries.

In addition to simulating outcomes for the thirteen representative investor prototypes, I also calculate the fraud-related gains and losses for an investor prototype that mimics the behavior of the S&P 500[10] during the 10-year study period. The trades of this prototype are not simulated, but rather mirror additions to, and deletions from, the S&P 500 from 1996-2005. I find that this investor prototype is essentially break-even with respect to fraud-related gains and losses over this period.

Below Section 2 provides a brief discussion of the probability theory relevant to the study's design and findings. Section 3 more fully describes the data, sample used, and simulation methodology and presents simulation results. Section 4 concludes and briefly considers the policy implications of the study's findings.

2. Probability Theory

The conventional wisdom espoused by securities regulation scholars is that though there may be some outliers (presumably at either the positive or negative extreme), generally, fraud-related gains and losses will be equal over the long term for diversified investors that are active traders. Recall the following statement by Alexander (1996):

> [M]any investors may not really need compensation from litigation, because they have diversified against the risk of securities violations…Institutional investors, in particular, hold diversified portfolios and trade frequently. The chance of being on the losing or winning side of a transaction when the stock price is distorted by a securities violation can be assumed to be random. The more trades investors make, the more likely that, in

---

[10] The S&P 500 is an index of 500 large stocks trading on U.S. markets. The performance of the index is used as an indicator of the state of the overall equity market.

6

the aggregate, their gains from trading while material facts are withheld will equal their

losses.

Alexander (1996, 1502, n.58) goes on to state:

Of course, the offsetting of gains against losses reflects statistical probabilities; some

individual investors will be net gainers or net losers. The more trades that are made and

the more diversified the investments, however, the more an individual's experience is

likely to approach the statistical mean.

The premise of this argument is that an investor must be both well diversified and trade frequently for

gains and losses to be approximately equal over the long term. Recall that, in the aggregate, gains and

losses from secondary market fraud are equal, so the expected net gain or loss is $0. Those with the best

chance of having an experience that equals the mean (that is, no net losses or net gains), Alexander

argues, are those that reduce the variance by holding many stocks and trading frequently.

In this argument, the compensation critics appear to be analogizing fraud-related trading outcomes to

outcomes in the familiar fair coin toss game in which a player wins $1 if the coin comes up heads and

loses $1 if it comes up tails. It is well known that the longer the player plays (that is, the higher the

number of trials), the more likely the *proportion* of the times she wins will approach 50%; her expected

gain/loss from this game is $0 ((.5*$1)+(.5*-$1)). On average, players are neutral in this game.[11]

However, the use of this analogy in the context of fraud-related trading suffers from a number of

deficiencies. First, though it is true that with increased trials, the likelihood of the proportion of gains and

losses being roughly equal will increase, probability theory holds that as the number of trials increases,

the probability of the player being *break-even decreases* and eventually goes to zero. Thus, active traders

are less likely to be break-even with respect to fraud-related gains and losses than those that trade

infrequently.

Consider again the example of the fair coin toss game. Again, players are neutral in this game on

average. However, the longer the player plays (that is, there are more trials), the greater the chance for an

---

[11] For discussions of the probability concepts described in this section, see Appendix C and Appendix D.

extreme result (that is, something far from zero - either positive or negative). If she just plays the game twice there are only three possibilities: she may be up $2 (HH) (25% chance), down $2 (TT) (25% chance) or neutral ($0 gain/loss -- HT or TH) (50% chance). However, if she plays, for example, a million times, she could be up by $1,000 or more (approximately 2.3% probability), down by $200 or more (approximately 34% probability) or any number of other possibilities. The probability of being exactly break-even is vanishingly small. Thus, even though players are neutral in this game on average, any single player can have an outcome that is far from zero.

Coin tosses follow the rules of a Bernoulli distribution with probability p=.5. Therefore, over a large enough number of trials (*n*) the probability density function will converge on a normal distribution with μ (mean) = np and Sigma (standard deviation) = $\sqrt{(np(1-p))}$. Players in this game will be break-even on average, so μ = 0. However, sigma (standard deviation) allows us to understand the probability of departing from the expected average, and it increases as the number of trials (*n*) increases. So, for example, there is virtually a 0% probability of losing $25 or more in a game with 100 coin tosses, but that probability increases to 31% in a game with 10,000 coin tosses.[12] This is because sigma (standard deviation) increases as a function of $\sqrt{(np(1-p))}$ (that is, the higher the *n*, the higher the standard deviation).

Applied to the current context, if we assume for present purposes that trading in fraud-tainted stocks is analogous to a coin toss game, then the outcome from such trades also should follow the rules for a Bernoulli distribution. Therefore, investors will break even on average, so adherents to the conventional wisdom are correct in their assessment that μ = 0. However, this view overlooks the importance of sigma, which suggests that the more trials (i.e., fraud-tainted trades) in which an investor engages, the greater the sigma (i.e., the less likely she is to break even from fraud-related gains and losses).

---

[12] Such probabilities may be calculated by building probability density functions or perhaps more easily with the use of normal distribution probability calculators such as the ones available at http://www.analyzemath.com/statistics/normal_calculator.html or http://stattrek.com/Tables/Normal.aspx#faq. For a further discussion of why the number of trials decreases the probability of being breakeven in the context of a coin toss game with possible gains and losses, see also Feller 1950, 238-239 for a discussion of the "Return to Equilibrium" and Appendix E.

To determine which investors are least likely to be break-even from fraud, one must ask which investors are likely to have *more* trials, which, again, in this case, means *more fraud-tainted* trades.[13] The following predicts the number of fraud-tainted trades in which an investor will engage over the course of the simulation described in this Article[14]:

BUY SIDE: **(A)** [(Number of stocks in initial portfolio) * (Probability of purchasing a fraud-tainted stock on day 1 of the simulation)] ***PLUS*** **(B)** [[(Probability of purchasing a fraud-tainted stock during a single day of the simulation) * (Total number of purchases on that day)], Summed for each trading day of the simulation][15]

****PLUS****

SELL SIDE: **(C)** [(Number of stocks in initial portfolio) * (Probability of purchasing a future fraudster on day 1 of the simulation) * (Probability of selling that future fraudster stock while fraud is ongoing)] ***PLUS*** **(D)** [(Probability of purchasing the stock of a future fraudster during a single day of the simulation) * (Total number of purchases on that day) * (Probability of selling that future fraudster stock while fraud is ongoing)], Summed for each trading day of the simulation][16]

The formula above demonstrates that the number of relevant trials will vary by number of trades during the simulation period. At a given turnover ratio, an investor with more stocks in her portfolio will engage in more trades (i.e., an eight-stock portfolio and a 100% annual turnover ratio equals eight trades/year; a 100-stock portfolio and a 100% annual turnover ratio equals 100 trades/year). In general, therefore, institutional investors, which generally hold more stocks and trade more frequently than retail investors, are more likely to engage in fraud-related trades than individual investors. Moreover, staying invested in the market for a longer term (assuming the pre-existing rate of trading remains constant or

---

[13] We are concerned about the number of fraud-tainted trades and not total trades because trades that do not involve fraud-tainted stocks are not relevant to fraud-related gain and loss calculations.

[14] Note: As described in Section 3.3, I do not include losses stemming from purchases made on the first day of the simulation.

[15] The probability of purchasing a fraud-tainted stock must be calculated for each day of the simulation because, on any given day, there are different numbers of fraudsters in the market and different total numbers of stocks available for purchase.

[16] The probability of purchasing the stock of a future fraudster must be calculated for each day of the simulation because, on any given day, there are different numbers of future fraudsters in the market and different total numbers of stocks available for purchase.

9

increases) increases the number of trials, which, again, decreases the probability of an investor being break-even from fraud over the entire investment period.

Second, one may be unconcerned with investors being *exactly* break-even and may believe that being close to break-even, say, 49% win/51% loss, is sufficient.[17] Indeed, a reasonable reading of the compensation critics' argument is that increased trading gets investors roughly, not perfectly, to break-even. However, even under this view, an additional issue remains. It is true that increasing the number of trials increases the probability of achieving outcomes within the narrow band of 49/51.   With 100 coin tosses (trades), we would expect the number of heads (or tails) to fall within that band 24% of the time. With 1,000 tosses, we would expect it to occur approximately 50% of the time, and with 10,000 tosses, 95% of the time. (See Appendix C) What should be clear is that in order to eventually reach the proportion of, say, 49% wins/51% losses, one needs to have a large number of trials (fraud-tainted trades) in order to achieve outcomes within that band with a reasonable amount of certainty. For 49/51, 10,000 trials are necessary for a 95% success rate. According to data collected for this study (see Table 1), the average mutual fund makes 700 trades over a ten-year period, and a highly active trader makes 3,100 trades over the same time frame, far short of the 10,000 trials required for a 95% probability of success. However, what is relevant for these purposes is not total number of trades, but rather *fraud-tainted* trades, a far lower number. As described in Section 3.4.3, the simulation prototype for the typical mutual fund makes, on average, 12.7 fraud-tainted trades over a ten-year period, and the simulation prototype for the highly active trader makes 29.4 fraud-tainted trades over the same time period, on average. Of course, the rate at which the mutual funds in the simulation encounter fraud may understate the rate at which such investors trade fraud-tainted stocks in the real world because the simulation assumes random trading. In the real world, large institutions are not trading randomly, but rather trading stocks of large companies with a media following and research coverage. Though such companies are not necessarily more likely to engage in fraud than small companies (indeed, the opposite may be true), the high profile nature of the

---

[17] Of course, to an investor in the real world, 2% net losses on a $1 billion portfolio can be quite meaningful. See Section 4 for further discussion.

10

company may make it more likely for fraud to be uncovered ex post, leading to the large post-revelation stock price drop that does not occur in frauds that are never uncovered by the marketplace. Therefore, institutional investors may have higher exposure to fraud than represented by the simulation.[18] That said, even if the likelihood of engaging in a fraud-tainted stock trade were 10 times higher in the real world than my simulation suggests (translating into 127 fraud-tainted trades for the average mutual fund and 294 of such trades for a highly active fund), having a 95% chance of being within the desired 49/51 band (which requires 10,000 trades) still is out of the reach of investors over any reasonable time horizon.

Of course, a similar point applies to the argument above with respect to the probability of being break-even from fraud-related gains and losses. The expected probability of being exactly break-even declines as the number of trials increases, but the size of the decline is much larger if one goes from 100 trials to 1,000,000 trials than if one were to move from, say, one trial to 30 trials, which roughly reflects the range of fraud-related trades experienced by the prototypes in my simulation.

Finally, the argument of the compensation critics assumes that trading fraud-tainted stocks is akin to the coin toss game described above. However, stock trading is different because gains and losses from fraud are not the same for each instance of fraud. In the coin toss game, the player always loses $1 or gains $1 on a flip. This is not true in securities fraud cases. For example, an investor may suffer a large loss (or enjoy a large gain) from buying (or selling) an Enron-like stock while the fraud is ongoing or could face a more modest gain or loss by trading the stock of a company whose stock price dips following revelations of stock option backdating. So, imagine an investor that engages in two fraud-tainted trades (one "win" and one "loss") over some investment period. If the investor's "win" is selling stock in a company while the price was inflated due to a 2% overstatement of revenue, but the investor's "loss" is buying Enron stock at the height of its price inflation, the investor will not be break-even and will be a net loser, despite having a win to match its loss.

---

[18] Indeed, using the fraud cases in the instant study and institutional ownership data, I find a statistically significant positive relationship between the incidence of fraud over a prior ten-year period and level of institutional ownership.

11

The following illustration considers the consequences of the coin toss game if gains and losses per flip differ over time. Assume a game results in the following six flips:

Win for heads (loss for tails):	$4(-$4)	$3(-$3)	$2(-$2)	$1(-$1)	($5)(-$5)	$4(-$4)
Result:	T	H	T	H	T	H
Payment:	-$4	$3	-$2	$1	-$5	$4

The end result is the expected three heads and three tails, but, because the gain or loss per flip varied, the net payment is not $0 (as would be the case in the coin-toss game originally described above with consistent payouts over time), but rather -$3.   Given that this variation in gains in losses occurs in the context of trading in fraud-tainted stocks, it is very difficult to predict ex ante how an investor will fare over time with certainty.[19]

3. Data Sources, Sample, Methodology and Results of Analysis

3.1 Data and Sample

The period of this study is January 1, 1996–December 31, 2005. This project requires a sample of stocks with prices inflated by fraud.[20] To generate this sample, I begin with the list of all post-Private Securities Litigation Reform Act (PSLRA) class action settlements from 1996–2006 involving allegations of common stock price inflation due to fraud, as reported by Cornerstone Research (Simmons and Ryan 2007) and identified in the Institutional Shareholder Services' Securities Class Action Services (SCAS) database. From this list of 827 settlements, I exclude those involving cases (i) with class action periods of fewer than three days,[21] (ii) in which the alleged fraud did not affect secondary market trading prices, and (iii) with class action periods that did not overlap with the study period. I also add to the sample additional class action settlements for firms with multiple suits during the study period. Following these adjustments, a sample of 818 settlements remains. From this list of 818 settlements, I exclude 14 firms

---

[19] Of course, it is possible that (1) the amount of the gain or loss from each fraud is random and does not vary with any particular pattern and (2) the few frauds that do result in major losses do not skew the outcomes in a material way. If this is the case, the coin toss analogy, though still imperfect, would be reasonable. Further research on this issue is required.

[20] This study does not include any "good news" fraud cases.

[21] A meaningful inflation ribbon cannot be computed for a class action period of such short duration. See Section 3.2.2 for a discussion of the calculation of fraud-related gains and losses.

for whom the Center for Research in Security Prices (CRSP) has no stock price data and 151 settlements in which there was less than a 5% decline of the defendant corporation's stock at the end of the class action period.[22] The final sample consists of 653 settlements.

In this simulation, "fraud" includes only instances involving settled securities fraud class action lawsuits. This is an imperfect proxy for the incidence of fraud because there are cases in which no fraud actually occurs, but a corporation settles a meritless suit to avoid the time and expense of litigation. Conversely, there are cases in which fraud occurs, but no investor brings suit because the fraud cannot be proven or because the potential recovery is not high enough to attract the attention of a plaintiffs' lawyer.[23] In this study, this is of no concern. Because I am calculating both gains and losses from fraud, an over- or underinclusive fraud sample has no effect on the final outcome.[24]

For this study, I obtain split-adjusted closing stock prices, shares outstanding and trading volume data from the CRSP database. Data on the composition of the S&P 500 at the beginning of the study period (January 2, 1996), as well as on S&P 500 composition changes from 1996–1999, are from Standard & Poor's Index Services. Data on composition changes from 2000–2005 are from the Standard & Poor's website (www.standardandpoors.com). I obtain data on class action periods from SCAS. The universe of stocks included in the simulation (serving as a proxy for the overall market) consists of 14,696 different stocks -- every common stock appearing in the CRSP database during the 10-year study period.[25] The Morningstar[26] database provides data on mutual fund characteristics.

---

[22] Excluding frauds with associated stock price declines of under 5% is consistent with the approach taken by Thakor et al. (2005). Thakor et al. write:

> An end-of-period stock price decline of less than 5.0 percent for a defendant company suggested to us a set of circumstances in which the constant inflation ribbon method would not produce an inflation estimate that was generally consistent with the claims made in the class action litigation (app. I, 3).

[23] Also, a plaintiff could bring suit, but without sufficient evidence of fraud for the case to make it past the motion to dismiss, the plaintiff would have no hope of settling with the defendants.

[24] Thakor et al. (2005) make the same claim with respect to their study.

[25] The CRSP database contains information on virtually every publicly traded stock. Thus, stocks in the database are a good proxy for the overall equity market.

[26] Morningstar generates and publishes highly influential mutual fund ratings. See Section 4 for additional information on Morningstar ratings.

13

In this study, I separate and analyze fraud-related gains and losses by type of fraud. I include frauds that purportedly affected only secondary market trading prices ("secondary market frauds"), as well as other types of frauds that affected secondary market prices, even though not exclusively. I read the class action allegations for each of the 827 settlements initially considered for this study and coded each settlement by type of fraud ((1) fraud affecting secondary market prices only, (2) secondary market combination (that is, frauds affecting secondary market prices, as well as some other transaction such as a primary offering of securities), (3) primary market offering fraud, and (4) fraudulent statements made in connection with mergers and acquisitions transactions). All of these types of fraud make up the fraud sample in this study because these frauds affect secondary market prices until the fraud is uncovered. In most cases, I obtained the class action allegations from the SCAS database. However, if the SCAS database description was not sufficiently clear to allow me to code the type of fraud with confidence, I obtained allegation descriptions from the Stanford Securities Class Action Clearinghouse or from press reports.

3.2 Methodology

3.2.1 Investor Types

In this study, I simulate the trading behavior of 14 investor types. To help determine whether typical investors are likely to be net losers or net gainers from fraud over time, I use characteristics that reflect, to the extent possible, characteristics of actual investors. Each investor type differs based on: (1) trading strategy, (2) initial capital invested, (3) percentage of initial invested capital held in cash, (4) number of stocks in the portfolio and (5) annual turnover rate. Table 1 describes the characteristics of these investor types.

<u>Trading Strategy</u>

I employ three types of trading strategies in the simulation: (1) random selection, (2) random selection of popular stocks only and (3) S&P 500 parallel. A computer program developed specifically for this project controls the investor prototypes' trading decisions (see Appendix A for a description of the

14

program and sample code from the program). Investors employing the random selection and random selection of popular stocks strategies begin the simulation with a computer-generated random initial portfolio of stocks. The program then randomly selects stocks for these investors to buy and sell.[27] The investor employing the random selection of popular stocks strategy randomly purchases stocks chosen only from among a universe of the 25 most actively traded stocks (excluding exchange-traded funds) in a given year.

Because personal finance experts often advise individual investors to invest in index funds as a means of inexpensive diversification, this study also provides evidence, using only observational data, on the gains and losses of an investor whose trades track the changes in the S&P 500 over a 10-year period. The investor following the S&P 500 parallel strategy purchases, on the first trading day of the simulation (January 2, 1996), all 500 stocks comprising the S&P 500 on that date in the proportions each stock represents of the S&P 500. Throughout the simulation, this prototype sells stocks in its portfolio on the dates Standard and Poor's removes the stocks from the index, and the prototype purchases new stocks on the dates Standard and Poor's adds these stocks to the index. In addition, this investor rebalances its portfolio once a quarter to maintain relative weightings, by market capitalization, of the stocks in its portfolio. If the investor purchases a stock in the middle of a quarter, the newly acquired stock takes on the weight of the stock it replaces.

The S&P 500 index, on which I base this investor prototype, does not rebalance on a fixed schedule. However, replicating the precise schedule of rebalancing for this study is not practicable given data acquisition costs. Therefore, this investor follows the rebalancing practice of the S&P 500 Equal Weight Index (see Standard & Poor's 2006 at 6), for which rebalancings occur on the third Friday of each month that ends a calendar-year quarter (March, June, September and December).

---

[27] There is one exception to this general rule. When a stock held in an investor's portfolio ceases trading (for example, because it is acquired), the investor sells that stock on its last trading date and then immediately re-invests the proceeds in another randomly selected stock.

15

Initial Capital Invested and Cash Reserves

"Initial Capital Invested" represents the amount of cash an investor has available to purchase stocks for its initial portfolio. At the beginning of the simulation, the investor purchases stocks with total dollar values equal to, as near as possible, the amount of initial capital available for investment. When the investor sells a stock, it purchases shares of another stock as quickly as possible in an amount equal, as near as possible, to the proceeds it received for the sold stock. In some cases, certain investor types do not invest 100% of their available assets, but rather hold a certain percentage in cash. These amounts are "cash reserves." Typically, mutual funds hold a small percentage of assets in cash or other highly liquid securities to meet shareholder redemption demand.

Number of Stocks in the Portfolio

At the beginning of the simulation, each investor begins with a portfolio consisting of a pre-set number of stocks and continues to hold this number of stocks throughout the simulation.[28]

Number of Trades per Year and Over the Course of the Simulation

In the simulation, a "trade" is a simultaneous (or near simultaneous) sale and purchase of a stock. A prototype's annual turnover rate determines the number of trades in a given year. During the simulation, the investors execute this pre-set number of trades each year in equal trading day intervals over the course of the year. Multiplying this yearly total by 10 yields the total number of trades during the simulation period. There is one exception to this general trading rule. As described above, the prototype mimicking the S&P 500 trades on the dates stocks are added to, or removed from, the S&P 500.

Investor Type Characteristic Selection

I employ investor characteristics that are consistent with characteristics observed among actual investors. Because mutual funds are large, diversified investors, Investors 1, 2, 3, 4, 5 and 6 share characteristics with an actively managed mutual fund. To obtain information on mutual fund trading behavior, I performed a search in the Morningstar database for all domestic equity mutual funds in the

---

[28] The number of stocks held by the investor mimicking the S&P 500 may not always total 500 because the S&P 500, on occasion and for very brief periods, is composed of fewer than 500 stocks.

16

United States.[29] Appendix B describes the search parameters. In the simulation, I use prototypes that are representative (as determined by the Morningstar search results) of the typical mutual fund (average and median), as well as those at both extremes (that is, mutual funds that hold portfolios with the highest and lowest number of stocks and those that trade most and least frequently). Among the 333 funds in the search results, the average fund has total assets of approximately $547 million, with 3.5% in cash reserves, 89 stocks in its portfolio, and an 81% annual turnover rate.[30] I model Investor 1 in the simulation on the average mutual fund. The median fund has total assets of $109 million, 1.5% in cash reserves, 60 stocks in its portfolio, and a 59% annual turnover rate and provides the characteristics for Investor 2. The average fund in the top decile by diversification, represented by Investor 3, has total assets of approximately $1.1 billion, 2.5% in cash reserves, 300 stocks in its portfolio, and a 97% annual turnover rate.[31] The average fund in the top decile by trading frequency has total assets of approximately $325 million, 1.9% in cash reserves, 120 stocks in its portfolio, and a 263% annual turnover rate and provides the characteristics for Investor 4. The average fund in the bottom decile by diversification has total assets of approximately $532 million, 7.1% in cash reserves, 24 stocks in its portfolio, and a 45% annual turnover rate. I model Investor 5 on this investor type. Finally, the average fund in the bottom decile by trading frequency, represented by Investor 6, has total assets of approximately $1.1 billion, 4.4% in cash reserves, 76 stocks in its portfolio, and a 9% annual turnover rate.

Investors 7, 8, 9, 10, 11, 12 and 13 share characteristics with individual investors. According to survey data produced by the Investment Company Institute and the Securities Industry Association

---

[29] I performed the search after the market close on July 19, 2007. Though these observed characteristics apply only to domestic equity mutual funds, they are representative of actively trading institutional investors overall. Note that available evidence does suggest that mutual funds generally trade more actively than pension funds (see figures in Davis Evans 2007, app. 2). In addition, the characteristics simulated are those of mutual funds today. The portfolio composition and trading frequency of investors during the study period (1996-2005) could differ significantly from today's investor characteristics. However, since this is a simulation that attempts to demonstrate how investors of varying types might fare with respect to fraud, and not an attempt to specify how any particular investor fared during the study period, this is of limited concern.

[30] Morningstar, consistent with general practice, defines turnover as the lesser of the amount of purchases or sales of securities during a year divided by average monthly net assets. For the simulation, I translate the turnover statistics into a number of trades per year.

[31] To generate the characteristics for Investor 3, I rank order all mutual funds by number of stocks held. I use the average characteristics of all funds in the top 10% of this list as Investor 3's characteristics. I employ a similar process (using either number of stocks held or annual turnover rate and funds either in the top 10% or bottom 10%) to generate the characteristics for Investors 4, 5 and 6.

17

("ICI/SIA"), in 2005, of 744 retail investors surveyed, the average investor owned eight stocks and had individual stock assets of \$199,400 outside of employer retirement plans (Equity Ownership in America Appendices, 7, fig. E.4). The median retail investor held four stocks and had \$35,000 invested in equity securities. A survey of 2,187 individual investors reveals that 60% of these investors did not engage in any trades in 2004 (Equity Ownership in America, 25, fig. 34). Of the 40% that did trade, 57% traded five or fewer times during the year, and 79% traded 12 or fewer times during the year (Equity Ownership in America, 25, fig. 34).[32] In the simulation, Investors 7 – 10 reflect the average and median stock ownership attributes described above and either trade five times per year or 12 times per year. The simulation does not include prototypes that do not trade at all. Buy and hold investors (in the literal sense of this expression) can only be net losers from fraud if they buy a fraud-tainted stock; without sales, there are no opportunities for gains (see Davis Evans 2007 for further discussion on this point).

According to Meir Statman (2004), though a portfolio consisting of approximately 300 stocks is optimal for diversification today, in prior periods, according to researchers, portfolios consisting of 20 stocks or 30 stocks could afford sufficient, or even optimal, diversification. Therefore, to represent retail investors who attempt diversification consistent with historical standards, I include Investor 11, which holds a portfolio of 20 stocks, and Investor 12, which holds a 30-stock portfolio, in the simulation. Barber and Odean (2000) find that the average retail investor (household) portfolio has a 75% annual turnover rate. Therefore, to model this trading behavior, in the simulation, Investors 11 and 12 each have a 75% annual turnover rate. Their total assets for investment are \$200,000, approximately equal to the invested assets of the average individual investor surveyed by the ICI/SIA.

Investor 13 also has \$200,000 available for investment, but holds a 10-stock portfolio. It chooses its investments from among a group of popular stocks. As described previously, in the simulation, I define popular stocks as the 25 most actively traded stocks (excluding exchange-traded funds) in a given year. Barber and Odean (2008) find that individual investors are attracted to stocks that capture their

---

[32] This study defines an equity trade as either the sale or purchase of a corporate stock or shares in a mutual fund. Thus, the numbers used in the simulation likely overstate the level of direct trading activity by individual investors.

18

attention, for a number of reasons, including abnormally high trading volume. Investor 13 mimics, to some extent, an investor that trades only attention-grabbing stocks. Finally, Investor 14 mimics the behavior of an S&P 500 index fund and begins with an initial portfolio of $361.5 million, an amount equal to the invested assets of the median S&P 500 Index Fund.[33]

3.2.2 Calculating Gains and Losses from Fraud

I calculate gains and losses from all types of fraud, as well as for gains and losses exclusively from fraud coded as Type 1 (that which affects secondary market prices only). Consistent with the approach taken by Thakor et al. (2005), an investor's loss (gain) from fraud is the amount by which the price of the stock is inflated because of the fraud, multiplied by the number of shares purchased (sold) while the fraud is ongoing (that is, during the reported class action period).[34] The amount of inflation per share is equal to the difference between the stock price one trading day before the end of the class action period and the price on the trading day after the class action period ends.[35] As Thakor et al. explain, this is the constant inflation ribbon method, which assumes that the decline in the company's stock price at the end of the class action period (purportedly when the bad news is disclosed) reflects the amount by which the price was inflated because of prior misrepresentations. This study assumes this level of inflation is constant throughout the class period.[36]

Thakor et al. note that the price drop upon disclosure is not a consistently reliable way to measure the level of inflation in a stock due to fraud. Moreover, assuming the amount of inflation stays constant throughout the study period also is not realistic. However, just as in the Thakor et al. study, since I use

---

[33] This figure is based on the results of a Morningstar database search for S&P 500 index funds (n=38) on December 17, 2007.

[34] No gain or loss is recognized on in-and-out trades. An "in-and-out" trade is one involving both a purchase and subsequent sale or both a sale and subsequent repurchase during a class action period.

[35] Similar to the approach taken by Thakor et al. (2005), in this simulation, with respect to the gain or loss for any one investor, the amount of gain or loss is limited by the price at which the investor buys or sells the stock. For example, if the amount of inflation in a fraud-tainted stock is assumed to be $8 per share, and an investor purchases the stock during the class action period for $7 per share, the investor's loss from fraud is equal only to $7, not $8 per share. This investor is not permitted to "lose" a greater amount than its initial investment in any one firm.

[36] As noted in Section 3.1 above, stocks that experienced a price decline of less than 5% at the end of the class action period are excluded from the study.

19

the same method, albeit flawed, to compute both gains and losses from fraud, the results should not be biased and should provide a rough approximation of how investors benefit and lose from fraud.[37]

3.3 Study Limitations

When interpreting this study's results, one must be mindful of certain limitations. First, except with respect to Investor 14, which mimics the S&P 500, random selection guides the trading decisions of all investor types. Thus, the results for the investor types profiled in this study may not be representative of the results an investor that follows a particular strategy (for example, buying only large cap stocks, growth stocks or healthcare stocks) would achieve. In addition, because there are innumerable potential combinations of stock selections, this study, through its use of 2,000 trials per investor type, provides a narrow view of the potential outcomes that any investor could expect.[38] However, the trials do represent a sufficiently large sample to provide a sense for the range of potential outcomes for investors.

It is possible, though, in my view, unlikely that the use of random trading strategies in the simulation results in institutional investor losses being overstated. Some institutional investors are able to identify fraud-prone companies and avoid buying shares in such companies, but I am skeptical of most portfolio managers' ability to do this with precision and regularity. However, even if this is possible, employing a random trading strategy in my simulation should not bias my results downward. Let us assume Institutional Investor A, because of its superior research capabilities and manager insight, is able to avoid investing in companies that have a propensity to engage in fraud. If this is the case, Institutional Investor A will not be on the losing side of trades tainted by fraud, but it will not be on the winning side of such trades either. Hence, gains and losses from fraud will be equal ($0) for this investor. In order to believe that using a random trading strategy in my simulation would overstate losses relative to what an

---

[37] The gains and losses from fraud also are not inflation- or market-adjusted in my simulation. Therefore, the gains and losses in the late 1990's during the tech bubble are likely to be large relative to gains and losses in, for example, 1996. However, since I am calculating both gains and losses, the difference in the likely dollar value of gains and losses over the years should not bias the overall conclusion with respect to whether investors are on the winning side, by dollar volume, of trades tainted by fraud as often as they are on the losing side. Indeed, this is the way investors actually would experience gains and losses. I also ignore the effect of tax rate changes on the net proceeds received by investors in my gain and loss calculations.

[38] Recall that there are 14,696 different stocks an investor prototype could potentially purchase during the simulation period.

20

investor would experience in the real world, one would have to believe that a typical institutional investor in the real world is capable of avoiding purchasing stocks of companies currently engaging in fraud, and is also adept at identifying future fraudsters and timing its sales of the stocks of these firms to occur when the firm actually is engaging in fraud (so as to enjoy a fraud-related gain). This strikes me as unlikely for the typical investment manager. There is some evidence that short sellers, acting on presumably leaked information, do take positions in companies engaged in fraud (see, for example, Dyck, Morse and Zingales (2009), but short selling is not an investment strategy modeled in the simulation.

Second, the simulation requires prototypes to maintain a consistent investment strategy during the 10-year study period. In reality, investors sometimes change strategies, particularly in response to negative trading outcomes. Incurring substantial fraud-related losses is one example of a negative trading outcome.

Third, the reported gain and loss figures are negatively biased because of anomalies stemming from the artificial nature of the simulation design. In this study, the investor prototypes purchase an entire portfolio of stocks on the first trading day of the simulation (January 2, 1996), but do not sell all of the stocks in their portfolios at the end of the simulation.[39] This creates a negative bias because it would be possible for an investor prototype to suffer a loss from fraud at the beginning of the simulation by purchasing a fraud-tainted stock, but, because the investor does not sell its entire portfolio at the end of the simulation, significantly less likely for the investor to enjoy a fraud-related gain on the simulation's final day. To guard against this, I do not include in the simulation results losses incurred upon the purchase of the initial portfolio of stocks.[40]

However, a significant additional bias remains. The defined 10-year time horizon of the simulation biases the results in favor of losses, rather than gains, as during the simulation, the probability of selling a fraud-tainted stock is lower than the probability of buying a fraud-tainted stock. To buy a

---

[39] In reality, an investor usually does not liquidate its entire portfolio and exit the equity markets completely on a single day. Of course, an investor also usually does not purchase an entire portfolio (particularly in the case of a large portfolio of, for example, 300 stocks) on one day.

[40] An alternative approach, of course, would be to include losses on initial purchases and have a sale of the entire portfolio at the end of the simulation.

21

fraud-tainted stock and suffer a loss from fraud, an investor must purchase a stock with a price that is artificially inflated by misrepresentations. However, to sell a fraud-tainted stock and enjoy a gain from fraud, an investor must (i) purchase stock in a future (not current) fraudster *and* (ii) sell that stock at the time when the stock price is inflated because of fraud. Ending the simulation at a date certain compounds the problem because any "future fraudsters" purchased near the end of the simulation have little chance of being sold and resulting in a gain for an investor. This asymmetry biases the average results downward and explains in large part why the mean outcomes for the first thirteen investor prototypes are all negative. If I could continue the simulation into infinity, the average outcomes I report in this simulation would be reliable. Still, the principal conclusions of this piece, which relate to the variance of outcomes for investors, not average results, remain the same, despite the negative bias. [41]

Fourth, the investor prototypes in the simulation do not employ hedging strategies (for example, short sales, derivatives transactions) to minimize downside exposure. If the prototypes did so, and were able to so effectively, the net losses reported in this Article likely would be lower. That said, the failure to model hedging strategies is not likely to distort likely outcomes for mutual funds because the use of such strategies does not appear to be widespread among such investors. Indeed, one study (Koski and Pontiff 1999) finds that 79.2% of equity mutual funds do not use derivatives at all. This is likely due to the significant regulatory restrictions on the use of these and other hedging instruments by mutual funds (see Sjogren et al. 2005 for a discussion of some of these limitations). The unhedged trading strategy

---

[41] Another negative bias stems from the way the simulation program recognizes gains and losses. There is a natural symmetry to all fraud-tainted secondary market trades; each "loser" is matched with a "winner." Thus, because of this symmetry, one would expect two traders taking exactly opposite positions to have exactly opposite damage from fraud-related trades (that is, if one trader has a net loss of $5, the other trader must have a net gain of $5). However, given the design of this simulation, that would not necessarily be the case; losses can be more heavily weighted. An example follows. Assume the class action period for fraud involving Stock A is years 1-3. Imagine hypothetical Trader #1 purchasing Stock A as part of its initial portfolio at the beginning of year 1 and selling it at the end of year 3. This is an in-and-out trade for Trader #1, and Trader #1, therefore, suffers no fraud-related damage. Imagine further a hypothetical Trader #2 that always takes the opposite position of Trader #1. Trader #2 buys Stock A at the end of year 3 and suffers a fraud-related loss on that trade (Trader #2 purchases Stock A while the fraud is ongoing, but before it is revealed). Trader #2 has no opportunity for an offsetting (vis-à-vis Trader #1) gain because it cannot sell Stock A at the beginning of the simulation (short sales are not permitted under the simulation rules). As a result, Trader #1 has no gain or loss on the purchase and subsequent sale of Stock A, but Trader #2 has a net loss on the purchase of Stock A. This is an asymmetry and results in a negative bias.

22

modeled in this simulation is most like that of a mutual fund, but may differ substantially from that of a hedge fund or other type of institutional investor with limited regulatory oversight.

Finally, this study assumes that the amount of gain to an investor that unknowingly sells a stock while fraud is ongoing equals the amount of loss to the investor that buys that stock while the fraud is ongoing. However, as I discuss elsewhere (Davis Evans 2007, citing Hegde et al. 2003 and Booth 2007), there actually is a fundamental asymmetry between purported gains and losses. When the market discovers fraud, the firm's stock price usually does not decline only to the price that should prevail in the absence of fraud. Instead, the price typically declines by an additional amount to account for the uncertainty surrounding what additional negative disclosures may be forthcoming or what effect the legal process, in the form of lawsuits and government investigations, will have on the company. Therefore, because this study assumes that gains and losses from fraud are equal, I underestimate the extent of net losses from fraud an investor in the real world might experience.

3.4 Results

3.4.1 All Frauds – Institutional Investors

Table 2 presents the results for simulations that calculate gains and losses from all four types of fraud. As shown in Table 2, the simulation results suggest the typical institutional investor, as well as those that are highly diversified and have high turnover rates, are almost certain to buy or sell a stock tainted by fraud over an extended period. In all or almost all of the 2,000 iterations performed in the simulation, Investors 1, 2, 3 and 4 buy or sell at least one stock whose price was inflated due to fraud. However, Investors 5 and 6, representing the institutions that hold the fewest stocks and trade the least, only encounter a fraud-tainted stock 50% or 60% of the time, respectively. This result is consistent with the intuition that the more stocks one holds and the more frequently one trades, the more likely one is to buy or sell a fraud-tainted stock. Investor 14, whose trades follow actual changes in the S&P 500 index, encounters, as one might expect, fraud-tainted stocks during the simulation.

The simulation also reveals that, over the 10-year simulation period and among only the iterations in which a fraud-tainted stock is encountered, the institutional investors are essentially break-even with

23

respect to fraud-related gains and losses. For example, Investor 1, modeled on the average mutual fund and which begins the simulation with $547.1 million in invested capital, suffers a mean net loss of only $1.0 million. This loss is trivial when compared to Investor 1's initial capital investment, as the percentage of initial capital represented by the average net loss of Investor 1 is 0.2%. Investors 2-6 experience similar outcomes.

As shown in Table 2, Investor 14, which mimics the S&P 500, does not suffer a fraud-related net loss, but rather enjoys a net gain of $2.5 million. Investor 14 is still effectively break-even, however, as its net gain represents only 0.7% of initial capital invested.[42]

Though the simulation suggests that, on average, institutional investors are effectively break-even with respect to fraud, it also shows that investors may experience significant fraud-related losses or gains. The average mutual fund prototype, Investor 1, has simulated results that reveal, over the 2,000 iterations performed, a maximum net loss of $624.7 million and a maximum net gain of $827.2 million.[43] The maximum net loss represents 114.2% of Investor 1's initial invested capital, and the maximum net gain represents 151.2%. Investors 2-6 also experience extreme highs and lows. None has maximum gain or loss figures that could be considered trivial.

Moreover, the simulation reveals that investor loss and gain amounts steadily decrease or increase from approximately $0 to very large values. Thus, the results suggest an investor's gain or loss experience could take on any number of values, including very large ones – negative or positive. Figure 1

---

[42] One may be tempted to conclude that Investor 14's broad level of diversification or infrequent trading is responsible for it achieving a net gain from fraud. After all, as I discuss more fully below, diversification and trading levels play a significant role in investor outcomes. However, the evidence suggests that one cannot make broad generalizations about Investor 14's performance. To determine whether the positive performance of Investor 14 stems from a particular number of stocks held or trading frequency, I run the simulation for an additional prototype. The additional prototype I employ, which I name Investor 14A, is identical to Investor 14 in all respects, including specific trading dates, with one exception. Rather than trading stocks that have been added to or deleted from the S&P 500 index, Investor 14A selects stocks randomly. Though Investor 14A suffers, in a simulation for "all frauds," a small mean net loss of $8.5 million (approximately 2% of initial capital), the standard deviation of the outcomes is $25.6 million and, over 2,000 iterations, the prototype's outcomes range from a low of a $385.4 million net loss to a high of a $65 million net gain. Investor 14A experiences a wide range of possible outcomes and is not a net gainer or even neutral in every case. These results suggest that Investor 14's outcome is not the result of merely holding a large number of stocks or even trading with a particular frequency. It is far more likely that Investor 14's simulation result is just one of several potential outcomes for a broadly diversified investor.

[43] An investor may have a net loss or gain that exceeds its initial capital investment because over the 10-year period the value of the investment portfolio grows significantly. This growth exposes investors to the risk of large net losses or gains.

24

is a histogram of the middle 80% of net gain and loss figures over 2,000 iterations for Investor 1.[44] For Investor 1, the modal gain or loss hovers near zero, but there is not near universal clustering of net gains and losses near zero. Instead, the distribution of gains and losses appears to be almost normal, contrary to the conventional wisdom.[45] Figure 2 is a scatter plot of all the values that fall at the bottom 10% of the distribution for Investor 1. These values represent the iterations with the greatest losses.[46]

What becomes apparent after reviewing the histogram and scatter plot is that the net loss values steadily progress from a negligible amount to a significant value for Investor 1. For example, in Figure 2, the scatter plot representing the bottom 10% of the iterations for Investor 1, one is able to see that there is a continuous stream of points all the way to a $275 million net loss before the line breaks immediately before a series of scattered observations.[47] Values up to the break represent a net loss of up to 50% of Investor 1's initial capital invested. I observe somewhat similar breakpoint values for Investors 2, 3, 4, 5 and 6 at $62 million (57% of initial investment), $329 million (31% of initial capital), $192 million (59% of initial capital), $147 million (28% of initial capital), and $174 million (16% of initial investment), respectively.[48] Because over a 10-year period, the size of investors' portfolios can grow substantially, making comparisons between net losses and *initial* capital amounts may be misleading. Therefore, I also calculate breakpoint values as a percentage of the average final capital amount, calculated over all 2,000 iterations for each investor. As a percentage of final capital, Investor 1's breakpoint value equals 16%,

---

[44] Unlike the net gain or loss figures in the tables, which report the results only of those iterations in which an investor encounters fraud, this histogram includes the results of the middle 80% (in order by magnitude of gain or loss) of all 2,000 iterations. This histogram includes only the middle 80% of values so the reader may more clearly see the range of values of the gains and losses for Investor 1. Because of the large dispersion of values (that is, the extreme values on either end), including 100% of the iterations would result in the differences in the middle 80% being lost due to the scaling of the graph. I only include a histogram for Investor 1 in this Article, but the histograms for the other institutional investor prototypes are similar to that of Investor 1.

[45] Note that, despite appearances, the Shapiro-Wilk W test I perform rejects the null hypothesis that these net gain and loss data are from a normal distribution.

[46] I only provide a scatter plot for the values at the bottom of the distribution in this piece. However, a scatter plot for values at the top of the distribution (that is, the largest gains) reveals a similar pattern to the graph shown in this Article. In addition, I include a scatter plot only for Investor 1 in this Article, but the scatter plots for the other institutional investor prototypes are similar to that of Investor 1.

[47] There are sophisticated analytical techniques that one may use to determine break points or discontinuities in a series of data. I have chosen not to employ those methods here. The discussion below on probability of these outcomes should provide sufficient detail on an investor's likely experience in this regard.

[48] As noted in Footnote 46, scatter plots for Investors 2, 3, 4, 5 and 6 do not appear in this Article.

25

Investor 2's equals 18%, Investor 3's equals 11%, Investor 4's equals 18%, Investor 5's equals 9%, and Investor 6's equals 6%.

Of course, the fact that we observe these results after 2,000 simulations means that these outcomes are possible, but it does not necessarily mean an investor in the real would necessarily face a high probability of these occurrences. Instead, the simulations provide useful data points for building the probability density function that one can use to determine the probability that an investor type will experience a specific fraud-related gain or loss amount. For example, in unreported results, based on the simulation outcomes, I calculate the probability[49] of Investor 1 experiencing a net loss of $275 million or more (what I term the "breakpoint" above) as close to 0%. However, Investor 1 has a 1.5% chance of having a net loss of $100 million or more, a 5% chance of having a net loss of $75 million or more, a 14% chance of having a net loss of $50 million or more, and a 33% chance (a significantly more meaningful risk) of having a net loss of $20 million or more. Investor 1 also, of course, faces the equivalent prospect of net gains. A gain or loss of $100 million represents 18.3% of Investor 1's initial capital and 5.8% of its final capital amount, and $20 million is 3.7% of Investor 1's initial capital and 1.2% of its final capital amount.

3.4.2 All Frauds – Individual Investors

The simulation results, as shown in Table 2, for Investors 7–12 (all retail investor prototypes) conform to the expectations of researchers. Individual investors with small portfolios who trade infrequently generally encounter stocks tainted by fraud less frequently than institutional investors do. Among these investors, the percentage of iterations in which fraud is encountered ranges from 22.7% for Investor 9, representing the median individual investor with a four-stock portfolio and five trades a year, to 97.1% for Investor 12, representing a reasonably well-diversified retail investor who holds 30 stocks and makes 23 trades per year.

Table 2 also shows the average outcome for the retail investor prototypes. Among only the iterations in which a fraud-tainted stock is encountered, each of Investors 7-12 experiences a fraud-related

---

[49] I assume, for ease of analysis, that fraud-related trading outcomes have a standard normal probability distribution.

26

net loss. The average net losses over the 10-year period appear to be modest, but, as a percentage of initial invested capital, they are substantial. The net losses range from 0.9% of initial capital for Investor 11, an atypical and the second most diversified individual investor prototype with a 20-stock portfolio, to 15.2% of initial capital for Investor 10, one of the least diversified investors with a four-stock portfolio. One should note, as discussed previously in Section 3.3, that these figures are biased downward because of the artificial nature of the simulation. The negative bias inherent in the simulation design affects all investor results, but has a greater effect on the results of the retail investor prototypes. If a retail investor prototype encounters a fraud-tainted stock, because it holds a portfolio consisting of just a few stocks, it will buy or sell more, on a percentage basis of its total investment, of that fraud-tainted stock – hence the downside of lack of diversification. Given the simulation's negative bias, the retail investor prototype is slightly more likely to be on the losing side of that trade and hence to have bigger simulated net losses overall.

As with the institutional investor prototypes, there are cases of extreme highs and lows. As a percentage of initial capital, the maximum net gain for these investors ranges from 97.0% (for Investor 7 with an eight-stock portfolio and moderate turnover rate) to 403.2% (for Investor 8 with an eight-stock portfolio and high turnover rate). The maximum net loss for this group of investors ranges from 175.2% (for Investor 11 with a 20-stock portfolio and high turnover rate) to 1,351.4% (for Investor 10, one of the least diversified investors (holder of four stocks) who also has a high turnover rate).

Investor 13 (10-stock portfolio, $200,000 in initial capital, high turnover, and ownership of only popular stocks) encounters fraud in 99.9% of the iterations – a higher percentage than any other individual investor prototype. This is understandable given Investor 13's strategy of only purchasing the 25 most actively traded stocks in a given year; large companies with heavy stock trading volume often are attractive targets of securities fraud litigation.[50] Investor 13 experiences an average net loss of $26,899 (13.5% of initial invested capital). The highest recorded net gain for Investor 13 is $1,590,390

---

[50] Recall, as described in Section 3.1, that instances of "fraud" in this simulation include only alleged fraud in connection with settled securities class actions.

(795.2% of initial invested capital), and the highest net loss is $5,089,922 (2,545.0% of initial invested capital). Recall that Investor 13 is not typical in that its behavior is not that of the average or median individual investor. However, it does follow a similar strategy to that of individuals that purchase stocks that are in the news frequently or that are household names.

3.4.3 All Frauds – Relative Results

To understand the relative performance of the prototypes representing typical institutions and individuals, I perform simulations in which all investor characteristics are identical to the ones described in Section 3.4.1 above, with two key exceptions: I hold initial capital constant at $1 million and cash reserves constant at 0%. Table 3 contains the results of these simulations.

The simulations reveal that, in all cases, the investor prototypes modeled after large institutional investors fare better, on average, than the prototypes modeled after retail investors. Investors 1-6 and 14, the actively traded mutual fund and index prototypes, all have better average net fraud-related trading gains/losses than those of Investors 7-13, the individual investor prototypes.

The data in Table 3 also show that the retail investor prototypes in the simulation generally have larger standard deviations than the institutional investor prototypes. This suggests that retail investors have higher statistical variance than institutional investors and thus have the lowest probability of breaking even from investing in fraud-tainted stocks. This is explained largely by the lack of diversification of the retail investor prototypes' portfolios. One should also note that the inherent biases in the simulation design likely also play a role in this result.

The simulations[51] also allow one to assess whether the investors most likely to have a higher number of fraud-tainted trades and, in turn, higher variance in outcomes are those that trade frequently. I find that, over the 2,000 iterations in which fraud is encountered, on average, Investors 1-6 engage in 12.7, 5.2, 28.8, 29.4, 1.91, and 1.87 fraud-tainted trades, respectively. Thus, Investor 4, with a 120-stock portfolio, a 263% annual turnover rate (the highest among the mutual fund prototypes) and 3,100 total

---

[51] Though the parameters are the same as those for the simulation otherwise described in this Section, the results reported in this and the succeeding paragraph reflect the results of another 2,000-iteration simulation.

28

trades, has the most fraud-tainted trades over the 10-year simulation period. The prototype with the second highest number of fraud-tainted trades is Investor 3, which holds 300 stocks (the highest among the mutual fund prototypes), has an annual turnover rate of 97% and 2,830 total trades. The prototypes with the least fraud-tainted trades over the 10-year simulation period are Investor 6, with a 76-stock portfolio, 9% annual turnover (the lowest among the prototypes) and 60 total trades, and Investor 5, which holds 24 stocks (the lowest among the prototypes), has a 45% annual turnover rate and makes 100 total trades.

I then calculate the variance of outcomes from this simulation to assess whether, as expected, the institutional investor prototypes with the highest number of average fraud-tainted trades have higher variance across outcomes than the other institutional investor prototypes. Based on the average number of fraud-tainted trades, one might expect Investor 4 (with 29.4 trades) to have the highest variance across outcomes, followed by Investor 3 (with 28.8 trades), Investor 1 (with 12.7 trades), Investor 2 (with 5.2 trades), Investor 5 (with 1.91 trades) and Investor 6 (with 1.87 trades). However, the results reveal one deviation from that expected pattern: Investor 3 has the highest variance, followed by Investor 1, Investor 4, Investor 2, Investor 5 and Investor 6. The likely explanation for the single anomalous result (that is, Investor 4's ranking) is that the number of transactions is not fixed for any single prototype. I am reporting averages, which themselves are random variables over the range of iteration outcomes, so I cannot compare this average number to the variance. For example, though Investor 4 engages in, on average, 29.4 fraud-related trades, the range across 2,000 iterations is $10 - 59$ fraud-related trades. In addition, another source of variation exists in this simulation. The dollar amount of gains and losses also is random and not identical across trials (as, described previously, is the case in a coin-toss game), but instead varies by fraud-tainted stock encountered. This, too, explains, in part, the anomaly.

The simulations also provide data that may be used to assess the probability of various outcomes for the institutional investors with varying trading levels. For example, Table 3 reveals that the standard deviation of the trading gain/loss for Investor 6, an infrequent trader that holds 76 stocks in its portfolio and makes six trades per year, is $26,244. In contrast, the standard deviation of the trading gain/loss for

29

Investor 4, a highly active trader that holds 120 stocks and makes 310 trades per year, is significantly higher at $108,199. Investor 6 (the infrequent trader whose trading outcomes have lower variance), therefore, has a lower probability of experiencing a material net loss due to fraud than Investor 4 (a frequent trader). Specifically, the probability that Investor 6 would lose, for example, 10% ($100,000) or more of its initial capital on a net basis from fraud-related trades is essentially 0%, whereas the probability that Investor 4 would experience such a loss is approximately 18%.

3.4.4 Secondary Market Frauds Only

Tables 4 and 5 show the results of simulations that are similar to the ones described above with one key exception – the only instances of fraud recognized are those frauds affecting only secondary market trading prices (Type 1 frauds). With few exceptions (including a negligible loss for Investor 14, the S&P 500 prototype, rather than a slight gain), the gain and loss results, as presented in Table 4, are qualitatively identical to the ones described above for the simulations including all frauds. The relative performance of the investor prototypes in the standardized simulation (that is, one in which all prototypes have $1 million in initial capital), as shown is Table 5, is similar to that described above for simulations reflecting gains and losses for all fraud types.

4. Discussion and Conclusion

Easterbrook and Fischel (1985) observed that the loss one innocent investor suffers from buying a fraud-tainted stock is offset by the gain another innocent shareholder receives from selling the stock while the fraud is ongoing. In addition, the scholars asserted, over the long term, the "wins" and "losses" from fraud for any single actively trading diversified investor should be approximately equal. A number of leading scholars have used this insight to call for the elimination of compensation for securities fraud losses. The study described in this Article generates evidence that should prove useful to the ongoing debate.

First, this piece provides information on how broadly the securities fraud problem could affect investors. To my knowledge, no other study has done so. The simulation reveals that the typical institutional investor is almost certain to trade a fraud-tainted stock over a 10-year period and that most

30

institutions that trade infrequently or hold few stocks are, as well. The results also demonstrate that even individuals that hold small portfolios and trade infrequently have a reasonable likelihood of encountering a fraud-tainted stock if they participate in the market over an extended period. For example, in the simulation, 23% of investors with small, four-stock portfolios that make just five trades a year (the characteristics of Investor 9), over a 10-year period, trade a stock with a price artificially inflated by fraud. This is a significant figure[52] and demonstrates that securities fraud has the potential to affect a large number of investors – both institutional and individual.[53]

Of course, according to the data produced by the simulation, the probability of engaging in a fraud-tainted stock is low on an absolute basis. Recall that Investor 4, the most active trader with 3,100 trades over a 10-year period, only engages, on average, in 29.4 (0.95% of the total) trades tainted by fraud. However, despite the low absolute number of fraud cases to which an investor may be exposed, the potential losses from these cases can be significant.

Second, the simulated results in this study provide evidence that is consistent with Hypothesis 1: on average, investors are essentially economically neutral with respect to gains and losses from fraud. Moreover, using observational data only, I find that an investor in an S&P 500 index fund during the 10-year study period would have been essentially break-even from fraud. Thus, this study demonstrates that the conventional wisdom is correct in this regard.

However, probability theory and the results of this simulation also provide a view that differs significantly from the conventional wisdom. Though, on average, the investor prototypes suffer no net losses from fraud, consistent with Hypothesis 2, there are many investors, not just a few outliers, that

---

[52] This likely understates the matter because, as explained previously, this simulation does not capture all fraud – only fraud lawsuits that result in settlements. In addition, this figure excludes any purchases of fraud-tainted stocks at the beginning of the simulation.

[53] These results stem from simulations involving the random selection of stocks and do not reflect outcomes for investors that make conscious investment decisions. Nonetheless, because of the *random* selection, the simulated outcomes provide a sense of the likely experience of the average investor of a certain type.

suffer net losses or enjoy net gains from fraud. Thus, not every (or even virtually every) institutional investor prototype breaks even "over the long term" as the compensation critics assert.[54]

In the real world, things could be even more dire for investors than the simulated results suggest. The argument of the compensation critics rests in part on the assertion that, *over the long term*, gains and losses from fraud are equal. However, some investors will not stay invested in the market for the long term either because they run out of money or because they lose the will or ability to continue losing money after suffering substantial losses. This phenomenon can affect institutions (whose funds close because of poor performance, for example) and individuals alike.

This simulation requires investor prototypes to stay in the market for the entire 10-year study period. Therefore, the behavior of the prototypes, by design, does not reflect the "risk of ruin." The risk of ruin refers to the possibility that even though an investor can achieve expected returns over some pre-defined period, the standard deviation (variability of outcomes) can result in the investor losing an unacceptable amount in the earlier periods before it is able to achieve the returns expected over the entire investment period. So, in the current context, even if an investor would be essentially break-even from fraud over a 10-year period (which, I argue, is not at all assured), the investor may never achieve that outcome because it exits the market after suffering unacceptably high fraud-related losses in the early years.[55] In the real world, unlike in this simulation, there generally is nothing to keep investors from exiting the market. Thus, real world outcomes could be worse than the simulated ones.[56]

---

[54]Of course, one must be mindful of the fact that frequent traders have not only the potential for losses, but also the potential for significant outsize gains from fraud.

[55] An investor also may leave the market long before its net fraud-related losses reach extreme levels. This could offset somewhat the risk of ruin effect discussed above in that losses in the real world would be lower than they appear in these simulated results.

[56] Of course, some investors who would have broken even had they remained in the market for 10 years could exit the market early after enjoying early large gains, as a way to "quit while they are ahead." However, early gains – even related to fraud -- are more likely to make an investor more confident in her investing abilities and keep her invested in the market.

32

The findings described in this Article, which focus only on the narrow question of whether any one investor's gains and losses from fraud are equal over time,[57] do not necessarily lead to clear policy prescriptions. On average, investors are break-even from fraud. Thus, if one were concerned only with investor profit overall, one might argue that, on average, investors suffer no net losses from fraud so providing compensation to them is inefficient. However, though an investor's expected loss from fraud is zero, a number can suffer extreme losses. Therefore, those concerned with corrective justice likely would find the potential for uncompensated losses of the magnitude revealed to be possible by this simulation troubling and would be unlikely to change their views even in light of claims that many investors enjoy significant net gains from fraud.[58]

A concern for overall social welfare also cautions against eliminating compensation for securities fraud losses because of the potential effect of uncompensated fraud losses on allocative efficiency. Because an investor does not know ex ante where her experience will fall on the distribution of fraud-related gains and losses, she has an incentive to guard against the risk of fraud. As I argue elsewhere (see Davis Evans 2007), investors are likely to take precautions (above the optimal amount)[59] to minimize their downside exposure. This can lead to allocative inefficiencies, as argued by Easterbrook and Fischel (1985), that result from investors expending energy to ferret out fraud risk, at the expense of other, more productive investment-related activities. Indeed, as they argue, the greater the dispersion of potential outcomes (and possibility of extreme losses), the greater the likely preventative and monitoring costs of investors.[60] Fear of fraud also can lead to investors opting out of the market altogether, which has negative effects on capital formation.

One might argue that the likely net losses from fraud will be modest in relation to capital under management and thus be of little concern to institutional investors. Recall, for example, from Section

---

[57] Recall that, as stated in the introduction, diversification generally provides protection against idiosyncratic risk (of which suffering losses from fraud is one type) without respect to whether or not fraud-related losses are offset by fraud-related gains over time.

[58] For a discussion of the law's asymmetry with respect to benefits conferred and harms imposed, see Hershovitz (2006 , 1149) ("…on the model of harms, the harm-benefit asymmetry is a natural expression of the moral difference between harming and benefiting").

[59] Some level of shareholder monitoring is, of course, beneficial.

[60] Of course, not all investors have the ability to invest in precaution or an interest in doing so.

33

3.4.1 that Investor 1 faces a 14% chance of losing $50 million or more from fraud, but this net loss amount represents 9.1% of initial capital and only 2.9% of final capital. However, first, one should also keep in mind that net loss figures reflect offsets of fraud-related gains. Because fraud-related gains are not something an investor realizes are gains at the time it achieves them (or perhaps ever), the investor will feel as though it has lost a great deal more than the net loss figures suggest.

Second, it is, of course, true that investors will spend resources taking precautionary steps only as long as they believe it will provide benefits. However, there is reason to believe that even principals of actively managed investment funds have incentives to take excessive precautions in this context. Even seemingly small losses can be economically significant to these investors, whose managers are extraordinarily focused on returns. For example, Del Guercio and Tkac (2002) find that for every additional 1% of lagged excess return (over the S&P 500), mutual fund managers and pension fund managers enjoy 0.93% and 0.67%, respectively, in additional net asset growth.

Perhaps even more important to mutual funds than actual returns are ratings from Morningstar, Inc. Del Guercio and Tkac (2008) find that Morningstar ratings affect investment in a fund, independently from other fund performance measures. Morningstar's highly influential mutual fund rating system[61] gives the top 10% of mutual funds 5-star ratings, the next 22.5% 4-star ratings, the next 35% 3-star ratings, the next 22.5% 2-star ratings, and the bottom 10% 1-star ratings. Researchers generally find that individual investors flock to 5-star and 4-star funds and sell funds if they fall to 3-star status. Indeed, Del Guercio and Tkac (2008) find that, though not all ratings changes have a significant effect on fund asset flow, some changes can have a substantial impact. In particular, a star rating upgrade from 4-stars to 5-stars results in a 25% above average expected flow over the seven-month period following the rating change, and a downgrade from 4-stars to 3-stars results, on average, in a fund outflow that is almost 1,200% larger than the expected normal flow.

---

[61] Morningstar takes risk-adjusted returns relative to peer group into account when assigning ratings. See Del Guercio and Tkac (2008) and sources referenced therein for a more comprehensive discussion of the Morningstar rating system.

34

According to Morningstar (2007), in 2004, domestic equity funds rated 4-star outperformed funds rated 3-star, with respect to one-year total return, by only 49 basis points (0.49%).[62] Thus, what some may view as modest net losses from fraud could have an economically significant influence on a fund's inflows and outflows if losses from fraud caused a fund to move from one star rating to another. Mutual fund managers are likely to care a great deal about these losses.

Moreover, the Morningstar rating system penalizes funds not only for poor overall returns, but also takes downward variation specifically into account. As Morningstar (2008) describes:

> The Morningstar Rating[TM] is based on "expected utility theory," which recognizes that investors are a) more concerned about a possible poor outcome than an unexpectedly good outcome and b) willing to give up some portion of their expected return in exchange for greater certainty of return. The rating accounts for all variations in a fund's monthly performance, with *more emphasis on downward variations*. (Emphasis supplied)

Therefore, fund managers who market to retail investors have incentives not only to maintain strong overall fund performance, but also to avoid negative surprises like losses from securities fraud.

In this Article, I describe how historically scholars have mischaracterized the likelihood of an institutional investor breaking even with respect to fraud-related gains and losses and demonstrate that though investors, on average, are break-even with respect to fraud, the variance of outcomes can be high for all types of investors. The evidence provided in this Article, therefore, should prompt a rethinking of the argument that continues to dominate the debate on the efficiency of compensating securities fraud victims.

---

[62] This represents an ex post look at how funds rated by Morningstar fared with respect to returns, not an ex ante input to the formula that determined the rating (though returns are an important part of the rating formula). However, it does demonstrate that small differences in return can separate mutual funds in the minds of investors in meaningful ways.

References

Alexander, Janet Cooper. 1996. "Rethinking Damages in Securities Class Actions." *Stanford Law Review* 48:1487-1537.

Barber, Brad M., and Terrance Odean. 2000. "Trading is Hazardous to Your Wealth." *Journal of Finance* 55:773-806.

Barber, Brad M., and Terrance Odean. 2008. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies* 21:785-818.

Booth, Richard A. 2007. "The End Of The Securities Fraud Class Action As We Know It*." Berkeley Business Law Journal* 4:1-36.

Campbell, John, Martin Lettau, Burton Malkiel, and Yexiao Xu. 2001. "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk." *Journal of Finance*, 56:1–43.

Cox, James D., and Randall S. Thomas. 2005. "Letting Billions Slip Through Your Fingers: Empirical Evidence and Legal Implications of the Failure of Financial Institutions to Participate in Securities Class Action Settlements." *Stanford Law Review* 58:411-54.

Davis Evans, Alicia. 2007. "The Investor Compensation Fund." *The Journal of Corporation Law* 33:223-96.

Del Guercio, Diane, and Paula A. Tkac. 2002. "The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds vs. Pension Funds." *Journal of Financial and Quantitative Analysis* 37:523-57.

Del Guercio, Diane, and Paula A. Tkac. 2008. "Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow." *Journal of Financial and Quantitative Analysis* 43:907-36.

Dyck, Alexander, Adair Morse, and Luigi Zingales. 2009. "Who Blows the Whistle on Corporate Fraud?" forthcoming *Journal of Finance*. Available at http://www.afajof.org/afa/forthcoming/4820.pdf.

Easterbrook, Frank H., and Daniel R. Fischel. 1985. "Optimal Damages in Securities Cases." *University of Chicago Law Review* 52:611-652.

Feller, William. 1950. *An Introduction to Probability Theory and its Applications*, Vol. I. New York: John Wiley & Sons, Inc.

Glater, Jonathan D. 2005. "Critics of Shareholder Suits Aim at Big Holders." *New York Times*, October 27.

Hegde, Shantaram P., Chris B. Malone, and John D. Finnerty. 2003. "The Financial Impacts of Fraud and Securities Class Action Suits." Working Paper.

Hershovitz, Scott. 2006. "Two Models of Tort (and Takings)." *Virginia Law Review* 92:1147-88.

Investment Company Institute and Securities Industry Association. 2005. *Appendices: Additional Figures for Equity Ownership in America, 2005*. Available at http://www.ici.org/pdf/rpt_05_equity_owners_append.pdf.

36

Investment Company Institute and Securities Industry Association. 2005. *Equity Ownership in America, 2005.* Available at http://www.ici.org/shareholders/dec/1rpt_05_equity_owners.pdf.

Koski, Jennifer Lynch and Jeffrey Pontiff. 1999. "How are Derivatives Used? Evidence from the Mutual Fund Industry." *Journal of Finance* 54:791-816.

Langevoort, Donald C. 1996. "Capping Damages for Open-Market Securities Fraud." *Arizona Law Review* 38:639-664.

Lehn, Kenneth M. 2006. Commentary, "Private Insecurities" *Wall Street Journal,* February 15.

Morningstar. 2007. "How the Morningstar Rating Has Performed." Available at http://news.morningstar.com/PDFs/FSNStarTable.pdf.

Morningstar. 2008. "The Morningstar Rating^{TM} for Funds." Available at http://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf

Simmons, Laura E. and Ellen M. Ryan. 2007. Cornerstone Research. "Securities Class Action Settlements, 2006 Review and Analysis." Available at http://www.cornerstone.com/pdf/practice_securities/2006Settlements.pdf.

Sjogren, Dee Anne, Brian R. Wahlquist, and Seth B. Chertok. 2005. "Limits on Mutual Funds' Use of Hedging Strategies and Derivatives." *The Investment Lawyer* 12:1,14-19.

Standard & Poor's. 2007. "S&P 500" (Fact Sheet). Available at http://www2.standardandpoors.com/spf/pdf/index/500factsheet.pdf.

Standard & Poor's. 2006, "S&P Equal Weight Index: Index Methodology" available at http://www2.standardandpoors.com/spf/pdf/index/SP_Equal_Weight_Index_Methodology_Web.pdf.

Statman, Meir. 2004. "The Diversification Puzzle." *Financial Analysts Journal* 60:44-53.

Thakor, Anjan V., with Jeffrey S. Nielsen, and David A. Gulley. 2005. "The Economic Reality of Securities Class Action Litigation." U.S. Chamber Institute for Legal Reform.

37

Table 1
Investor Characteristics

Investor No.	Trading Strategy	Prototypical Investor Type	Initial Capital ($)	Cash Reserve (%)	No. Stocks in Portfolio	Turnover (%)	Trades Per Year	Total No. Trades
1	Random Selection	Mutual Fund (Average)	547,057,057	3.5	89	81.0	70	700
2	Random Selection	Mutual Fund (Median)	109,000,000	1.5	60	59.0	35	350
3	Random Selection	Mutual Fund (Highly Diversified)	1,069,424,242	2.5	300	97.0	283	2,830
4	Random Selection	Mutual Fund (Highly Active Trader)	325,121,212	1.9	120	263.0	310	3,100
5	Random Selection	Mutual Fund (Low Diversification)	531,636,364	7.1	24	45.0	10	100
6	Random Selection	Mutual Fund (Infrequent Trader)	1,094,393,939	4.4	76	9.0	6	60
7	Random Selection	Retail Investor (Average Assets/Diversification, Moderate Turnover)	199,400	0.0	8	62.5	5	50
8	Random Selection	Retail Investor (Average Assets/Diversification, High Turnover)	199,400	0.0	8	150.0	12	120
9	Random Selection	Retail Investor (Median Assets/Diversification, Moderate Turnover)	35,000	0.0	4	125.0	5	50
10	Random Selection	Retail Investor (Median Assets/Diversification, High Turnover)	35,000	0.0	4	300.0	12	120
11	Random Selection	Retail Investor (Average Assets, Modest Diversification, High Turnover)	200,000	0.0	20	75.0	15	150
12	Random Selection	Retail Investor (Average Assets, Modest Diversification, High Turnover)	200,000	0.0	30	75.0	23	230
13	Random Selection	Retail Investor (Average Assets, Low Divers., High Turnover) (Popular Stocks)	200,000	0.0	10	100.0	10	100
14	S&P 500	S&P 500 Index Fund	361,500,000	0.0	500	5.8	28	284

NOTE.— A "trade" is defined as a simultaneous (or near simultaneous) sale and purchase of a stock. Characterizations of "average," "median" or "low" with respect to prototypical investor type for Investors 7-13 refer to assets held and/or level of diversification and are based on Investment Company Institute/Securities Industry Association ("ICI/SIA") survey results. The characterizations are intended to reflect relative, not absolute behavior of individual investors. Similarly, turnover characterizations are based on the ICI/SIA survey, which finds that 60% of investors surveyed made no trades during the year in question, 22.8% made 1-5 trades during the year, 8.8% made 6-12 trades, and 8.4% made more than 13 trades. Level of diversification is characterized as "modest" if it conforms to prior research on minimum number of stocks to be held for proper diversification (20 or 30 stocks, depending on researcher and year of study). Investor 13 makes random buy decisions from a limited universe of "popular stocks." "Popular stocks" are defined as the 25 most actively traded stocks (excluding exchange traded funds) in any given year. Trades per year for Investor 14 reflect average (rounded) yearly trades. Trades by year in 1996-2005, respectively, are: 21; 27; 41; 41; 56; 30; 22; 9; 19; 18. Note: Last sale of 1999 is accompanied by buy at the beginning of 2000. Figures may not reflect number of additions/deletions to the index in any given year because trade total excludes deletions/additions that merely reflect a name change, merger or other reorganization that does not require a sale and subsequent purchase of stock. Three such trades are excluded from 1996 total.

38

Table 2
Simulation Results - All Frauds

Investor No.	Prototypical Investor Type	% Iterations Fraud Encountered	Of Those Affected by Fraud					
			% Net Gainers	% Net Losers	Avg. Trading Gain (Loss) ($)	Median Trading Gain (Loss) ($)	Low Trading Gain (Loss) ($)	High Trading Gain (Loss) ($)	Std. Dev. Trading Gain (Loss) ($)
1	Mutual Fund (Average)	99.9	46.0	54.0	(1,029,342)	(1,221,978)	(624,717,809)	827,231,679	45,862,944
2	Mutual Fund (Median)	97.4	43.8	56.2	(238,014)	(205,749)	(103,125,872)	107,493,535	7,328,497
3	Mutual Fund (Highly Diversified)	100.0	47.7	52.4	(1,848,759)	(1,267,290)	(405,356,965)	550,735,210	38,353,418
4	Mutual Fund (Highly Active Trader)	100.0	46.0	54.1	(1,527,510)	(920,408)	(330,961,389)	109,790,191	21,391,216
5	Mutual Fund (Low Diversification)	50.4	36.6	63.4	(6,085,689)	(1,713,294)	(515,806,549)	232,819,412	38,633,017
6	Mutual Fund (Infrequent Trader)	60.0	36.3	63.7	(3,578,987)	(901,463)	(826,827,933)	292,208,500	36,719,093
7	Retail Investor (Average Assets/ Diversification, Moderate Turnover)	23.6	26.1	73.9	(12,855)	(1,724)	(548,253)	193,502	58,668
8	Retail Investor (Average Assets/ Diversification, High Turnover)	50.9	25.3	74.7	(10,721)	(2,631)	(1,209,431)	804,033	68,247
9	Retail Investor (Median Assets/ Diversification, Moderate Turnover)	22.7	14.3	85.7	(4,920)	(1,354)	(95,677)	99,801	15,175
10	Retail Investor (Median Assets/ Diversification, High Turnover)	48.9	16.8	83.2	(5,306)	(1,155)	(472,991)	118,175	23,344
11	Retail Investor (Average Assets, Modest Diversification, High Turnover)	62.6	36.5	63.5	(1,837)	(616)	(350,409)	385,265	28,406
12	Retail Investor (Average Assets, Modest Diversification, High Turnover)	97.1	46.1	53.9	(1,968)	(696)	(494,105)	272,366	25,382
13	Retail Investor (Average Assets, Low Divers., High Turnover) (Popular Stocks)	99.9	40.7	59.3	(26,899)	(3,910)	(5,089,922)	1,590,390	213,435
14	S&P 500 Index Fund	N.M.	100.0	0.0	2,524,956	2,524,956	2,524,956	2,524,956	N.M.

39

Table 3
Simulation Results - All Frauds
Standardized $1 Million Initial Capital and No Cash Reserves

Investor No.	Prototypical Investor Type	% Iterations Fraud Encountered	Of Those Affected by Fraud					
			% Net Gainers	% Net Losers	Avg. Trading Gain (Loss) ($)	Median Trading Gain (Loss) ($)	Low Trading Gain (Loss) ($)	High Trading Gain (Loss) ($)	Std. Dev. Trading Gain (Loss) ($)
1	Mutual Fund (Average)	99.9	44.3	55.7	(2,384)	(3,667)	(835,302)	1,373,715	77,065
2	Mutual Fund (Median)	96.2	43.7	56.3	(4,612)	(2,204)	(1,306,138)	755,498	73,515
3	Mutual Fund (Highly Diversified)	100.0	46.7	53.3	(945)	(1,434)	(363,521)	634,021	34,616
4	Mutual Fund (Highly Active Trader)	100.0	44.9	55.1	(7,281)	(4,059)	(3,449,398)	1,184,642	108,199
5	Mutual Fund (Low Diversification)	49.9	35.8	64.2	(5,505)	(2,916)	(867,430)	1,540,297	104,533
6	Mutual Fund (Infrequent Trader)	60.1	38.9	61.1	(2,894)	(656)	(137,945)	148,016	26,244
7	Retail Investor (Average Assets/ Diversification, Moderate Turnover)	23.3	28.8	71.2	(26,291)	(6,974)	(1,284,920)	3,250,263	246,877
8	Retail Investor (Average Assets/ Diversification, High Turnover)	50.1	27.5	72.5	(68,380)	(12,778)	(6,641,623)	1,328,353	359,899
9	Retail Investor (Median Assets/ Diversification, Moderate Turnover)	23.6	16.1	83.9	(113,191)	(32,451)	(2,902,360)	2,996,790	367,307
10	Retail Investor (Median Assets/ Diversification, High Turnover)	49.1	18.2	81.8	(123,141)	(27,627)	(5,763,802)	2,416,410	453,157
11	Retail Investor (Average Assets, Modest Diversification, High Turnover)	63.6	36.2	63.8	(15,788)	(4,234)	(1,358,136)	1,131,518	112,139
12	Retail Investor (Average Assets, Modest Diversification, High Turnover)	83.8	41.9	58.1	(7,386)	(2,533)	(1,399,186)	711,629	95,043
13	Retail Investor (Average Assets, Low Divers., High Turnover) (Popular Stocks)	99.8	40.2	59.8	(115,194)	(21,443)	(21,871,853)	5,085,514	862,274
14	S&P 500 Index Fund	100.0	100.0	0.0	6,949	6,949	6,949	6,949	N.M.

40

Table 4
Simulation Results – Secondary Market Frauds Only

Investor No.	Prototypical Investor Type	% Iterations Fraud Encountered	Of Those Affected by Fraud					
			% Net Gainers	% Net Losers	Avg. Trading Gain (Loss) ($)	Median Trading Gain (Loss) ($)	Low Trading Gain (Loss) ($)	High Trading Gain (Loss) ($)	Std. Dev. Trading Gain (Loss) ($)
1	Mutual Fund (Average)	99.8	46.8	53.2	(168,798)	(1,011,641)	(624,717,809)	827,231,679	44,223,680
2	Mutual Fund (Median)	96.4	44.4	55.6	(177,297)	(170,484)	(101,324,107)	107,493,535	7,176,765
3	Mutual Fund (Highly Diversified)	100.0	48.4	51.6	(623,012)	(658,590)	(406,227,518)	550,982,132	37,197,829
4	Mutual Fund (Highly Active Trader)	100.0	47.0	53.1	(875,953)	(552,792)	(330,961,389)	113,342,113	20,529,699
5	Mutual Fund (Low Diversification)	47.7	36.9	63.1	(5,713,580)	(1,708,004)	(515,806,549)	151,200,117	35,923,461
6	Mutual Fund (Infrequent Trader)	57.6	37.2	62.8	(3,440,251)	(827,317)	(826,827,933)	292,334,329	36,323,030
7	Retail Investor (Average Assets/ Diversification, Moderate Turnover)	22.1	27.0	73.0	(10,849)	(1,708)	(548,253)	193,502	50,943
8	Retail Investor (Average Assets/ Diversification, High Turnover)	48.0	25.5	74.5	(9,023)	(2,657)	(1,209,431)	804,033	63,417
9	Retail Investor (Median Assets/ Diversification, Moderate Turnover)	21.1	14.7	85.3	(4,958)	(1,291)	(95,677)	99,801	15,451
10	Retail Investor (Median Assets/ Diversification, High Turnover)	45.2	16.3	83.7	(5,481)	(1,169)	(472,991)	118,175	24,015
11	Retail Investor (Average Assets, Modest Diversification, High Turnover)	60.3	36.7	63.3	(1,718)	(608)	(350,409)	385,265	28,348
12	Retail Investor (Average Assets, Modest Diversification, High Turnover)	96.4	47.0	53.0	(1,492)	(585)	(494,105)	271,511	24,390
13	Retail Investor (Average Assets, Low Divers., High Turnover) (Popular Stocks)	99.6	42.2	57.8	(25,148)	(3,436)	(5,089,922)	1,590,390	213,528
14	S&P 500 Index Fund	100.0	0.0	100.0	(847,930)	(847,930)	(847,930)	(847,930)	N.M.

41

Table 5
Simulation Results - Secondary Market Frauds Only
Standardized $1 Million Initial Capital and No Cash Reserves

Investor No.	Prototypical Investor Type	% Iterations Fraud Encountered	Of Those Affected by Fraud					
			% Net Gainers	% Net Losers	Avg. Trading Gain (Loss) ($)	Median Trading Gain (Loss) ($)	Low Trading Gain (Loss) ($)	High Trading Gain (Loss) ($)	Std. Dev. Trading Gain (Loss) ($)
1	Mutual Fund (Average)	99.9	45.3	54.7	(1,967)	(2,654)	(846,816)	1,373,886	74,509
2	Mutual Fund (Median)	95.3	43.9	56.1	(4,394)	(2,018)	(1,306,138)	754,778	72,956
3	Mutual Fund (Highly Diversified)	100.0	48.0	52.0	(254)	(841)	(358,410)	632,563	33,897
4	Mutual Fund (Highly Active Trader)	100.0	46.5	53.5	(5,222)	(2,675)	(3,448,292)	1,185,612	104,545
5	Mutual Fund (Low Diversification)	47.3	37.5	62.5	(3,221)	(2,252)	(867,430)	1,540,297	105,502
6	Mutual Fund (Infrequent Trader)	57.7	39.5	60.5	(2,935)	(687)	(137,945)	128,735	25,777
7	Retail Investor (Average Assets/ Diversification, Moderate Turnover)	21.6	29.9	70.1	(22,991)	(7,401)	(1,284,920)	3,250,263	248,097
8	Retail Investor (Average Assets/Diversification, High Turnover)	47.1	28.5	71.5	(65,630)	(11,700)	(6,641,623)	1,328,353	367,385
9	Retail Investor (Median Assets/ Diversification, Moderate Turnover)	22.2	16.0	84.0	(113,673)	(30,932)	(2,902,360)	2,996,790	380,520
10	Retail Investor (Median Assets/Diversification, High Turnover)	45.5	18.8	81.2	(120,719)	(26,464)	(5,763,802)	2,533,697	451,574
11	Retail Investor (Average Assets, Modest Diversification, High Turnover)	61.1	37.2	62.8	(14,644)	(3,778)	(1,358,136)	1,131,518	110,432
12	Retail Investor (Average Assets, Modest Diversification, High Turnover)	81.3	42.0	58.0	(6,913)	(2,106)	(1,399,186)	711,629	93,960
13	Retail Investor (Average Assets, Low Divers., High Turnover) (Popular Stocks)	99.4	40.7	59.3	(106,175)	(16,421)	(21,871,853)	5,078,309	859,821
14	S&P 500 Index Fund	100.0	0.0	100.0	(2,354)	(2,354)	(2,354)	(2,354)	N.M.

42

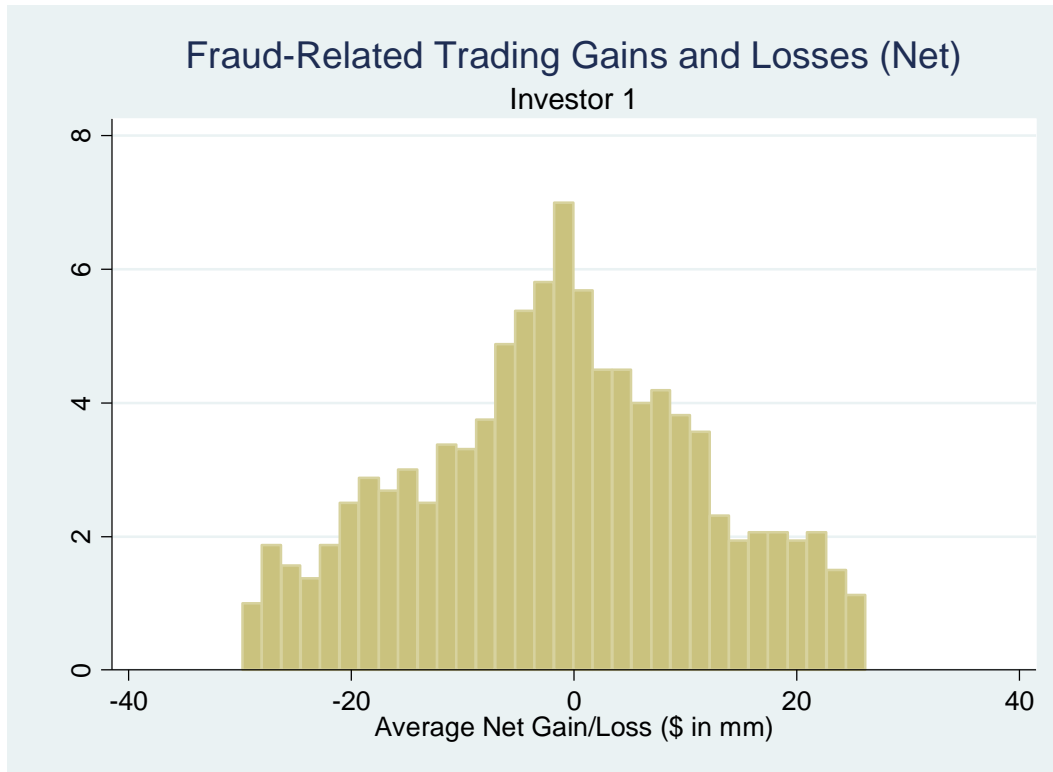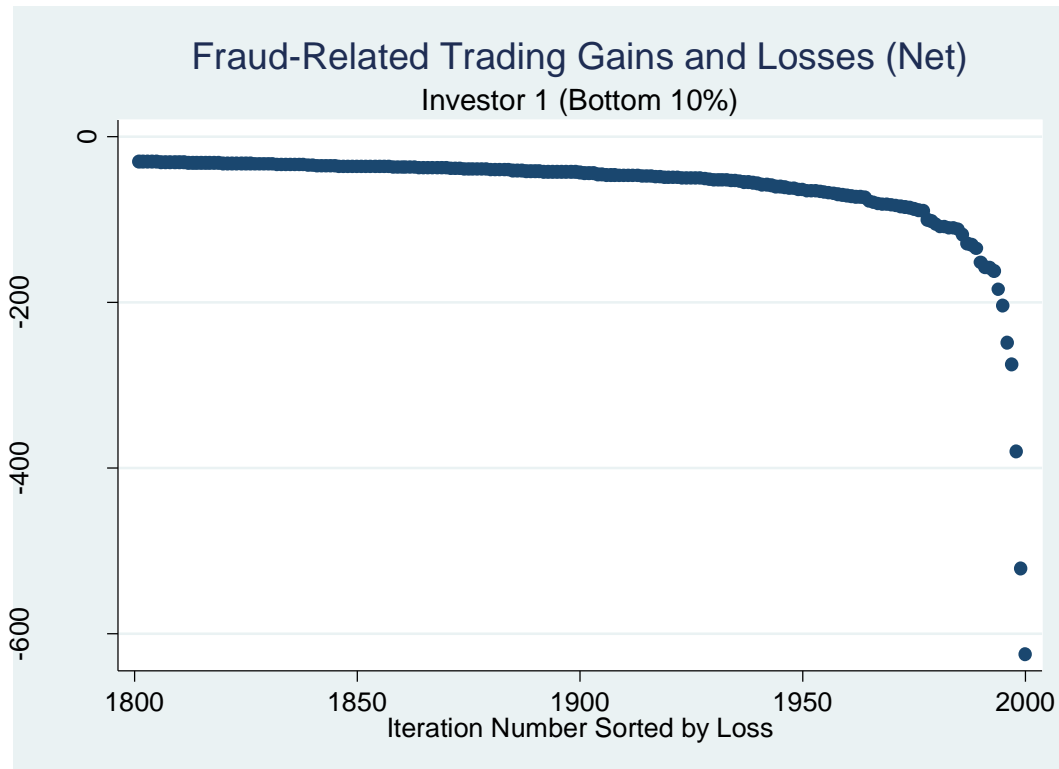Figure 1– Histogram of Middle 80% of Trials for Investor 1

Figure 2 – Scatter Plot of Bottom 10% of the Distribution for Investor 1

43

Appendix A
Description of Simulation Program and Sample Code

<u>Description</u>

The program consists of two main loops that ensure each investor performs the required actions. The outer loop controls the date of the program and will load the appropriate stock prices, stock offerings, and markers for fraudulent stocks. Within the inner loop, each investor's portfolio is subject to a check, to ensure that it does not hold a stock for a company that is out of business or going to be out of business on the next trading day. Next, a simple function is performed to determine how many actions an investor should perform with respect to buying or selling on that particular day. With this number calculated, a third loop is entered that performs the number of actions previously calculated. Each sale is followed by a check to ensure that no fraud was encountered and then a similar buy procedure is followed that also performs this check. Following these loops, although not shown, the net fraud encountered by each investor is calculated using the appropriate "fraud ribbon."

<u>Sample Code</u>

```
For i = 0 To UBound(days)
LoadDay()
  For j = 1 To NumInvestors
      StockGoneCheck = StockGone(j)
      If StockGoneCheck <> "" Then
          StocksGone = Split(StockGoneCheck, ",")
        FraudCheck()
          InvestorCapital(j) = InvestorCapital(j) + (NumShares * Stock2SellPrice)

        For k = 1 To goneBUY
          Stock2Buy = BuyStockSelector(intelS & intelV & intelR, j)
          Stock2BuyPrice = StockPrice(Stock2Buy)
          If Stock2BuyPrice = 0 Then GoTo buyRetry2
          If StockGone("*" & Stock2Buy & "/" & Stock2BuyPrice) <> "" Then GoTo buyRetry2
            NumShares = InvestorCapital(j) / (Stock2BuyPrice * goneBUY)
            FraudCheck()
            InvestorCapital(j) = AddStock(InvestorCapital(j), j, Stock2Buy,True)
        Next k
      End If
    End If
        tradesToday = DayTrader(TradeF(j), i, UBound(days), LastDay)
        TradeF(j) = TradeF(j) - tradesToday
      End If

  For k = 1 To tradesToday
    Stock2Buy = BuyStockSelector(intelS & intelV & intelR, j)
    Stock2BuyPrice = StockPrice(Stock2Buy)
    If Stock2BuyPrice = 0 Then GoTo buyRetry
    If StockGone("*" & Stock2Buy & "/" & Stock2BuyPrice) <> "" Then GoTo buyRetry

    Stock2Sell = SellStockSelector(intelS & intelV & intelR, j)
    Stock2SellPrice = StockPrice(Stock2Sell)
```

44

```
      FraudCheck()

      InvestorCapital(j) = InvestorCapital(j) + (NumShares * Stock2SellPrice)
      EventHolder = InvestorCapital(j)
      NumShares = Int((InvestorCapital(j) - (CashReservesA(j) * StartingCapitalA(j))) /
Stock2BuyPrice)
      InvestorCapital(j) = AddStock(InvestorCapital(j), j, Stock2Buy, CashReservesA(j) *
StartingCapitalA(j), True)
      FraudCheck()
    Next k

  Next j

If i = UBound(days) Then
   LastDay = True
End If

Next i
```

45

Appendix B
Morningstar Mutual Fund Search Parameters

     (% Bond = 0)
and (% Other = 0)
and (% Non-US Stock = 0)
and (% Cash not = 100)
and (Index Funds = No)
and (Enhanced Index Funds = No)
and (Fund of Funds = No)
and (Distinct Portfolio Only = Yes)

      Number of funds: 333

The table above provides the search parameters used in the Morningstar mutual fund database search on July 19, 2007 to generate the sample of actively managed domestic stock mutual funds for inclusion in the "mutual fund" sample. The sample only includes 100% domestic stock funds (that is, funds with no securities other than stock of U.S. corporations). I exclude index funds and funds of funds. To avoid double counting, I ensure that I have a sample of unique funds, rather than different share classes of the same fund by searching for distinct portfolios only.

46

Appendix C
Haigh Excerpt – Probability of Being Break-Even from Fraud

<u>**Excerpt from John Haigh, Taking Chances: Winning with Probability. Oxford University Press. 2003.**</u>

"[I]t is useful to know what to expect in a long sequence of repetitions, such as in tossing a fair coin. Here, the probability of Heads on any toss will be ½, but what do we expect to see over thousands or even millions of tosses? Will the proportion of Heads settle down to one-half? The short answer is 'Yes.' But that needs some elaboration.

Demanding exactly 500 000 H and 500 000 T in one million tosses is rather unreasonable. But if you ask for a fixed *percentage* on either side of the middle, the picture changes. Suppose you want the number of Heads to be between 49% and 51% of the total. In an experiment with just 100 tosses, this is asking for either 49, 50 or 51 Heads. If you repeated this experiment a large enough number of times, you would find the proportion of Heads fell in this narrow range about 24% of the time.

Increase the size of the experiment to 1000 tosses. Then anything from 490 to 510 is acceptable, and this will happen about 50% of the time. With 10 000 tosses, the permitted range is 4900 to 5100, and now the success rate is over 95%. With a million tosses, the Romanovs will regain absolute power in Russia before the proportion of heads falls outside the 49% to 51% range.

The same sort of thing holds true with even more stringent proportions: if you look for the proportion of Heads to fall between 49.9% and 50.1%, you are asking a lot with 1000 tosses (only 499 or 500 or 501 Heads fit [sic] the bill), but you would be astonished not to meet the requirement with ten million tosses. The proportion of Heads will, eventually, be as close to ½ as you wish. What you cannot expect is exact equality of Heads and Tails, or for the number of Heads always to be within any fixed number, say 20, of its average, when a coin is tossed a million times. In fact, the reverse holds; if you toss a fair coin often enough, the absolute difference in the numbers of Heads and Tails will eventually exceed any number you care to name. It is the *proportion* of Heads that stabilizes." (pp. 7-9)

"Plainly to have equal numbers of H and T in any sequence, the total number of tosses must be even. Let's dispose of one fallacious piece of folklore—that as you increase the number of tosses as 2, 4, 6, 8, 10,..., so equal numbers of H and T become more likely (another wrong version of the Law of Averages). This is not just wrong, it is wrong in a spectacular way—the more tosses in this sequence, the *less* likely are equal numbers of H and T!

It is not hard to show this must be so. For any (even) number of tosses, we count up how many outcomes there are altogether, and how many of them have equal numbers of H and T. The chance of equality is the ratio of the second to the first—e.g., 20/64 = 5/16 as we saw for six tosses...[63] [T]he first few answers are shown in Table 4.1. To make comparisons easy, the bottom row of the table gives all the probabilities as fractions with a common denominator. It is easy to see that these probabilities decrease as the number of tosses increases…For future reference, the general result [using Sterling's Formula] is

the chance of *n* H and *n* T in 2*n* tosses of a fair coin is very close to $1/\sqrt{n\pi}$

---

[63] In a prior section, the author calculates this figure. Six tosses of a fair coin yields 64 different possible outcomes (calculated as $2^6 = 64$). He then goes on to explain, "To find the probability of equal numbers of Heads and Tails, we have to count how many of these outcomes have exactly three Heads. To do this, we count how many different ways there are to select the positions of the three Hs…[T]here are $^6C_3 = 20$ ways of choosing where the Hs go, which means that the chance of equal numbers is 20/64 = 5/16."

47

especially for large numbers of tosses. This approximation always exaggerates the chance of equal numbers by a small amount." (pp. 54-56)

Table 4.1. For different numbers of tosses of a fair coin, the corresponding chances of equal numbers of Heads and Tails.

Number of Tosses	2	4	6	8	10
Chance of Equality	1/2	3/8	5/16	35/128	63/256
(Common Denominator)	128/256	96/256	80/256	70/256	63/256
[Percentage]*	[50%]	[37.5%]	[31.25%]	[27.34%]	[24.61%]

*Added by author

48

Appendix D
Simonoff Excerpt – Probability of Being Break-Even from Fraud

Excerpt from Jeffrey S. Simonoff, "The Binomial Distribution." 2010. Available at http://pages.stern.nyu.edu/~jsimonof/classes/1305/pdf/binomial.pdf

"Note that these properties of the binomial help debunk the idea of the "law of averages." Say you flip a fair coin 1000 times, and get 450 heads. What is the probability of heads on the next flip? Of course, it is .5, by independence. But what then of the "law of averages," which says that you should end up with about half heads and half tails, so you somehow have to "make up" the extra tails with extra heads? Let X be the number of heads in n flips of a fair coin, and let $\bar{p}$ = X/n be the observed proportion of heads. We know that E(X) = [(n)(p)] = (n)(.5) = n/2, and that V (X) = [(n)(p)(1-p)] = (n)(.5)(.5) = n/4, or SD(X) = $\sqrt{n}/2$.

Further, we know that

$$E(\bar{p}) = E(X/n) = E(X)/n = .5$$

and

$$V(\bar{p}) = V(X/n) = V(X)/n^2 = 1/(4n),$$

or SD($\bar{p}$) = $1/(2\sqrt{n})$.

Now, what about those 1000 flips of the coin? It is certainly true that the expected number of heads is half, or 500, and the standard deviation of the number of heads is $\sqrt{1000}/2$ = 15.8. What if we flipped the coin 2000 times? The expected number of heads would be half, or 1000, but the standard deviation would be $\sqrt{2000}/2$ = 22.4. That is, *as the number of flips gets larger, the variability around the expected number of heads gets larger, not smaller.* There is no law of averages "pushing" the number of coins back towards half of the total, since the variability of the number of heads **increases** with more flips, not decreases.

All is not lost, however. What is the expected proportion of heads in 1000 flips of the coin? It's *E($\bar{p}$)*, or .5. What is the standard deviation of that proportion? It's *SD($\bar{p}$)*, or .0158. What if there are 2000 flips? The expected observed proportion of heads is still .5, but the standard deviation of the observed proportion is .0112. That is, the observed proportion of heads gets closer and closer to the expected proportion as the number of binomial trials gets larger, in the sense that the observed proportion is centered around the expected proportion with prog[r]essively smaller standard deviation. This is not the "law of averages," but rather the Law of Large Numbers, and virtually any statistic satisfies a version of it: as the size of a random sample increases, the observed value of a statistic that can be written as an average of sample values gets closer and closer to the common expected value of each sample value in this probabilistic sense."

49

Appendix E
Feller Excerpt - Probability of Being Break-Even from Fraud

Excerpt from Feller, _Introduction to Probability Theory and its Applications_, 1950, pp. 238-239.

"*(a) Return to Equilibrium in Coin Tossing.* In a sequence of independent tossings of a symmetric coin let έ stand as an abbreviation for "the accumulated numbers of heads and tails are equal." For the time-honored bettor who at each trial loses or gains a unit amount, the occurrence of έ means that his accumulated gain is zero. Here it is clear that the occurrence of έ means a return to the initial situation, or, as we prefer to call it, a *return to equilibrium.* Clearly έ can only occur at an even number of trials. If u_n is the probability that έ occurs at the *n*th trial, then u1=u3=u5=u7=…=0; for έ to occur at the 2*n*th trial it is necessary and sufficient that the 2*n* trials produce *n* heads and *n* tails, and hence:

$$u_{2n} = \binom{2n}{n} \frac{1}{2^{2n}}$$ [NB: u_{2n} is the probability that the bettor is break even at the 2*n*th trial]

If the coin is tossed until έ occurs for the first time, we get the following sequences (arranged according to length): *HT, TH; HHTT, TTHH; HHHTTT, HHTHTT, TTTHHH,TTHTHH; HHHHTTTT, HHHTHTTT, HHHTTHTT, HHTHHTTT, HHTHTHTT, TTTTHHHH, TTTHTHHH, TTTHHTHH, TTHTTHHH,TTHTHTHH*; etc. If f_n is the probability that έ occurs *for the first time* at the *n*th trial, then obviously $f_n = 0$ whenever *n* is odd. An inspection of the above sequence shows that $f_2 = 1/2$ [50.0%], $f_4 = 1/8$ [12.5%], $f_6 = 1/16$ [6.25%], $f_8 = 5/128$ [3.9%]. A further enumeration leads to the values f_{10}=7/256 [2.7%], f_{12}=21/1024 [2.1%], f_{14}=33/2048 [1.6%],…"

This excerpt from Feller (1950) describes the phenomenon by which the probability of being even at the 2*n*th trial goes to zero as *n* goes to infinity. Thus, the longer one plays (that is, the more trials), the less likely one is to break even.

50