How Widespread and Predictable is Stock Broker Misconduct?

by
Craig McCann, PhD, CFA, Chuan Qin, PhD
and Mike Yan, PhD, CFA, FRM [1]

In this paper we reconcile widely diverging recent estimates of broker misconduct. Qureshi and Sokobin report that 1.3% of current and past brokers are associated with awards or settlements in excess of a threshold amount. [2] Egan, Matvos, and Seru find that 7.8% of current and former brokers have financial misconduct disclosures including customer complaints, awards, and settlements.[3]

We replicate and extend the analysis of broker misconduct in these studies. Qureshi and Sokobin arrive at their low estimate by excluding 85% of all brokers, including those brokers most likely to have engaged in misconduct. Applying Qureshi and Sokobin's restrictive definition of potential misconduct to all brokers, we find that misconduct is much more widespread.

We also evaluate Qureshi and Sokobin's claim that its BrokerCheck website provides helpful information to investors seeking to avoid bad brokers and answer the question posed by Egan, Matvos, and Seru: If BrokerCheck data can identify broker misconduct, why don't investors use that data to protect themselves? We find that BrokerCheck is worthless in its current hobbled form, but that it could easily be modified so that market forces might substantially reduce broker misconduct.

1. Introduction

FINRA is a self-regulatory organization tasked with policing registered representatives of brokerage firms ("brokers"). It maintains a database of investor complaints and disciplinary and employment history for over 1,200,000 current and past brokers and publishes some of this information on its BrokerCheck website.

 Craig McCann can be reached at 703-246-9381 or at CraigMcCann@SLCG.com. Chuan Qin can be reached at 703-539-6778 or ChuanQin@slcg.com Mike Yan can be reached at 703-539-6780 or MikeYan@slcg.com.

[2] Qureshi, H. and Sokobin, J. (2015), "Do Investors Have Valuable Information About Brokers?" http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2652535

[3] Egan, M., Matvos, G. and Seru, A. (2016), "The Market for Financial Adviser Misconduct." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2739170

Two recently published research reports have reported wildly different estimates of the extent of investment fraud perpetrated by brokers based on the same BrokerCheck data. These reports also evaluated the potential of BrokerCheck data to predict future broker misconduct, thereby allowing investors to discern good brokers from bad brokers.

Qureshi and Sokobin [2015] (hereafter "Qureshi and Sokobin") analyze publicly and non-publicly available BrokerCheck data on 181,133 brokers and find that 2,349 of the brokers, or 1.3% of the total studied, had at least one customer complaint during the period from 2000 to 2013 which resulted in an award or settlement above a $10,000 threshold before May 18, 2009 and above a $15,000 threshold thereafter. Thus, it appears from Qureshi and Sokobin, that misconduct is rare in the brokerage industry.

Egan, Matvos and Seru [2016] (hereafter "Egan, Matvos and Seru") using the BrokerCheck data covering a slightly different time period - 2005 to 2015 - found that that 7.8% of brokers have misconduct disclosures on their record and that brokers remain in the industry despite repeated misconduct. While Egan, Matvos and Seru use a more expansive definition of broker misconduct than Qureshi and Sokobin, this difference doesn't explain the dramatically different assessment of brokerage industry misconduct.

Qureshi and Sokobin only report the incidence of awards and settlements by 15% of brokers who were registered between 2000 and 2015. The 85% who were also registered during this period and which are excluded by Qureshi and Sokobin have a much higher incidence of customer awards or settlements than the brokers they included. Including all brokers increases the number of brokers with disclosed awards or settlements more than ten-fold, from 2,349 to 27,494.

Qureshi and Sokobin investigated whether BrokerCheck provides investors with information that can predict future investor harm and concluded that information available on BrokerCheck significantly predicts future meritorious investor claims.

> Overall, our results suggest that BrokerCheck provides valuable information to investors, thereby allowing them to discriminate between brokers with a high propensity for investor harm from other brokers. [p.4]

Egan, Matvos and Seru also find information on past broker misconduct can be used to predict future misconduct. They find stock brokers with recent customer complaints are more likely to be terminated by their employer, subsequently have longer spells of unemployment, and are thereafter re-employed at lower compensation and by less prestigious firms than brokers who do not have customer complaints. Egan, Matvos and Seru's results confirm that, rather than weeding bad brokers out of the industry, the regulatory environment and labor market sifts bad brokers down the quality ladder over time into brokerage firms with loose hiring practices and lax compliance ethics and that these bad brokerage firms specialize in preying on unsophisticated investors.[4]

We use the same BrokerCheck data as Qureshi and Sokobin and Egan, Matvos and Seru to reassess whether BrokerCheck provides information to retail investors that helps them avoid bad brokers. We find that the BrokerCheck data does not help investors protect themselves because BrokerCheck cannot, in its current hobbled form, be used to discern good brokers from bad brokers as claimed by Qureshi and Sokobin.

We fit two regression models, a probit model and a random forest model, to the BrokerCheck data and evaluate the models' predictive performance. Random forest models generally achieve much better predictive performance than probit models, demonstrating the importance of selecting appropriate statistical models to make the most of the vast amount of BrokerCheck data.

We consider both characteristics of the individual broker (available from the BrokerCheck website, one broker at a time) and characteristics of those working with the broker at the same brokerage firm, (calculated using individual broker characteristics) as our models' input variables. While the models using only individual broker characteristics have power to discriminate brokers with a high propensity for investor harm from others brokers, adding coworker characteristics significantly improves our models' predictive performance. As we explain below, the results of our analysis – and of the analyses performed by Qureshi and Sokobin and Egan, Matvos and Seru – do <u>not</u> support the conclusion the BrokerCheck provides <u>any</u> useful information to investors.

[4] Dimmock, Gerken and Graham [2015], in a related study find that financial fraud is contagious. They find that a broker's propensity to commit financial fraud is significantly influenced by his or her co-workers' propensity to commit fraud after controlling for firm culture, branch atmosphere, market conditions and state regulatory environment.

FINRA promotes a perception of BrokerCheck that is a classic example of the fallacy of composition. It suggests that since information on each individual of the 1.2 million brokers is accessible, the information on all 1.2 million brokers is accessible. FINRA actually goes to great lengths to make information which is ostensibly public, effectively non-public. FINRA could provide investors with the results of statistical modeling of all the BrokerCheck data on which Qureshi and Sokobin base their study rather than the infinitesimal portion of the data it currently provides retail investors. While this step would greatly enhance the usefulness of BrokerCheck, a much simpler solution is obvious: FINRA should simply make BrokerCheck information truly publicly available and allow the disinfecting power of sunshine to reduce broker misconduct.

The remainder of this paper is organized as follows. In Section 2, we summarize and reconcile the estimates of potential brokerage misconduct reported in Qureshi and Sokobin (1.3%) and in Egan, Matvos and Seru (7.8%). In Section 3, we replicate the main Qureshi and Sokobin results using data on 1.2 million brokers, downloaded one broker at a time from the BrokerCheck website. In Section 4, we apply a more sophisticated statistical technique, random forests, to the BrokerCheck data and demonstrate that BrokerCheck data could be even more useful than suggested by Qureshi and Sokobin if the data were truly made publicly available. In Section 5, we conclude with an explanation of why BrokerCheck data in its current form is virtually useless to investors trying to protect themselves from bad brokers and how it could be dramatically improved at little or no cost.

2. How Prevalent is Broker Fraud?

Brokers and investment advisers make recommendations and take orders. While the mix of activities varies from relationship to relationship, a broker or investment adviser can engage in misconduct that harms their customers. Some instances of broker misconduct are followed by customer complaints, arbitration filings or both. These customer complaints or arbitration filings may allege that unsuitable recommendations were made, important risks were not disclosed, accounts were churned or some other bad acts occurred. Broker financial misconduct might not directly involve a retail investor. For instance, a broker might have unsatisfied liens or personal bankruptcies which reflect on the broker’s fitness to manage or guide other people's investments and, as an empirical matter, help predict future customer complaints, arbitration filings.

Additionally, some brokers are disciplined by regulators such as the Securities and Exchange Commission, FINRA, state securities regulators, and state Attorneys General. The conduct underlying these regulatory actions may have already been subject to customer complaints or arbitration filings but regulators have the ability to enforce larger systemic remedies on brokerage firms than investors can accomplish by filing individual complaints.

FINRA maintains a database of registration, employment, complaint and disciplinary history for each brokerage firm and broker, the Central Registration Depository, or CRD. The CRD includes each broker's involvement in customer disputes, financial, disciplinary and criminal events, employment history, and qualifications. FINRA makes a portion of the information in the CRD public through its BrokerCheck website.

a. Qureshi and Sokobin [2015]

Qureshi and Sokobin analyzed BrokerCheck information on 181,133 brokers registered with FINRA between 2000 and 2013 and found only 2,349, or 1.3%, of these brokers have been associated with an award or settlement above a threshold dollar amount.

Qureshi and Sokobin analyze the BrokerCheck data only for brokers first registered in 2000 or later. Thus, a broker first registered in 1998 and still active in 2008 is excluded by Qureshi and Sokobin. Of the 552,016 brokers with at least one current state registration in BrokerCheck on December 31, 2015, 57% or 314,652 were first registered in 2000 or later and 43% or 237,364 were first registered sometime earlier. Thus, Qureshi and Sokobin excluded nearly half of the currently registered brokers from their study.

Qureshi and Sokobin also only include brokers that were registered with four or more states for at least half of their careers because such brokers are asserted to be more likely to have client interactions. Of the 552,016 brokers with at least one current state registration in BrokerCheck as of December 31, 2015, only 53.6% are registered with four or more states. Combined, the two restrictions on the sample imposed by Qureshi and

Sokobin limit their sample to only 181,133 brokers – only 15% of the 1.2 million brokers on BrokerCheck.

Qureshi and Sokobin also have a highly restrictive assumption about what constitutes an investor harm event. Qureshi and Sokobin define the initial filing of a grievance that subsequently results in an arbitration award in favor of the customer or in a settlement in excess of $10,000 prior to May 18, 2009 and in excess of $15,000 thereafter as an investor harm event. Their definition of an investor harm event assumes that settlements below these thresholds do not evidence any investor harm but are in fact entered into by brokerage firms to avoid further litigation costs. This assumption is overly restrictive. Many meritorious claims are not brought because the potential recovery is too small and too uncertain to warrant investors and their attorneys expending the effort to prosecute a case or because the investor does not know about the opportunities for redress. Also, low settlements and awards likely reflect the low wealth of these investors not the lack of merit of their claims.

b. Egan, Matvos and Seru [2016]

Egan, Matvos and Seru analyze BrokerCheck data for all 1.2 million brokers registered at some point in time between 2005 and 2015 without regard for when the broker was first registered and for how many states he or she has been registered with. They include as misconduct disclosures arbitration filings resulting in awards and lower settlements and other reported events as indicative of broker misconduct.

Egan, Matvos and Seru using BrokerCheck data from 2005 to 2015 find that 46,900 currently registered brokers have misconduct disclosures and nearly as many brokers no longer registered have disclosed misconduct compared to the 2,349 current and past brokers Qureshi and Sokobin find to have been associated with awards and settlements above their thresholds. They also find that misconduct in the brokerage industry is persistent; in any given year 0.60% of active brokers report a misconduct disclosure in the current year and 7.8% have a misconduct disclosure at some point in their career. That 13 times as many brokers have a misconduct disclosure than on average make a disclosure in any given year means brokers with misconduct disclosures remain in the industry rather than being weeded out by regulators or market forces. That 1.62% of brokers have some more broadly defined financial and disciplinary disclosure in any given year and 12.7% have such disclosures at some point in their career. [5]

Egan, Matvos and Seru find that broker misconduct can be predicted by disclosures of the broker's past misconduct. They find 38% of brokers that engaged in misconduct had previous misconduct disclosures. That is, brokers who engage in misconduct are not drawn randomly from brokers with clean or checkered pasts. They document that brokers with a misconduct disclosure at some point in their career previously are approximately five times as likely to have a misconduct disclosure in the current year. They find recidivist rate in the first year after a misconduct disclosure is nearly 20 times the average rate of misconduct and remains more than five times the average rate five years after the most recent misconduct disclosure.

Egan, Matvos and Seru find that brokers with recent customer complaints are more likely to be terminated by their employer, have longer spells of unemployment, and are re-employed at lower compensation than brokers who do not have customer complaints. Their results confirm that, rather than weeding bad brokers out of the industry, the regulatory environment and labor market sifts bad brokers down the quality ladder over time into brokerage firms with loose hiring practices and compliance ethics. Supplementing the BrokerCheck data with Census Bureau data, they find these bad brokerage firms which accumulate bad brokers specialize in preying on unsophisticated investors.

c. Incidence of Bad Brokers is between 3 and 4 Times FINRA's Estimate

Qureshi and Sokobin report that only 1.3% of brokers had been associated with an award or a settlement in excess of the $10,000/$15,000 thresholds described above. Adopting Qureshi and Sokobin's definition of an investor harm event, but including the brokers excluded by Qureshi and Sokobin results in a much higher estimate of the prevalence of bad brokers.

[5] Egan, Matvos and Seru group six CRD disclosures in FINRA classifications Customer Dispute-Settled, Regulatory-Final, Employment Separation After Allegations, Customer Dispute-Award/Judgment and Civil-Final as *a priori* "misconduct" disclosures. They group disclosures in the remaining 17 CRD categories, including Customer Dispute-Pending, are grouped treated as "Other" disclosures.

Table 1 reports the number of brokers, number of investor harm events and the number of brokers associated with investor harm events by number of state registrations as of December 31, 2015. The BrokerCheck database contains 27,494 current and former broker that have customer complaints resulting in settlements or awards meeting the definition of investor harm used by Qureshi and Sokobin. 14,351 of these 27,494 brokers with a disclosed investor harm event are currently registered with four or more states. Qureshi and Sokobin's conclusion that only 2,349 current and former brokers registered with at least 4 states during half of their career had investor harm events is only 8.5% (i.e. 2,349 ÷ 27,494) of the brokers who meet their definition investor harm events in the BrokerCheck database.

Table 1: Investor Harm Events by State Registrations as of December 31, 2015

# of State Registrations	Brokers	Investor Harm Events	Brokers Associated with Investor Harm	% Bad Brokers
0	648,657	19,464	10,676	1.65%
1	166,567	1,407	1,012	0.61%
2	57,893	1,055	767	1.32%
3	31,786	894	688	2.16%
>= 4	295,770	20,010	14,351	4.85%
Total	1,200,673	42,830	27,494	2.29%

The very low incidence of investor harm reported by Qureshi and Sokobin is primarily the result of their exclusion of brokers who were first registered before 2000. Qureshi and Sokobin excluded these brokers because the pre-2000 data available in electronic format was converted from a paper-based legacy system and may not completely reliable. This might justify excluding the brokers first registered before 2000 from the regression analysis which includes length of time in the industry but does not justify excluding these brokers – 48.6% of all brokers active in the 2000-2015 period – from the estimation of how widespread broker misconduct was in 2000-2015. Since the likelihood a broker will have a disclosed settlement or award increases with time in the industry, by excluding brokers first registered before 2000 Qureshi and Sokobin significantly understates the incidence of bad brokers.

Table 2 reports the number of brokers and the number of brokers associated with investor harm events by number of state registrations for brokers first registered in 1999 or earlier and in 2000 or later. Our dataset only includes the number of state registrations held by each broker as of December 31, 2015, not the number of state registrations throughout a broker's career used by Qureshi and Sokobin so we can not exactly replicate their sample.[6] Nonetheless, the left panel of Table 2 closely tracks Qureshi and Sokobin's results. They found that 1.30% of the 181,133 current and previously licensed brokers first registered after 1999 and with 4 or more state registrations for more than half their career were associated with a customer award or settlement above a dollar threshold. Consistent with their estimate, we find 1,943, or 1.35%, of the 144,178 brokers first registered after 1999 with 4 or more current state registrations have been associated with a customer award or settlement above a dollar threshold.

The currently registered brokers with four or more state registrations are nearly evenly divided between 151,592 brokers first registered before January 1, 2000 and 144,178 brokers first registered after January 1, 2000. 12,408 (8.19%) of the brokers first registered before 2000 and currently registered with four or more state have at least one investor harm event compared to only 1,943 (1.35%) of the brokers first registered after 1999. Thus, the brokers excluded by Qureshi and Sokobin are approximately six times as likely to have been associated misconduct disclosures (8.19% vs. 1.35%).

Table 2: Investor Harm Events of 2000-2014 by State Registrations as of December 31, 2015

	First Registered In or After 2000 (included by FINRA)		**First Registered Before 2000** (excluded by FINRA)	
# of State Registrations	Brokers	Brokers Associated with Investor Harm	Brokers	Brokers Associated with Investor Harm
0	436,583	1,847 (0.42%)	212,074	8,829 (4.16%)
1	116,888	235 (0.20%)	49,679	777 (1.56%)
2	35,599	149 (0.42%)	22,294	618 (2.77%)
3	17,987	139 (0.77%)	13,799	549 (3.98%)
>= 4	144,178	1,943 (1.35%)	151,592	12,408 (8.19%)
Total	751,235	4,313 (0.57%)	449,438	23,181 (5.16%)

In total, 4,313 brokers (0.57%) of the 751,235 brokers first registered after 1999 have reported a settlement or award in excess of Qureshi and Sokobin's thresholds. On the other hand, 23,181 brokers (5.16%) of the 449,438 brokers first registered before 2000 have reported a claim in 2000 or later which resulted in an award or settlement above Qureshi and Sokobin's threshold amounts. Brokers registered before 2000 are therefore

[6] It appears that Egan, Matvos and Seru extracted current information from the BrokerCheck website in a similar manner to our method. If FINRA made the BrokerCheck data truly publicly available, important research could be much more easily performed to the benefit of investors.

nine times as likely to have reported a claim in the 2000 to 2014 period as brokers first registered after 1999.

As Table 2 shows, the brokers excluded by Qureshi and Sokobin are between six and ten times as likely to have had a settlement or award in excess of the FINRA thresholds as those FINRA included in its study, regardless of the number of state registrations the broker had as of December 31, 2015.

Figure 1(a) presents the numbers of active brokers in each year from 2000 to 2014 who were first registered before 2000 and those who were first registered in 2000 or later. The number of active brokers first registered in 2000 or later exceeded the number of those first registered before 2000 starting in 2008.

Figure 1(a): Number of Brokers by Year

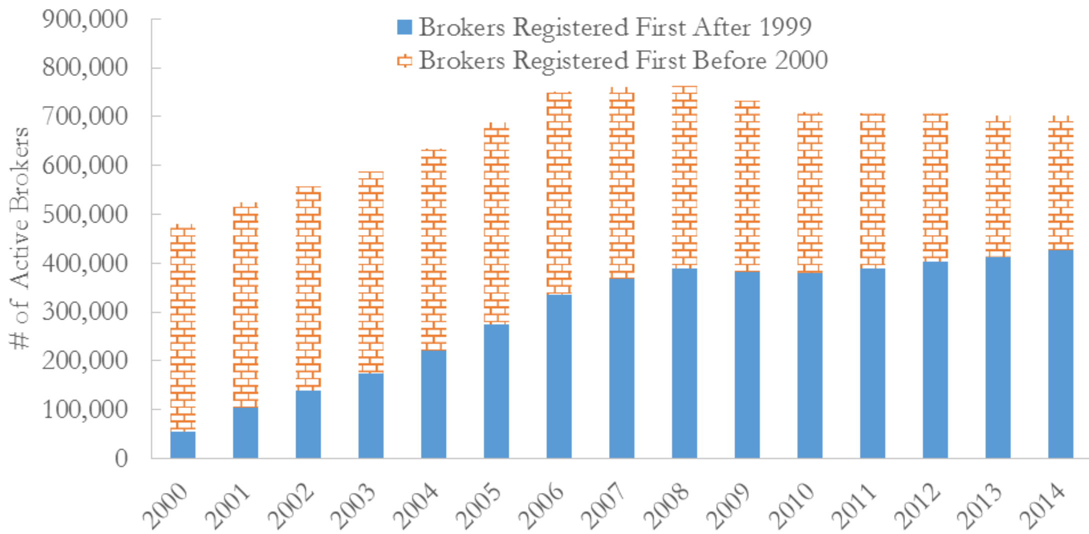
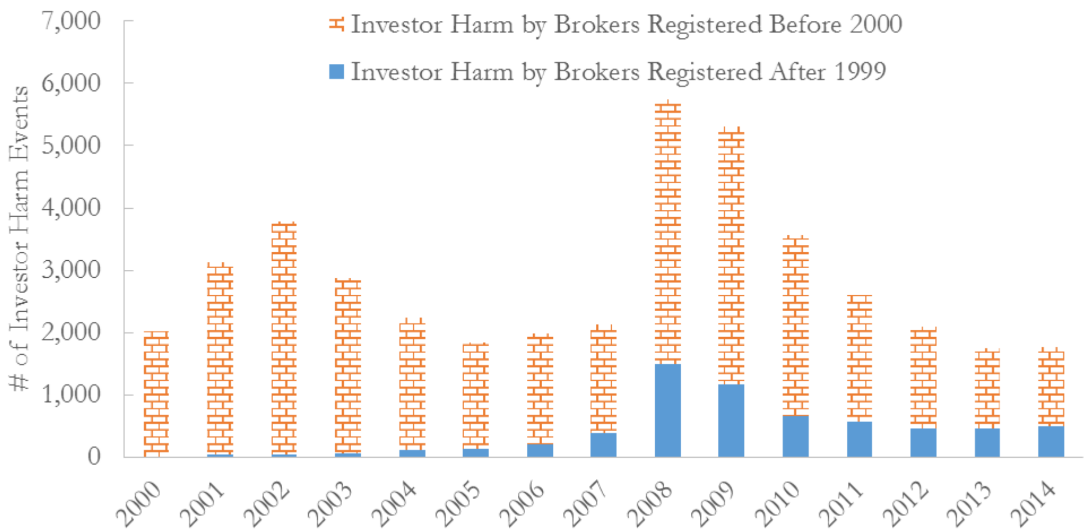


Figure 1(b) contrasts the numbers of investor harm events associated with brokers in these two groups each year. Much fewer of the investor harm events over the entire period of 2000-2014 were attributable to brokers first registered in 2000 or later than were associated with brokers first registered before 2000. This is not simply as result of brokers first registered earlier having a longer time period to accumulate customer complaints. Even though the brokers registered in 2008 are roughly equally divided between those first registered before 2000 and those first registered in 2000 or later, the brokers registered first in the earlier time account for 75% of the investor harm events in 2008.

Figure 1(b): Number of Investor Harm Events by Year



3. Can Broker Misconduct Be Predicted

In this section, we replicate the Qureshi and Sokobin results subject to some data limitations. Qureshi and Sokobin use data on bankruptcies within 10 years prior to each year's observances of broker misconduct. Since personal bankruptcies are only available on BrokerCheck for 10 years, we can't observe bankruptcies prior to 2005 in the data we observe as of December 31, 2015. Also, Qureshi and Sokobin limit their sample to brokers registered in more than three states for at least half their career. The BrokerCheck website only lists current state registrations so we approximate this filter by requiring brokers to be currently registered in four or more states. As a result, we exclude brokers who are no longer in the industry from our analysis in this section of the paper. Also, Qureshi and Sokobin use the broker's gender as an explanatory variable even though it gender is not one of the data items available on BrokerCheck. As we show below, these and other data limitations do not seem to matter; we are able to closely replicate Qureshi and Sokobin's results.

a. Data

As we explain more fully below, FINRA does not meaningfully make publicly available the BrokerCheck data it is required to make publicly available. With great effort over several weeks in early January 2016, we downloaded all data on the BrokerCheck website. The BrokerCheck website included information on 616,243 currently registered and 584,430 further formerly registered brokers who were not currently licensed as of December 31, 2015. To conform closely to the Qureshi and Sokobin sample, we then limited the sample to brokers first registered with FINRA after 1999 and who were registered with at least four states on December 31, 2015.

We aggregate the characteristics (disclosures, employment, qualifications, etc.) of each broker in each calendar year of the broker's tenure over the 2000-2014 period. This gives rise to an annual panel of broker-year observations spanning 15 years. We do not include information customer claims filed in 2015 and 2016 because Qureshi and Sokobin's outcome-based-filter requires enough time to pass after the filing of an arbitration claim to observe whether it resulted in an award or settlement in excess of their thresholds. For a disclosure event associated with more than one date, we assign the event with the earliest date when the underlying event was reported. For example, if a customer complaint was filed in August 2013 and arbitrated or settled in January 2015, we add the complaint to the 2013 panel. Thus, our sample is an annual panel dataset including the complete career history through 2014 of all the brokers who first registered with FINRA no earlier than 2000 and were registered with four or more states on December 31, 2015. There are 144,178 unique brokers and 1,163,927 broker-year observations in our sample.

b. Investor Harm and Model Features

We predict the occurrence of investor harm events in each year by building statistical models on BrokerCheck historical data. The model inputs, called *features*, are constructed from the raw data to reflect the characteristics of each broker and the brokerage firms where they were employed. We expect these features to contain useful information for predicting investor harm.

Customer complaints filed against a broker may result in a settlement or award, or remain unresolved, or they may be denied. Following Qureshi and Sokobin, we assume that arbitration filings that fail to lead to an award or to a settlement above $10,000 before May 18, 2009 and above $15,000 thereafter do not reflect investor harm or broker misconduct. The time when the investor harm event occurs is approximated by the arbitration filing year. We associate each year in a broker's career with an indicator variable which equals 1 if the broker discloses an investor harm event that year and 0 otherwise.

Table 3 summarizes the distribution of investor harm events during the 2000-2014 period. A small portion of brokers (1.35%) in our sample are associated with investor harm as defined in the FINRA study. The majority of the brokers associated with investor harm only had one complaint that resulted in an award or settlement above the threshold.

Table 3: Summary of Investor Harm Events

Investor Harm Events	Brokers Associated with Investor Harm Events
1	1,627
2	226
3 or more	90
Total	1943

We are interested in predicting the occurrence of investor harm events associated with a broker in a given year using BrokerCheck information prior to that year. The disclosure history of a broker to have a significant power to predict future investor harm caused by the broker if past offenders are more likely to commit similar offenses in the future. The six broker disclosure features and five qualifications and employment features listed in Table 4 are compiled based on the BrokerCheck data in each year of each broker's career.

Table 4: Broker Disclosure Features

Features	Symbol	Description of Characteristic
Customer Dispute Settlements and Awards	SA	Number of costumer disputes that led to a settlement or award against the broker above $10,000/$15,000, from first registration to the year under consideration
All Customer Disputes	CD	Total number of costumer disputes against the broker regardless of status, from first registration to the year under consideration
Disclosed Judgements and Liens	JUDG	Total number of judgments and liens that are not satisfied until the year under consideration

Disclosed Bankruptcy Disclosures	BKCY	Total number of bankruptcies and compromises that occurred within the past 10 years of the year under consideration
Disclosed Disciplinary Events	DPRY	Total number of regulatory actions, investigations, and employment separation after allegations available on BrokerCheck through the year under consideration
Criminal Events	CRIM	Total number of criminal disclosures through the year under consideration
Exams Passed	EXAM	Number of exams (S6, S7, S63, S66) passed through the year under consideration
Association with Expelled Firm	EXPEL	An indicator that equals 1 if the broker had been registered, by the year under consideration, with any firm that has been expelled from the industry
Number of Prior Employers	PREMPL	Number of firms the broker worked at and separated from by the year under consideration
Number of Employment Years	EMPLYR	Number of years registered as a broker until the year under consideration
Dual Registration	DUAL	An indicator that equals 1 for brokers registered with SEC as an investment advisor

Table 5 presents the average values of broker features listed in Table 4 for brokers subsequently associated with investor harm events and for brokers not subsequently associated with investor harm events. The p-values for the two-sample t-tests suggest that brokers associated with investor harm have a higher average number of past customer disputes that led to an award or settlement above the thresholds, judgments and liens, disciplinary events, and criminal events. On average, the brokers associated with investor harm had passed more exams and had more previous employers and a longer registration history.[7] Brokers associated with investor harm events are also more likely to be SEC-registered investment advisors and are more likely to have been previously affiliated with an expelled firm. The only feature that does not appear to be statistically significantly different across the two subsets of brokers is the incidence of prior personal bankruptcies.

[7] These differences in means across the subsets of brokers are interesting but should not be overly interpreted. For example, brokers associated with investor harm events have passed more exams on average than brokers who are not associated with investor harm. This difference is statistically significant due to the very large sample size but it may not be meaningful. The brokers associated with investor harm on average have passed only 5.6% more exams than other brokers but 21 times as many have been associated with prior settlements and awards. Also, as we show below in the regression analysis, after controlling for other differences across brokers, differences in the number of exams passed provides no useful information for predicting which brokers will be associated with an investor harm event.

Table 5: Summary of Broker Features. The first column records average for all brokers, while the second and the third columns record averages for brokers associated with investor harm and those without such association, respectively. The fourth column reports the difference between the second and the third columns. The last column reports the p-value from a two-sample t-test on the characteristics of brokers with and without investor harm. The symbol *** denotes significance at the 0.001 level.

Features	All Brokers	Brokers Associated with Investor Harm	Brokers Not Associated with Investor Harm	Difference	p-values
SA	0.0100	0.1800	0.0097	0.1703	0.0000***
CD	0.0369	0.4308	0.0360	0.3948	0.0000***
JUDG	0.0080	0.0415	0.0079	0.0336	0.0003***
BKCY	0.0223	0.0310	0.0222	0.0088	0.1939
DPRY	0.0060	0.0244	0.0060	0.0184	0.0000***
CRIM	0.0226	0.0449	0.0226	0.0223	0.0002***
EXAM	2.0854	2.1996	2.0852	0.1144	0.0000***
EXPEL	0.0083	0.0530	0.0082	0.0448	0.0000***
PREMPL	0.8742	1.4255	0.8731	0.5524	0.0000***
EMPLYR	5.4154	7.1533	5.4119	1.7414	0.0000***
DUAL	0.6865	0.8434	0.6862	0.1572	0.0000***

Coworker Features

The previous studies ([2], [3] and [4]) all found that firm culture and coworker misconduct influenced on the amount of fraud committed by individual brokers. For each year in a broker's career we construct the eight coworker characteristics listed in Table 6 which describe the disclosure and employment profile of the broker's coworkers.

Table 6: Summary of Coworker Features

Feature	Symbol	Description of Feature
Harm Associated with Coworkers	HAC	Average number of investor harm events per broker for all other brokers at the same firm, averaged over the entire career of the brokers and across all brokers at the firm the broker is employed by in the year under consideration
Customer Disputes Associated with Coworkers	CDAC	Average number of customer disputes against each broker for all other brokers at the same firm, averaged over the entire career of the brokers and across all firms the broker is employed by in the year under consideration
Average Number of Coworkers	CONUM	Average number of all other brokers at the same firm, averaged over all firms the broker is employed by in the year under consideration

Coworker Disclosed Judgements and Liens	COJUDG	Average number of unsatisfied judgments and liens through the year under consideration per broker for all other brokers at the same firm, averaged over the entire career of the brokers and across all firms the broker is employed by in the year under consideration
Coworker Disclosed Disciplinary Events	CODPRY	Average number of disclosed disciplinary events per broker for all other brokers at the same firm, averaged over the entire broker career and across all firms the broker is employed by in the year under consideration
Coworker Criminal Events	COCRIM	Average number of criminal disclosures per broker for all other brokers at the same firm, averaged over the entire broker career and across all firms the broker is employed by in the year under consideration
Coworker Affiliation with Expelled Firm	COEXPEL	Percentage of coworkers that were previously registered with an expelled firm, averaged over all firms the broker is employed by in the year under consideration
Average Number of Prior Employers for Coworkers	COPREMPL	Average number of prior employers through the year under consideration per broker for all other brokers at the same firm, averaged over all firms the broker is employed by in the year under consideration
Average Number of Employment Years for Coworkers	COEMPLYR	Average number of years registered with FINRA through the year under consideration per broker for all other brokers at the same firm, averaged over all firms the broker is employed by in the year under consideration

Table 7 presents average values of coworker features for brokers associated with investor harm and for brokers not associated with investor harm in the current year. The extremely small p-values reflect the significant difference in past investor harm or customer dispute history between the coworkers of the brokers associated with investor harm and the coworkers of those without such association. Brokers associated with investor harm are more likely to have more coworkers with past judgements and liens, disciplinary events, criminal charges, and prior affiliations with expelled firms than brokers not associated with investor harm. The coworkers of brokers associated with investor harm are also more likely to have more previous employers and a longer employment history than the coworkers of brokers not associated with investor harm. This indicates that the coworker features listed in Table 6 may contain valuable information for distinguishing brokers likely to be associated with future investor harm from other brokers and therefore should be incorporated in statistical models that predict investor harm.

Table 7: Differences in Coworker Features. The symbols ***, ** and * denote significance levels at 0.001, 0.01 and 0.05, respectively.

Features	All Brokers	Brokers Associated with Investor Harm	Brokers Not Associated with Investor Harm	Difference	p-values
HAC	0.0012	0.0031	0.0012	0.0019	0.0000***
CDAC	0.0046	0.0106	0.0046	0.0060	0.0000***
CONUM	2359.40	2262.29	2359.60	-97.31	0.0824
COJUDG	0.0018	0.0039	0.0018	0.0021	0.0000***
CODPRY	0.0012	0.0020	0.0012	0.0008	0.0000***
COCRIM	0.0071	0.0081	0.0071	0.0010	0.0059**
COEXPEL	0.0089	0.0413	0.0088	0.0325	0.0000***
COPREMPL	0.9544	1.1910	0.9540	0.2370	0.0000***
COEMPLYR	5.3717	5.5721	5.3713	0.2008	0.0000***

The Probit Models

The first statistical model we use to estimate the propensity of any given broker to cause investor harm is the following probit regression model:

$$P(Y_{i,t} = 1 | X_{i,t-1}) = \Phi(\beta_0 + \beta_1 X_{i,t-1}), \qquad \mathrm{i} = 1,2,\dots,\mathrm{N}, \qquad \mathrm{t} = 1,2,\dots,n_i,$$

where $P(Y_{i,t} = 1 | X_{i,t-1})$ denotes the probability of investor harm in year *t* given $X_{i,t-1}$, a vector of model features containing BrokerCheck information prior to year *t*, and Φ is the cumulative distribution function ("CDF") of a standard normal random variable. The dependent variable $Y_{i,t}$ equals 1 if at least one investor harm event caused by broker *i* occurred in year t. The model relates the probability of broker *i* causing investor harm in each year *t* of broker *i*'s career to a linear combination of model features via the standard normal distribution function. Since the broker features in each year are associated to the occurrence of investor harm in the subsequent year, the brokers first registered in 2014 are effectively excluded from the sample input to the model. The number of broker-year observations used in the regression is 1,020,707, corresponding to 133,556 unique brokers. After obtaining the coefficient estimates $\hat{\beta}_0$ and $\hat{\beta}_1$, we can calculate the predicted probability of broker *j* causing investor harm in year *t* (denote by $\hat{Y}_{j,t}$) as the following

$$\hat{Y}_{j,t} = \Phi(\hat{\beta}_0 + \hat{\beta}_1 X_{j,t-1})$$

We fit a probit regression model with eleven features, including broker disclosure features and qualification and employment features (from Table 4), and prior year's annual index return on the S&P 500 index (denoted by MKRN). We refer to this model as the baseline model. The annual market return acts as a control for the macroeconomic conditions in the year under consideration.

Table 8 summarizes the estimated model coefficients (β_1) with corresponding z-scores and p-values. The signs and significance levels of the coefficients shed light on the value of information contained in each model feature. For example, since the coefficient for SA is positive and highly statistically significant, the past settlement and award information is useful in predicting future investor harm. Also the size of the coefficient is economically significant. For example, the impact of an additional previous expelled firm record on the predicted probability is equivalent to the impact of 1.57 additional settlements and award, or 3.30 additional disciplinary events, or 4.07 additional criminal records, or 5.01 additional disclosed judgements and liens.

Table 8: Coefficients of the Baseline Probit Regression Models

Feature	FINRA	SLCG
SA (Settlements and Awards)	0.332***	0.3550***
	(8.936)	(17.593)
JUDG (Judgments)	0.112***	0.1114***
	(5.748)	(5.310)
BKCY (Bankruptcy)	0.0226**	0.0124
	(2.324)	(0.453)
DPRY (Disciplinary event)	0.230***	0.1690***
	(5.038)	(3.801)
CRIM (Criminal)	0.170***	0.1371***
	(6.240)	(4.427)
EXAM (Exams Passed)	0.00663	0.0040
	(0.507)	(0.314)
EXPEL (Previous Expelled Firm)	0.432***	0.5578***
	(6.779)	(13.286)
PREMPL (Number of Prior Employers)	0.0861***	0.0253***
	(15.91)	(4.599)
EMPLYR (Years in Industry)	0.0222***	0.0424***
	(16.59)	(18.086)
DUAL (Also IA registered)	0.279***	0.2911***
	(18.81)	(14.933)
Gender (Male)	0.292***	
	(14.28)	
MKRN	-0.514***	-0.6214***
	(-14.84)	-17.790)
Model Chi-square	2303.9***	1769.74***
Observations	1,014,873	1,020,707

We calculate the predicted probability of investor harm for each year in each broker's career using the estimated model coefficients. Following Qureshi and Sokobin, we rank all broker-year observations and allocate these observations into quintiles according to their predicted probabilities. As a result, 51.29% of the investor harm events are associated with the broker-year observations allocated to the highest risk quintile, and only 3.72% of the investor harm events are attributed to the broker-year observations allocated to the lowest risk quintile. Qureshi and Sokobin found 55.5% of investor harm events were by brokers in the highest risk quartile and 3.8% in the lowest risk quartile.[8] Thus, as with the incidence of harm found, we are able to replicate the Qureshi and Sokobin regression results. If the model had no predictive ability, roughly 20% of the investor harm events should have been attributed to the broker-year observations in each of the lowest risk quintile and the highest risk quintile. Hence our findings show that the probit model is effective in discriminating brokers associated with investor harm from those without such association, and that the BrokerCheck information is useful for predicting investor harm *if all the data is available and subjected to sophisticated statistical analysis*.

Continuing to replicate Qureshi and Sokobin, we compare the predicted probability of each broker causing investor harm in a given year to the unconditional probability of investor harm, defined as the ratio of the number of broker-year observations associated with investor harm to the total number broker-year observations. The number of broker-year observations associated with investor harm is 2,094, and the unconditional probability of investor harm equals 2,094/1,020,707 = 0.21%. We predict broker j to cause investor harm in year t if and only if the predicted probability $\hat{Y}_{j,t}$ is greater than 0.21%. Using this criterion, 1,458 (69.63%) of the 2,094 broker-year observations associated with investor harm are correctly predicted to have investor harm, while 337,475 (33.13%) of the 1,018,613 broker-year observations without investor harm are erroneously predicted to have investor harm. The baseline probit model at cutoff of 0.21% has a true positive rate of 69.63% and a false positive rate of 33.13%. This provides further evidence that the baseline model is effective in predicting investor harm.

[8] The difference in our results in this baseline model and Qureshi and Sokobin's results appears to be their access to the brokers' gender which is not available in the data on BrokerCheck and on which we rely.

We explore alternative specifications by fitting 9 probit regression models on different sets of features to assess the importance of each feature in predicting investor harm. Model A1, the FINRA study's baseline model, includes prior settlements and awards, disclosed judgments and liens, bankruptcies within ten years, disciplinary and criminal events, exams passed, associations with expelled firms, dual registration and the previous year's stock market return as explanatory variables. Model A2 adds information on co-workers included in FINRA's baseline model for the subject including prior affiliation with expelled firms, number of prior employers and number of years in the industry to Model A1's list of explanatory variables. Model A3 adds harm associated with coworkers as well as coworker's disclosed judgements and liens, and disciplinary and criminal events. We apply an exponent of 1/3 to the coworker disclosure features to better fit the data.

Model A1 (FINRA's Baseline Model): SA + JUDG + BKCY + DPRY + CRIM + EXAM + EXPEL + PREMPL + EMPLYR + DUAL + MKRN
Model A2: **Model A1** + CONUM + COEXPEL + COPREMPL + COEMPLYR
Model A3: **Model A2** + HAC $^{1/3}$ + COJUDG $^{1/3}$ + CODPRY $^{1/3}$ + COCRIM $^{1/3}$

We construct Model B1 by excluding the qualification and employment history of individual brokers from the baseline model, and build Models B2 and B3 by successively adding two groups of coworker features to model B1.

Model B1: EXAM + EXPEL + PREMPL + EMPLYR + DUAL + MKRN
Model B2: **Model B1** + CONUM + COEXPEL + COPREMPL + COEMPLYR
Model B3: **Model B2** + HAC $^{1/3}$ + COJUDG $^{1/3}$ + CODPRY $^{1/3}$ + COCRIM $^{1/3}$

Models C1 and C2 are created by replacing settlements and awards above the FINRA thresholds (SA) with all customer disputes (CD) in Models A1 and A2, respectively. To obtain Model C3 from A3, we replace settlements and awards above the FINRA thresholds (SA) with customer disputes (CD) and all coworker settlements and awards above the FINRA thresholds (HAC) with all coworker customer complaints (CDAC).

Model C1: CD + JUDG + BKCY + DPRY + CRIM + EXAM + EXPEL + PREMPL + EMPLYR + DUAL + MKRN
Model C2: **Model C1** + CONUM + COEXPEL + COPREMPL + COEMPLYR
Model C3: **Model C2** + CDAC $^{1/3}$ + COJUDG $^{1/3}$ + CODPRY $^{1/3}$ + COCRIM $^{1/3}$

Table 9 reports the estimated coefficients with significance levels for the best models within each grouping above, Models A3, B3, and C3. In each model, most features are statistically significant. The features of past wrongdoing associated with the broker or the broker's coworkers (namely, SA, CD, HAC, and CDAC) achieve the highest z-scores in their respective models. The log-likelihood chi-square statistics indicate that C3 is the most statistically significant model.

This result has important implications. Model C3 includes all of the broker's and the broker's coworkers' customer complaints not just the ones which result in settlements and awards in excess of the FINRA thresholds. That is, even if all you are trying to predict is filings that result in settlements and awards in excess of the FINRA thresholds, the best set of explanatory features includes all customer complaints. Thus, as an empirical matter, cases that are dismissed or which result in settlements or awards below the FINRA threshold contain important information for differentiating good brokers from bad brokers.

Table 9: Coefficients with Significant Levels and Model Chi-Squares for Probit Regression Models. The values in the parentheses are z-cores for the coefficients. The symbols ***, ** and * denote significance levels of 0.001, 0.01 and 0.05, respectively.

Features	Model A3	Model B3	Model C3
SA	0.3126***		
	(15.400)		
CD			0.2422***
			(21.076)
JUDG	0.0657**		0.0588*
	(2.757)		(2.377)
BKCY	0.0178		0.0175
	(0.647)		(0.633)
DPRY	0.1181**		0.0558
	(2.597)		(1.157)
CRIM	0.1073***		0.0919**
	(3.302)		(2.733)
EXAM	0.0146	0.0098	0.0058
	(1.122)	(0.754)	(0.444)
EXPEL	0.1804***	0.2133***	0.1402**
	(3.394)	(4.121)	(2.581)
PREMPL	0.0007	0.0033	-0.0017
	(0.118)	(0.546)	(-0.281)
EMPLYR	0.0587***	0.0624***	0.0559***
	(20.238)	(21.774)	(18.936)
DUAL	0.3186***	0.3199***	0.2929***
	(14.810)	(15.062)	(13.294)
MKRN	-0.4489***	-0.4383***	-0.4598***
	(-12.070)	(-11.862)	(-12.404)

CONUM	-0.00003***	-0.00003***	-0.00002***
	(-8.655)	(-9.224)	(-6.993)
COEXPEL	0.5013***	0.5604***	0.2734*
	(4.273)	(4.911)	(2.239)
COPREMPL	0.0750***	0.0769***	0.0774***
	(4.282)	(4.457)	(4.307)
COEMPLYR	-0.0903***	-0.0924***	-0.0950***
	(-13.793)	(-14.225)	(-14.206)
HAC $^{1/3}$	2.548***	2.724***	
	(17.935)	(19.477)	
CDAC $^{1/3}$			2.333***
			(19.064)
COJUDG $^{1/3}$	0.3530**	0.4048***	0.2783*
	(2.871)	(3.342)	(2.212)
CODPRY $^{1/3}$	0.4239**	0.4188**	0.1713
	(3.011)	(2.994)	(1.194)
COCRIM $^{1/3}$	-0.0684	-0.0856	-0.2634*
	(-0.583)	(-0.732)	(-2.174)
Model Chi-square	2498.80***	2176.91***	2796.41***

Table 10 reports true positive rates, false positive rates, and distribution of investor harm events among quintiles of broker-year observations for each model. The best probit model sort brokers so that the highest risk quintile captures 60% of investor harm events and the lowest risk quintile is associated with less than 2% of the investor harm events.

Table 10: Within-Sample Predictive Performance of Probit Regression Models

Models	True Positive	False Positive	Investor Harm Events 1st Quintile	2nd Quintile	3rd Quintile	4th Quintile	5th Quintile
A1	69.63%	33.13%	3.72%	8.17%	11.99%	24.83%	51.29%
A2	72.54%	34.04%	2.44%	6.26%	13.85%	22.02%	55.44%
A3	71.30%	31.27%	2.34%	5.21%	13.80%	20.15%	58.50%
B1	69.15%	34.89%	3.96%	8.55%	13.80%	27.13%	46.56%
B2	72.40%	35.35%	2.58%	7.26%	13.94%	23.16%	53.06%
B3	71.20%	32.10%	2.58%	5.49%	13.42%	22.45%	56.06%
C1	68.19%	30.73%	3.44%	6.83%	12.23%	25.36%	52.15%
C2	72.16%	32.99%	2.44%	5.54%	13.99%	20.92%	57.12%
C3	72.87%	31.10%	1.77%	5.54%	11.13%	21.39%	60.17%

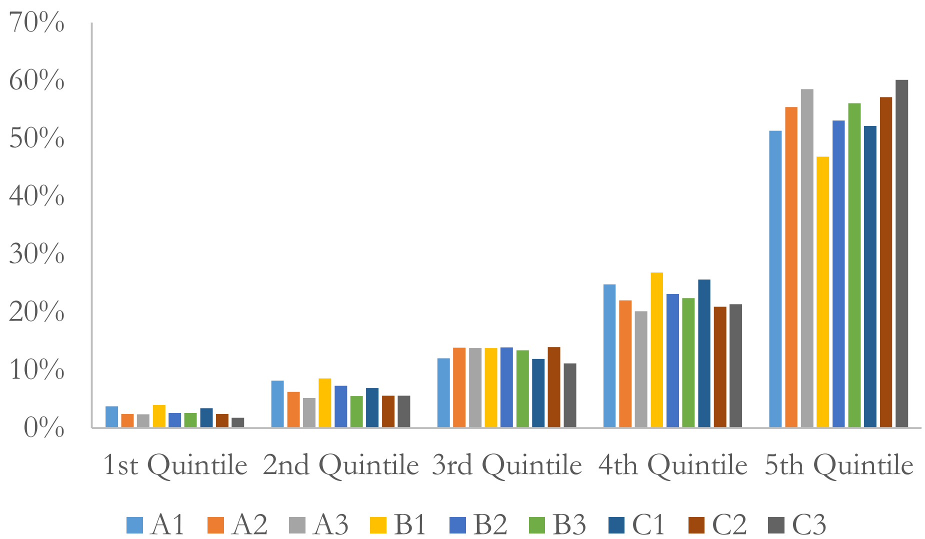
The within-sample measures might overstate the predictive performance since a model that fits one dataset well may not necessarily predict well when applied to a different dataset (called "overfitting"). To gauge the true predictive efficacy of the models we perform 5-fold cross-validations on the entire data. We randomly partition all the broker-year observations into five groups. Each group is left out and a probit model is fit to the remaining groups combined. The estimated model is then used to predict the occurrence of investor harm in the held-out group. The prediction accuracy measures for all five models are averaged to generate the accuracy measure for one cross-validation procedure. We repeat the procedure ten times and average the resulting ten groups of accuracy measures to get the final measures, which are reported in Table 11.

Table 11: Cross-Validated Predictive Performance of Probit Regression Models

			Investor Harm Events				
Models	True Positive	False Positive	1st Quintile	2nd Quintile	3rd Quintile	4th Quintile	5th Quintile
A1	69.46%	32.95%	3.72%	8.15%	12.03%	24.78%	51.33%
A2	72.59%	34.02%	2.44%	6.24%	13.87%	22.04%	55.41%
A3	71.31%	31.26%	2.36%	5.18%	13.78%	20.17%	58.51%
B1	69.11%	35.01%	3.97%	8.54%	13.79%	26.83%	46.86%
B2	72.48%	35.37%	2.58%	7.27%	13.92%	23.17%	53.07%
B3	71.26%	32.13%	2.56%	5.53%	13.39%	22.44%	56.08%
C1	68.35%	30.83%	3.44%	6.88%	11.91%	25.62%	52.15%
C2	72.27%	32.97%	2.43%	5.55%	13.99%	20.92%	57.11%
C3	72.81%	31.13%	1.76%	5.55%	11.15%	21.40%	60.14%

Figure 2 plots the proportion of broker misconduct disclosures in each quintile of broker risk as predicted by the probit models. All nine specifications can differentiate high risk brokers from low risk brokers.

Figure 2: Performance of Probit Regression Models



The true positive rates and the numbers of investor harm events captured in the highest quintile in Table 11 are, at worst, only slightly smaller than those in Table 10 and follow a similar pattern across the models. The models including both broker and coworker features (e.g., A2, A3, C2, C3) assign more than 55% of the investor harm events in the hold-out samples to the highest quintile and less than 3% of investor harm events to the lowest quintile, which confirms that the BrokerCheck database contains valuable information for discriminating potentially harmful brokers from non-harmful ones. The improvement in predictive performance from Model A1 to Model A3, and from Model B1 to Model B3 indicates that the coworker features play a significant role in making better predictions about investor harm. The improved performance of Models C1-C3 over Models A1-A3 suggests that not only the broker's disputes leading to award or settlement above a threshold amount, but also those pending, denied, or closed without action are useful in determining the likelihood of future investor harm event as defined by Qureshi and Sokobin.

4. The Random Forest Models

The probit model combines features in a linear fashion and may have poor predictive performance on data with complex nonlinear structures. To predict investor harm more effectively we consider a more sophisticated statistical model, the *random forest model*.

The random forest model is a tool for regression and classification that makes decisions based on the consensus of results from an ensemble of tree models (see [1]). It is known for its broad applications, fast implementation, and remarkable prediction accuracy on a wide range of problems. The impressive predictive strength of the random forest model results from the algorithm's efficient variance reduction, achieved by combining the bootstrap aggregating and random subspace techniques. The way that random forest models make predictions is similar to other more conventional tools of data mining. The model parameters are estimated using observations in the data set, each of which is associated with a vector of feature values and a response value (called *model training*). Then the trained model is applied to new data with known feature values to predict the unknown responses. The response and features used in random forest models for investor harm prediction are defined the same way as those in the probit models, and we continue to use information from prior years to predict the investor harm in the current year. The random forest model can be written in formula as

$$Y_{i,t} = RandomForest\left(X_{i-1,t}; \Theta\right), \qquad \mathrm{i} = 1,2,\ldots,\mathrm{N},\ \mathrm{t} = 1,2,\ldots,n_i,$$

where $Y_{i,t}$ equals 1 if a customer complaint leading to an award or settlement above $10,000 / $15,000 was filed against broker *i* in year *t* and 0 otherwise, $X_{i-1,t}$ is the feature vector for broker *i* and year *t*, and Θ represents model parameters. The function "RandomForest" encapsulates the decision rules of the random forest model. After obtaining the parameter estimates $\hat{\Theta}$, the predicted probability of investor harm by broker *j* in year *t* is calculated as

$$\hat{Y}_{j,t} = RandomForest(X_{j-1,t}; \hat{\Theta})$$

Two important tuning parameters of a random forest model are the number of trees or bootstrap samples used by the forest (denoted by "n_tree") and the number of features randomly selected when splitting each tree node (denoted by "m_try"). As n_tree grows the predictive power of the random forest model increases and eventually stabilizes. Due to limitation of computational resources we build each forest model with n_tree = 300. We let m_try be the number of features divided by 3 rounded down to previous integer, the default value recommended by the inventor of the model. We build 8 random forest regression models using different subsets of features.

Model RF1: SA, JUDG, BKCY, DPRY, CRIM, EXAM, EXPEL, PREMPL, EMPLYR, DUAL, MKRN.
Model RF2: **Model RF1** + CONUM, COEXPEL, COPREMPL, COEMPLYR, HAC.
Model RF3: CD, JUDG, BKCY, DPRY, CRIM, EXAM, EXPEL, PREMPL, EMPLYR, DUAL, MKRN, CONUM, COEXPEL, COPREMPL, COEMPLYR, CDAC.
Model RF4: EXAM, EXPEL, PREMPL, EMPLYR, DUAL, MKRN.
Model RF5: **Model RF4 +** SA.
Model RF6: **Model RF4 +** HAC.
Model RF7: **Model RF4 +** SA + HAC.
Model RF8: **Model RF4 +** CD + CDAC.

Table 12 reports measures of predictive strength of the random forest models based on *out-of-bag estimates*. Analogous to cross-validation predictive measures, the out-of-bag estimates provide good approximations to the true predictive performance of random forest models on new data.

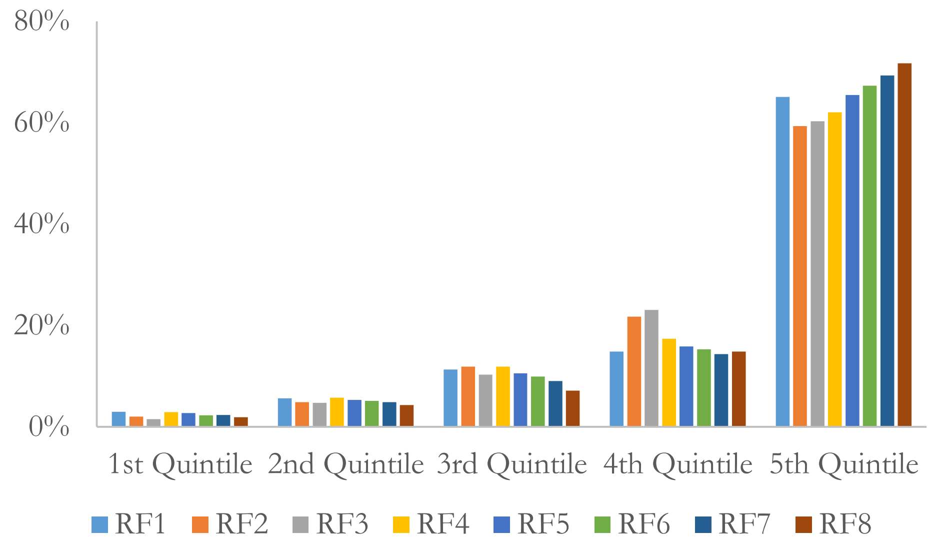
Table 12: Out-of-Bag Predictive Performance of Random Forest Models

Models	Investor Harm Events						
	True Positive	False Positive	1st Quintile	2nd Quintile	3rd Quintile	4th Quintile	5th Quintile
RF1	67.96%	23.73%	3.01%	5.68%	11.32%	14.90%	65.09%
RF2	48.90%	11.99%	2.05%	4.92%	11.89%	21.78%	59.36%
RF3	50.38%	11.47%	1.53%	4.78%	10.32%	23.11%	60.27%
RF4	70.87%	29.70%	2.91%	5.78%	11.89%	17.38%	62.03%
RF5	72.68%	27.53%	2.72%	5.35%	10.55%	15.90%	65.47%
RF6	69.48%	22.77%	2.29%	5.16%	9.93%	15.33%	67.29%
RF7	72.16%	22.88%	2.34%	4.87%	9.07%	14.37%	69.34%
RF8	73.97%	22.28%	1.91%	4.30%	7.16%	14.90%	71.73%

The best random forest models sort brokers so that the highest risk quintile is associated with over 70% of investor harm events and the lowest risk quintile is associated with less than 2% of the investor harm events. Broker and brokerage firm rankings based on the random forest models would be tremendously helpful to investors trying to avoid broker misconduct.

Figure 3 plots the proportion of broker misconduct disclosures in each quintile of broker risk as predicted by the random forest models. All eight random forest model specifications do a better job than the probit models at differentiating high risk brokers from low risk brokers.

Figure 3: Out-of-Bag Predictive Performance of Random Forest Models



Random forest models easily outperform probit models in predicting the most harmful brokers. All eight random forest models allocate more investor harm events to the highest risk quintile of brokers than the most effective probit model (C3). Model RF1 sees a significant improvement in prediction accuracy over the baseline probit model A1 which uses the same set of features as inputs. The best random forest model (RF8) captures 20% more investor harm events in the highest risk quintile of broker-year observations ranked by predicted probability of investor harm than the best probit model (C3). Thus, the effectiveness of BrokerCheck information for predicting investor harm based on all the data depends also on the sophistication of the models being used.

The gradual improvement of predictive power from Model RF4 to Model RF7 confirms that the information about past settlement and award for both the subject broker and the broker's coworkers is useful for making predictions with random forests. Comparing RF2 with RF3 and RF7 with RF8 shows that using prior customer complaints regardless of status in place of prior investor harm events as defined by Qureshi and Sokobin is more effective in predicting future investor harm. Somewhat surprisingly, the performance of Model RF1 drops when we add relevant coworker features to create Model RF2. This is possibly due to the relatively small value of n_tree. There is evidence that using more trees can improve prediction accuracy: Model RF2 with n_tree = 700 allocates around 62.9% investor harm events to the highest quintile. We conjecture that the performance of Model RF2 will eventually surpass that of Model RF1 as the number of trees gets large enough. On the other hand, increasing n_tree for Model RF8 may not lead to a significant gain in predictive power: the model RF8 with n_tree = 700 allocates around 72.1% investor harm events to the highest quintile.

While it is impossible to evaluate the strength or relevance of features in a random forest model by Z-scores and p-values, the model has its own built-in measurements of feature importance. The model can return an "importance score" for each feature, which measures the feature's relative contribution to the overall predictive power of the model. We report the importance scores for the features used in four random forest models in Table 13. Coworker settlements and awards (HAC) and coworker complaints (CDAC) achieve the highest scores in Models RF7 and RF8, respectively, showing the importance of coworker information in increasing the predictive power of our random forest models.

Table 13: Importance Scores from Random Forest Models

Features	Model RF1	Model RF3	Model RF7	Model RF8
SA	18.41		18.88	
CD		28.30		27.68
JUDG	6.63	8.94		
BKCY	3.42	6.90		
DPRY	3.93	7.88		
CRIM	5.39	12.88		
EXAM	7.38	41.55	8.22	10.06
EXPEL	3.02	4.96	3.37	3.69
PREMPL	20.20	56.46	20.99	25.57
EMPLYR	21.66	66.57	23.60	28.48
DUAL	3.88	11.63	4.12	4.77
MKRN	19.78	24.42	18.50	22.50
CONUM		80.17		
COEXPEL		65.91		
COPREMPL		130.44		
COEMPLYR		147.65		
HAC			89.62	
CDAC		101.28		122.91

5. Discussion

a. Current BrokerCheck Provides Little Useful Information

Qureshi and Sokobin conclude that BrokerCheck information "has significant power to discriminate between brokers associated with investor harm events and other brokers [because] [t]he 20% of brokers with the highest ex ante predicted probability of investor harm are associated with more than 55% of the investor harm events in our sample." While this statement may accurately reflect Qureshi and Sokobin's statistical analysis of more than a dozen variables for each of the 181,133 brokers in their study, it says nothing about whether BrokerCheck provides retail investors with any useful information. An investor in Dallas going to the BrokerCheck website to research a broker who just cold-called them from an impressive sounding Long Island brokerage firm cannot determine whether the broker is one Qureshi and Sokobin know to be one of "the 20% of brokers with the highest ex ante predicted probability of investor harm."

Qureshi and Sokobin's analysis and our analysis start with BrokerCheck records on over 1.2 million brokers. We both apply sophisticated statistic modeling. Investors need all the data and our sophisticated modeling to glean the information Qureshi and Sokobin attribute to the BrokerCheck data. Retail investors preyed upon by bad brokers do not have access to the all the ostensibly public BrokerCheck data or to our analytical capabilities.

Retail investors can only observe an infinitesimal portion of the BrokerCheck data. Investors querying BrokerCheck only see information on one broker at a time and so do not know whether a broker's reported characteristics are unusual or not and whether those characteristics portend a higher likelihood that the broker they are querying will engage in fraud.

Imagine that BrokerCheck displays data on a wall eight feet high and running the 120 yard length of a football field, from the back of one end zone to the back of the far end zone. Now imagine that FINRA places a black-out drape over the entire length of the wall. Qureshi and Sokobin use all the data on the wall but when investors want information from BrokerCheck, FINRA opens the drapes the thickness of two sheets of copier paper. If investors were to review the information on 100 brokers it would still only be less than ½ an inch of data on FINRA's shrouded 120 yard wall of data. Even if investors had the analytical capabilities we have they could never learn from that vanishing small sliver of the data what Qureshi and Sokobin derive from the BrokerCheck data.

b. Current BrokerCheck Information is Insufficient to Differentiate High Risk from Low Risk Brokers

Our analysis, and the analyses conducted by Egan, Matvos and Seru and Qureshi and Sokobin, show that association with past customer complaints and disciplinary events is a good indicator of higher propensity for future investor harm. While avoiding brokers with disclosure events may be a good rule of thumb for unsophisticated investors who have access to nothing more than public BrokerCheck information, it is not sufficient. Even at the highest risk firms, 80% of brokers don't have customer complaints. The 20% of brokers at these firms with a history of customer complaints do, though, increase the likelihood that another broker at the same firm with a clean record will cause investor harm in the future. Investors need to know the disciplinary history of a broker's co-workers.

To illustrate, consider two brokers with the same time in the industry and identically clean records at the end of 2014 – no customer complaints, no judgments or liens, no bankruptcies, no disciplinary events, no criminal record. Both have passed two exams, were never affiliated with an expelled, had only one prior employer and were dually registered. An investor using BrokerCheck to make an educated choice between the brokers would be at a loss. BrokerCheck could not help this investor select the more trustworthy broker. Perhaps she would toss a coin.

One of these two brokers ("Broker A") had a customer complaint in 2015 that is still pending as of 2016 while he maintained a clean disclosure record in 2014. The broker with a complaint in 2015 is currently employed at Aegis Capital Corp, and the broker without a customer complaint is employed at Pyramis Distributors Corporation LLC.

Noticeably, our statistical models are able to provide some additional insights into the quality of these two brokers. The probabilities of inflicting investor harm in 2014 predicted by the best performing probit model C3 for the Aegis broker was 0.50% and for the Pyramis broker was 0.007%. The probit analysis places the Aegis broker in the highest risk quintile on December 31, 2014 and the Pyramis broker in the lowest risk quintiles on December 31, 2014. The discriminative power of Model C3 results from its use of eight coworker features, none of which is accessible through BrokerCheck. The different coworker features of the two brokers on December 31, 2014 are presented in Table 14.

Table 14: Coworker Input and Coefficients of Two Brokers

	CONUM	COPREMPL	COEMPLYR	$CDAC^{1/3}$	$COJUDG^{1/3}$	$CODPRY^{1/3}$	$COCRIM^{1/3}$	COEXPEL
Broker A	58	2.73	6.62	0.3895	0.337	0.2267	0.1595	0.4398
Broker B	62	1.4	6.34	0	0.1136	0	0.2103	0
Δ	-4	1.33	0.28	0.3895	0.2234	0.2267	-0.0508	0.4398
C3 Coef	-0.00002	0.0774	-0.0950	2.333	0.2783	0.1713	-0.2634	0.2734
Δ ×Coef	0.00008	0.10294	-0.0266	0.90870	0.06217	0.03883	0.01338	0.12024

On average, the Aegis broker had a much more "harmful" group of coworkers than the Pyramis broker. Although there was no definitive evidence of a causal relationship between having distinctive coworker profile in 2014 and causing investor harm in 2015, adding coworker information has clearly helped the statistical model make a more educated (in this case, likely correct) guess about the likelihood of future investor harm. Among all the coworker features in this example, CDAC or Customer Disputes Associated with Coworkers has the higher impact, 74.50% on the final predicted probability.

While Qureshi and Sokobin also noticed that "investors may benefit from information about harm associated with brokers' coworkers" (which they also denoted by "HAC"), their analysis has left an impression that including coworker harm variable can only lead to a marginal increase in efficiency on top of the current BrokerCheck information.[9] The above example shows that information about coworkers, in particular CDAC is tremendously useful for investors to make wise decisions in choosing between brokers with clean personal disclosure records.

c. Fixing BrokerCheck and Reducing Misconduct Is Easy

FINRA could easily make the public-facing BrokerCheck data available in bulk to anyone interested in analyzing the data. FINRA and the SEC have already determined that this information is not confidential and should be disseminated to the public. FINRA has so thoroughly throttled the distribution of this important data as to make it virtually useless.

Our results above confirm the findings of Qureshi and Sokobin and of Egan, Matvos, and Seru that the risk a broker will commit misconduct is significantly increased if he or she works with co-workers who have previously committed misconduct. In fact, investors would be as well informed to know the average misconduct history of a broker's coworkers as they would be knowing the broker's own misconduct history. If the publicly available BrokerCheck information were truly publicly available researchers, third party vendors, ratings companies like Lipper and Morningstar, and news outlets like US News and World Report and BusinessWeek could rank brokerage firms on the risk of fraud. These rankings would generate substantial publicity and bad brokerage firms would no longer be able to prey on unsophisticated investors with relative impunity.

In Appendix 1 we list the 210 brokerage firms with 400 or more registered brokers sorted by the percentage of their brokers associated with investor harm events as defined by Qureshi and Sokobin as of December 31, 2015. We also report firm rankings by the percentage of brokers with misconduct disclosures as defined by Egan, Matvos, and Seru.

Table 15 excerpts the 30 firms with the highest percentage of brokers associated with investor harm events from Appendix 1. We have identified the firms with more than 1,000 brokers in bold font. These firms are the firms identified in Egan, Matvos, and Seru's Table 6. There are six firms with a higher percentage of brokers with associated with investor harm events than Oppenheimer, the highest risk firm with more than 1,000 brokers in the Egan, Matvos, and Seru study. The top six firms in Table 15 are the same whether

[9] Their baseline probit model allocated 55.5% of the investor harm events to the highest quintile, compared to 58.9% of the investor harm events captured in the top quintile in the Baseline + HAC prediction.

we rank them based on the Qureshi and Sokobin investor harm measure or Egan, Matvos, and Seru's financial misconduct measure.

These six firms – Aegis Capital, Summit Brokerage Services, National Securities, Centaurus Financial, Independent Financial Group and Kovack Securities employ a far higher percentage of bad brokers than other firms. These six highest-risk firms are also among the top ten firms ranked by percentage of current brokers who were previously fired by other firms after customer allegations of misconduct. 7.71% of the registered brokers in these six high risk firms have been fired at least once by a previous employer after allegations of misconduct, 10 times the average of 0.78% of the remaining 204 brokerage firms.[10] Given their coworkers' disclosure record as of 2014, 83.7% of the brokers at these six firms would be in the highest risk quintile as defined in the FINRA study and should be avoided by investors. The BrokerCheck reports for most of the brokers at these six firms should prominently display a skull and crossbones warning.

Table 15. Top 30 firms with 400 or more registered brokers ranked by percentage of brokers with investor harm events defined in FINRA study.

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers Previously Fired	Hired Fired Rate
1	15007	AEGIS CAPITAL CORP.	444	107	24.10%	21	4.73%
2	34643	SUMMIT BROKERAGE SERVICES	676	129	19.08%	65	9.62%
3	7569	NATIONAL SECURITIES CORP.	760	137	18.03%	42	5.53%
4	30833	CENTAURUS FINANCIAL, INC.	602	98	16.28%	39	6.48%
5	7717	INDEPENDENT FINANCIAL GROUP	638	90	14.11%	50	7.84%
6	44848	KOVACK SECURITIES INC.	434	58	13.36%	57	13.13%
7	**249**	**OPPENHEIMER & CO. INC.**	**2,217**	**276**	**12.45%**	**92**	**4.15%**
8	877	WEDBUSH SECURITIES INC.	634	77	12.15%	32	5.05%
9	30613	INVESTORS CAPITAL CORP.	641	72	11.23%	39	6.08%
10	2543	WUNDERLICH SECURITIES, INC.	459	51	11.11%	21	4.58%
11	**8174**	**UBS FINANCIAL SERVICES INC.**	**12,555**	**1,377**	**10.97%**	**94**	**0.75%**
12	**32444**	**FIRST ALLIED SECURITIES, INC.**	**1,179**	**122**	**10.35%**	**49**	**4.16%**
13	46214	NEXT FINANCIAL GROUP, INC.	796	75	9.42%	21	2.64%
14	14503	VSR FINANCIAL SERVICES, INC.	511	48	9.39%	7	1.37%
15	18456	STERNE AGEE FINANCIAL SERVICES	580	53	9.14%	35	6.03%
16	**11025**	**WELLS FARGO ADVISORS FINANC**	**1,993**	**176**	**8.83%**	**30**	**1.51%**
17	18487	AMERICAN PORTFOLIOS FINANC	838	73	8.71%	36	4.30%
18	**149777**	**MORGAN STANLEY SMITH BARNEY**	**23,782**	**2,065**	**8.68%**	**151**	**0.63%**
19	**705**	**RAYMOND JAMES & ASSOCIATES**	**5,812**	**488**	**8.40%**	**101**	**1.74%**
20	**463**	**JANNEY MONTGOMERY SCOTT**	**1,369**	**114**	**8.33%**	**29**	**2.12%**

[10] 274 out of 3,554, or 7.71% registered brokers in those top six companies have at least one "Employment Separation After Allegations" on BrokerCheck, while 4,543 out of 549,617, or 0.83% registered brokers in our 210 brokerage firms have at least one "Employment Separation After Allegations".

21	**793**	**STIFEL, NICOLAUS & COMPANY**	**4,588**	**380**	**8.28%**	**113**	**2.46%**
22	14303	SIGMA FINANCIAL CORPORATION	678	55	8.11%	19	2.80%
23	7684	INVESTACORP, INC.	500	39	7.80%	17	3.40%
24	**19616**	**WELLS FARGO ADVISORS, LLC**	**26,319**	**1,998**	**7.59%**	**309**	**1.17%**
25	**10205**	**SECURITIES AMERICA, INC.**	**2,662**	**191**	**7.18%**	**61**	**2.29%**
26	20804	UNITED PLANNERS' FINANC SERV	510	35	6.86%	19	3.73%
27	18697	GIRARD SECURITIES, INC.	477	31	6.50%	19	3.98%
28	**10299**	**CETERA ADVISORS LLC**	**1,618**	**105**	**6.49%**	**48**	**2.97%**
29	**29604**	**NATIONAL PLANNING CORP.**	**1,815**	**117**	**6.45%**	**38**	**2.09%**
30	**35747**	**PURSHE KAPLAN STERLING**	**1,229**	**78**	**6.35%**	**33**	**2.69%**

If FINRA unshackled BrokerCheck, researchers would come up with innovative ways to reach and inform unsophisticated investors about high risk brokers and brokerage firms. Releasing the potential of BrokerCheck to protect investors would also benefit some brokers and brokerage firms. Brokers with clean CRDs would have an incentive to move to firms with a lower proportion of bad brokers so they would not be penalized in the rankings for associating with bad brokers and brokerage firms would compete to hire better brokers and fire brokers with prior settlements and awards to improve their quality rankings.

Continuing with the analogy above, proposals to supplement data items available on BrokerCheck or add a search term miss the mark badly. These proposals amount to adding a few inches of height to the wall *and* to the drapes. If FINRA continues to only allow investors the benefit of a glimpse at 0.01 inch of the 120 yard long wall it won't matter if the wall of data is 8 feet tall or 8 feet, 2 inches tall. The only way to empower investors to protect themselves is for FINRA to take down the drapes.

Bibliography

[1] Breiman, L. (2001), "Random Forests". Machine Learning 45: 5-32.

[2] Dimmock, S., Gerken, W. and Graham N. (2015), "Is Fraud Contagious? Career Networks and Fraud by Financial Advisers." Working paper.

[3] Egan, M., Matvos, G. and Seru, A. (2016), "The Market for Financial Adviser Misconduct." Working paper, SSRN-id2739170.

[4] Qureshi, H. and Sokobin, J. (2015), "Do Investors Have Valuable Information About Brokers?" Working paper.

[5] United States Securities and Exchange Commission, "Study and Recommendations on Improved Investor Access to Registration Information about Investment Advisers and Broker-Dealers" January 2011.

Appendix 1: 210 Brokerage Firms with more than 400 Brokers as of December 31, 2015

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
1	15007	AEGIS CAPITAL CORP.	444	107	24.10%	156	35.14%	1
2	34643	SUMMIT BROKERAGE SERVICES	676	129	19.08%	209	30.92%	3
3	7569	NATIONAL SECURITIES CORP	760	137	18.03%	240	31.58%	2
4	30833	CENTAURUS FINANCIAL, INC.	602	98	16.28%	167	27.74%	5
5	7717	INDEPENDENT FINANCIAL GROUP	638	90	14.11%	175	27.43%	6
6	44848	KOVACK SECURITIES INC.	434	58	13.36%	124	28.57%	4
7	**249**	**OPPENHEIMER & CO. INC.**	**2,217**	**276**	**12.45%**	**433**	**19.53%**	**10**
8	877	WEDBUSH SECURITIES INC.	634	77	12.15%	133	20.98%	8
9	30613	INVESTORS CAPITAL CORP.	641	72	11.23%	131	20.44%	9
10	2543	WUNDERLICH SECURITIES, INC.	459	51	11.11%	100	21.79%	7
11	**8174**	**UBS FINANCIAL SERVICES INC.**	**12,555**	**1,377**	**10.97%**	**1,863**	**14.84%**	**19**
12	**32444**	**FIRST ALLIED SECURITIES, INC.**	**1,179**	**122**	**10.35%**	**201**	**17.05%**	**13**
13	46214	NEXT FINANCIAL GROUP, INC.	796	75	9.42%	137	17.21%	12
14	14503	VSR FINANCIAL SERVICES, INC.	511	48	9.39%	65	12.72%	30
15	18456	STERNE AGEE FINANCIAL SERVICES	580	53	9.14%	110	18.97%	11
16	**11025**	**WELLS FARGO ADVISORS FINANC**	**1,993**	**176**	**8.83%**	**298**	**14.95%**	**18**
17	18487	AMERICAN PORTFOLIOS FINANCIAL	838	73	8.71%	142	16.95%	14
18	**149777**	**MORGAN STANLEY SMITH BARNEY**	**23,782**	**2,065**	**8.68%**	**3,047**	**12.81%**	**29**
19	**705**	**RAYMOND JAMES & ASSOCIATES**	**5,812**	**488**	**8.40%**	**767**	**13.20%**	**26**
20	**463**	**JANNEY MONTGOMERY SCOTT**	**1,369**	**114**	**8.33%**	**190**	**13.88%**	**22**
21	**793**	**STIFEL, NICOLAUS & COMPANY**	**4,588**	**380**	**8.28%**	**605**	**13.19%**	**27**
22	14303	SIGMA FINANCIAL CORPORATION	678	55	8.11%	96	14.16%	20
23	7684	INVESTACORP, INC.	500	39	7.80%	81	16.20%	15
24	**19616**	**WELLS FARGO ADVISORS, LLC**	**26,319**	**1,998**	**7.59%**	**3,189**	**12.12%**	**33**
25	**10205**	**SECURITIES AMERICA, INC.**	**2,662**	**191**	**7.18%**	**364**	**13.67%**	**23**

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
26	20804	UNITED PLANNERS' FINAN SERV	510	35	6.86%	79	15.49%	16
27	18697	GIRARD SECURITIES, INC.	477	31	6.50%	73	15.30%	17
28	**10299**	**CETERA ADVISORS LLC**	**1,618**	**105**	**6.49%**	**220**	**13.60%**	**24**
29	**29604**	**NATIONAL PLANNING CORP**	**1,815**	**117**	**6.45%**	**252**	**13.88%**	**21**
30	**35747**	**PURSHE KAPLAN STERLING INVES**	**1,229**	**78**	**6.35%**	**150**	**12.21%**	**32**
31	115368	PARKLAND SECURITIES, LLC	464	29	6.25%	62	13.36%	25
32	199	D.A. DAVIDSON & CO.	914	57	6.24%	95	10.39%	46
33	**31194**	**RBC CAPITAL MARKETS, LLC**	**5,250**	**318**	**6.06%**	**507**	**9.66%**	**52**
34	**6694**	**RAYMOND JAMES FINANCIAL SERV**	**5,461**	**323**	**5.91%**	**619**	**11.33%**	**35**
35	3866	KMS FINANCIAL SERVICES, INC.	452	26	5.75%	48	10.62%	43
36	43100	QUESTAR CAPITAL CORPORATION	824	47	5.70%	102	12.38%	31
37	2225	SII INVESTMENTS, INC.	809	46	5.69%	87	10.75%	42
38	**23131**	**ROYAL ALLIANCE ASSOCIATES**	**2,153**	**121**	**5.62%**	**237**	**11.01%**	**38**
39	**7691**	**MERRILL LYNCH**	**33,288**	**1,841**	**5.53%**	**2,798**	**8.41%**	**66**
40	**133763**	**SAGEPOINT FINANCIAL, INC.**	**2,278**	**121**	**5.31%**	**269**	**11.81%**	**34**
41	453	J.J.B. HILLIARD, W.L. LYONS, LLC	732	37	5.05%	63	8.61%	63
42	**6363**	**AMERIPRISE FINANCIAL SERVICES**	**13,952**	**699**	**5.01%**	**1,472**	**10.55%**	**45**
43	**7461**	**FSC SECURITIES CORPORATION**	**1,503**	**74**	**4.92%**	**162**	**10.78%**	**41**
44	25803	TRIAD ADVISORS, INC.	838	41	4.89%	87	10.38%	47
45	3496	STEPHENS INC.	659	32	4.86%	55	8.35%	68
46	31243	VALMARK SECURITIES, INC.	455	22	4.84%	35	7.69%	81
47	**12984**	**IFC HOLDINGS, INC.**	**1,496**	**70**	**4.68%**	**158**	**10.56%**	**44**
48	1763	H. BECK, INC.	793	37	4.67%	103	12.99%	28
49	**39543**	**CAMBRIDGE INVESTM RESEARCH**	**3,538**	**162**	**4.58%**	**359**	**10.15%**	**48**
50	10641	CADARET, GRANT & CO., INC.	912	41	4.50%	103	11.29%	36
51	**42046**	**NFP ADVISOR SERVICES, LLC**	**1,917**	**86**	**4.49%**	**186**	**9.70%**	**51**
52	142785	BB&T SECURITIES, LLC	764	34	4.45%	62	8.12%	71

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
53	**14869**	**AMERITAS INVESTMENT CORP**	**1,463**	**64**	**4.37%**	**141**	**9.64%**	**54**
54	**17499**	**SUNTRUST INVESTMENT SERVICES**	**1,510**	**66**	**4.37%**	**164**	**10.86%**	**40**
55	**6413**	**LPL FINANCIAL LLC**	**18,315**	**796**	**4.35%**	**1,671**	**9.12%**	**58**
56	**3870**	**LINCOLN FINANCIAL SECURITIES**	**1,174**	**50**	**4.26%**	**102**	**8.69%**	**61**
57	41791	SANTANDER SECURITIES LLC	725	30	4.14%	71	9.79%	50
58	**8032**	**COMMONWEALTH EQUITY SERV**	**2,780**	**112**	**4.03%**	**219**	**7.88%**	**76**
59	13318	SECURITIES SERVICE NETWORK, INC.	477	19	3.98%	46	9.64%	53
60	265	EQUITY SERVICES, INC.	598	22	3.68%	60	10.03%	49
61	**6627**	**AXA ADVISORS, LLC**	**5,474**	**199**	**3.64%**	**486**	**8.88%**	**59**
62	**8158**	**ROBERT W. BAIRD & CO. INC.**	**2,429**	**87**	**3.58%**	**145**	**5.97%**	**105**
63	137115	BMO HARRIS FINANCIAL ADVISORS	477	17	3.56%	38	7.97%	74
64	16443	INVESTMENT CENTERS OF AMERICA	618	22	3.56%	42	6.80%	93
65	**13572**	**CETERA ADVISOR NETWORKS LLC**	**3,145**	**109**	**3.47%**	**226**	**7.19%**	**86**
66	**2882**	**VOYA FINANCIAL ADVISORS, INC.**	**2,806**	**92**	**3.28%**	**232**	**8.27%**	**69**
67	**421**	**WOODBURY FINANCIAL SERVICES**	**1,479**	**48**	**3.25%**	**161**	**10.89%**	**39**
68	**468**	**SIGNATOR INVESTORS, INC.**	**1,651**	**52**	**3.15%**	**142**	**8.60%**	**64**
69	37157	THE LEADERS GROUP, INC.	673	21	3.12%	54	8.02%	72
70	**4031**	**HORNOR, TOWNSEND & KENT, INC.**	**1,194**	**37**	**3.10%**	**93**	**7.79%**	**78**
71	**14251**	**METLIFE SECURITIES INC.**	**7,233**	**223**	**3.08%**	**570**	**7.88%**	**75**
72	**3978**	**LINCOLN FINANCIAL ADVISORS**	**2,436**	**74**	**3.04%**	**172**	**7.06%**	**88**
73	12963	MONEY CONCEPTS CAPITAL CORP.	527	16	3.04%	49	9.30%	57
74	11869	PLANMEMBER SECURITIES CORP.	572	17	2.97%	64	11.19%	37
75	13941	CUNA BROKERAGE SERVICES, INC.	785	23	2.93%	60	7.64%	83
76	4173	ONEAMERICA SECURITIES, INC.	879	25	2.84%	74	8.42%	65
77	35371	INFINEX INVESTMENTS, INC.	609	17	2.79%	51	8.37%	67
78	**5167**	**NYLIFE SECURITIES LLC**	**8,352**	**229**	**2.74%**	**680**	**8.14%**	**70**

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
79	**30349**	**1ST GLOBAL CAPITAL CORP.**	**1,034**	**28**	**2.71%**	**53**	**5.13%**	**113**
80	2936	THE O.N. EQUITY SALES COMPANY	944	25	2.65%	82	8.69%	62
81	**15708**	**PROEQUITIES INC**	**1,514**	**40**	**2.64%**	**121**	**7.99%**	**73**
82	29357	BANCWEST INVESTMENT SERVICES	422	11	2.61%	27	6.40%	99
83	**46173**	**PARK AVENUE SECURITIES LLC**	**2,841**	**74**	**2.60%**	**222**	**7.81%**	**77**
84	**5685**	**PRUCO SECURITIES, LLC**	**5,203**	**135**	**2.59%**	**403**	**7.75%**	**79**
85	**628**	**FIFTH THIRD SECURITIES, INC.**	**1,530**	**39**	**2.55%**	**100**	**6.54%**	**96**
86	**136300**	**KEY INVESTMENT SERVICES LLC**	**1,151**	**28**	**2.43%**	**72**	**6.26%**	**101**
87	43285	M HOLDINGS SECURITIES, INC.	922	22	2.39%	58	6.29%	100
88	**15296**	**SECURIAN FINANCIAL SERVICES**	**1,643**	**39**	**2.37%**	**117**	**7.12%**	**87**
89	42132	CUSO FINANCIAL SERVICES, L.P.	689	16	2.32%	53	7.69%	82
90	**10409**	**MML INVESTORS SERVICES, LLC**	**5,409**	**124**	**2.29%**	**475**	**8.78%**	**60**
91	**15340**	**CETERA INVESTMENT SERVICES**	**1,926**	**44**	**2.28%**	**124**	**6.44%**	**98**
92	128929	GWN SECURITIES INC.	660	15	2.27%	63	9.55%	55
93	**866**	**WADDELL & REED, INC.**	**2,859**	**64**	**2.24%**	**221**	**7.73%**	**80**
94	105108	FIRST REPUBLIC SECURITIES COMP	548	12	2.19%	22	4.01%	130
95	**5393**	**CHARLES SCHWAB & CO., INC.**	**7,616**	**163**	**2.14%**	**402**	**5.28%**	**110**
96	**16164**	**TRANSAMERICA FINANC ADVISORS**	**4,943**	**100**	**2.02%**	**328**	**6.64%**	**95**
97	**1137**	**PRINCOR FINANCIAL SERV CORP**	**3,654**	**71**	**1.94%**	**254**	**6.95%**	**90**
98	**17868**	**U.S. BANCORP INVESTMENTS, INC.**	**1,737**	**33**	**1.90%**	**100**	**5.76%**	**107**
99	33856	BB&T INVESTMENT SERVICES, INC.	753	14	1.86%	34	4.52%	122
100	**519**	**LINCOLN INVESTMENT PLANNING**	**1,130**	**21**	**1.86%**	**78**	**6.90%**	**91**
101	30999	LEGEND EQUITIES CORPORATION	594	11	1.85%	43	7.24%	85
102	**42803**	**VALIC FINANCIAL ADVISORS, INC.**	**1,769**	**30**	**1.70%**	**130**	**7.35%**	**84**
103	**39550**	**CITIZENS SECURITIES, INC.**	**1,126**	**19**	**1.69%**	**70**	**6.22%**	**103**
104	**16986**	**THE HUNTINGTON INVESTM CORP.**	**1,109**	**18**	**1.62%**	**55**	**4.96%**	**116**
105	112630	MWA FINANCIAL SERVICES, INC.	746	12	1.61%	34	4.56%	121

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
106	18464	FORESTERS EQUITY SERVICES, INC.	438	7	1.60%	41	9.36%	56
107	**250**	**EDWARD D. JONES & CO., L.P.**	**17,178**	**270**	**1.57%**	**798**	**4.65%**	**120**
108	**129052**	**PNC INVESTMENTS LLC**	**2,318**	**34**	**1.47%**	**117**	**5.05%**	**114**
109	14455	UNIONBANC INVESTMENT SERVICES	616	9	1.46%	42	6.82%	92
110	**816**	**CREDIT SUISSE SECURITIES (USA)**	**3,850**	**52**	**1.35%**	**87**	**2.26%**	**167**
111	**79**	**J.P. MORGAN SECURITIES LLC**	**25,691**	**340**	**1.32%**	**886**	**3.45%**	**138**
112	2908	NEUBERGER BERMAN LLC	908	12	1.32%	23	2.53%	158
113	1252	WILLIAM BLAIR & COMPANY L.L.C.	936	12	1.28%	27	2.88%	143
114	**11173**	**NATIONWIDE SECURITIES, LLC**	**1,753**	**22**	**1.25%**	**92**	**5.25%**	**112**
115	**19585**	**HSBC SECURITIES (USA) INC.**	**2,317**	**29**	**1.25%**	**87**	**3.75%**	**135**
116	7927	NORTHERN TRUST SECURITIES, INC.	417	5	1.20%	8	1.92%	178
117	**305**	**FORESTERS FINANCIAL SERVICES**	**1,020**	**12**	**1.18%**	**48**	**4.71%**	**118**
118	**2525**	**DEUTSCHE BANK SECURITIES INC.**	**3,697**	**42**	**1.14%**	**98**	**2.65%**	**156**
119	**27060**	**BBVA SECURITIES INC.**	**1,423**	**16**	**1.12%**	**88**	**6.18%**	**104**
120	**10358**	**CETERA FINANCIAL SPECIALISTS**	**1,604**	**18**	**1.12%**	**60**	**3.74%**	**136**
121	13704	PEOPLE'S SECURITIES, INC.	541	6	1.11%	17	3.14%	142
122	**18387**	**THRIVENT INVESTMENT MANAGEM**	**3,702**	**41**	**1.11%**	**163**	**4.40%**	**126**
123	**7059**	**CITIGROUP GLOBAL MARKETS INC.**	**7,136**	**79**	**1.11%**	**198**	**2.77%**	**148**
124	13158	AIG CAPITAL SERVICES, INC.	550	6	1.09%	30	5.45%	108
125	231	MBSC SECURITIES CORPORATION	651	7	1.08%	18	2.76%	149
126	**18272**	**ALLSTATE FINANCIAL SERVICES**	**8,511**	**85**	**1.00%**	**555**	**6.52%**	**97**
127	**13686**	**H.D. VEST INVESTMENT SECURITIES**	**4,819**	**47**	**0.98%**	**217**	**4.50%**	**123**
128	28832	JANUS DISTRIBUTORS LLC	412	4	0.97%	10	2.43%	163
129	**10111**	**PFS INVESTMENTS INC.**	**18,510**	**176**	**0.95%**	**927**	**5.01%**	**115**
130	**17358**	**M&T SECURITIES, INC.**	**1,398**	**13**	**0.93%**	**39**	**2.79%**	**147**
131	134	CANTOR FITZGERALD & CO.	646	6	0.93%	34	5.26%	111

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
132	45744	CAPITAL ONE INVESTING, LLC	544	5	0.92%	34	6.25%	102
133	**7870**	**TD AMERITRADE, INC.**	**3,514**	**31**	**0.88%**	**165**	**4.70%**	**119**
134	46106	FORESIDE FUND SERVICES, LLC	798	7	0.88%	21	2.63%	157
135	104474	SANFORD C. BERNSTEIN & CO., LLC	928	8	0.86%	18	1.94%	177
136	154431	T3 TRADING GROUP, LLC	571	4	0.70%	40	7.01%	89
137	28519	FIRST TRUST PORTFOLIOS L.P.	445	3	0.67%	11	2.47%	161
138	16853	ALPS DISTRIBUTORS, INC.	688	4	0.58%	28	4.07%	129
139	**2881**	**NORTHWESTERN MUTUAL INVEST**	**9,152**	**52**	**0.57%**	**361**	**3.94%**	**131**
140	**20472**	**TIAA-CREF IND & INS SERV**	**4,006**	**22**	**0.55%**	**158**	**3.94%**	**132**
141	**2347**	**JEFFERIES LLC**	**1,686**	**9**	**0.53%**	**31**	**1.84%**	**180**
142	611	MUTUAL OF OMAHA INVESTOR SERV	765	4	0.52%	51	6.67%	94
143	**7784**	**FIDELITY BROKERAGE SERVICES**	**14,007**	**71**	**0.51%**	**465**	**3.32%**	**139**
144	25900	AXA DISTRIBUTORS, LLC	407	2	0.49%	18	4.42%	125
145	**19714**	**BARCLAYS CAPITAL INC.**	**3,470**	**17**	**0.49%**	**58**	**1.67%**	**185**
146	**29106**	**E*TRADE SECURITIES LLC**	**1,650**	**8**	**0.48%**	**69**	**4.18%**	**128**
147	**8206**	**SCOTTRADE, INC.**	**2,069**	**10**	**0.48%**	**77**	**3.72%**	**137**
148	840	COLUMBIA MANAGEM INVESTM DIST	418	2	0.48%	8	1.91%	179
149	42405	EVERCORE GROUP L.L.C.	664	3	0.45%	12	1.81%	181
150	566	KEYBANC CAPITAL MARKETS INC.	677	3	0.44%	15	2.22%	170
151	**7110**	**NATIONWIDE INVESTMENT SERV**	**2,086**	**9**	**0.43%**	**49**	**2.35%**	**165**
152	15356	MUTUAL OF AMERICA LIFE INSURAN	464	2	0.43%	13	2.80%	146
153	**129035**	**USAA FINANCIAL ADVISORS, INC.**	**1,673**	**7**	**0.42%**	**46**	**2.75%**	**151**
154	665	PIPER JAFFRAY & CO.	757	3	0.40%	19	2.51%	159
155	109064	LEGG MASON INVESTOR SERVICES	525	2	0.38%	7	1.33%	197
156	6247	AMERICAN FUNDS DISTRIBUTORS	539	2	0.37%	12	2.23%	169
157	36368	MACQUARIE CAPITAL (USA) INC.	554	2	0.36%	11	1.99%	176
158	128351	SG AMERICAS SECURITIES, LLC	905	3	0.33%	12	1.33%	198

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
159	**107622**	**METLIFE INVESTORS DISTRIBUTION**	**1,523**	**5**	**0.33%**	**60**	**3.94%**	**134**
160	**126292**	**WELLS FARGO SECURITIES, LLC**	**3,073**	**10**	**0.33%**	**51**	**1.66%**	**187**
161	**103863**	**FARMERS FINANCIAL SOLUTIONS**	**6,542**	**21**	**0.32%**	**258**	**3.94%**	**133**
162	**145**	**LINCOLN FINANCIAL DISTRIBUTOR**	**1,331**	**4**	**0.30%**	**43**	**3.23%**	**141**
163	21570	PRUDENTIAL ANNUITIES DISTRIBUTORS	713	2	0.28%	11	1.54%	190
164	11643	HORACE MANN INVESTORS, INC.	746	2	0.27%	33	4.42%	124
165	35350	NYLIFE DISTRIBUTORS LLC	757	2	0.26%	16	2.11%	172
166	38030	MML DISTRIBUTORS, LLC	766	2	0.26%	25	3.26%	140
167	133366	WELLS FARGO FUNDS DISTRIBUTOR	400	1	0.25%	11	2.75%	150
168	7616	COWEN AND COMPANY, LLC	400	1	0.25%	7	1.75%	183
169	4452	PACIFIC SELECT DISTRIBUTORS, LLC	815	2	0.25%	12	1.47%	193
170	**5309**	**FBL MARKETING SERVICES, LLC**	**1,668**	**4**	**0.24%**	**47**	**2.82%**	**145**
171	**361**	**GOLDMAN, SACHS & CO.**	**7,868**	**18**	**0.23%**	**69**	**0.88%**	**209**
172	149823	KCG AMERICAS LLC	442	1	0.23%	9	2.04%	175
173	3641	FIRST COMMAND FINANC PLANNING	900	2	0.22%	24	2.67%	154
174	32205	TRANSAMERICA INVESTORS SECUR	908	2	0.22%	48	5.29%	109
175	**7452**	**VANGUARD MARKETING CORP**	**5,906**	**13**	**0.22%**	**129**	**2.18%**	**171**
176	612	ALLIANZ LIFE FINANCIAL SERVICES	463	1	0.22%	8	1.73%	184
177	37693	ADP BROKER-DEALER, INC.	479	1	0.21%	13	2.71%	152
178	**12060**	**COUNTRY CAPITAL MANAGEMENT**	**2,062**	**4**	**0.19%**	**100**	**4.85%**	**117**
179	15647	PNC CAPITAL MARKETS LLC	525	1	0.19%	5	0.95%	206
180	**40178**	**JACKSON NATIONAL LIFE DISTRIBU**	**1,057**	**2**	**0.19%**	**26**	**2.46%**	**162**
181	154957	PIMCO INVESTMENTS LLC	616	1	0.16%	7	1.14%	203
182	**43036**	**STATE FARM VP MANAGEM CORP.**	**15,325**	**24**	**0.16%**	**649**	**4.23%**	**127**
183	**13109**	**GWFS EQUITIES, INC.**	**2,330**	**3**	**0.13%**	**62**	**2.66%**	**155**

Investor Harm Ranking	CRD	Company Name	Registered Brokers	Investor Harm Brokers	Investor Harm Rate	Brokers With Misconduct	Misconduct Rate	Misconduct Ranking
184	16686	BMO CAPITAL MARKETS CORP.	819	1	0.12%	12	1.47%	194
185	34815	VOYA FINANCIAL PARTNERS, LLC	883	1	0.11%	21	2.38%	164
186	**7654**	**UBS SECURITIES LLC**	**1,885**	**2**	**0.11%**	**30**	**1.59%**	**189**
187	7369	INVESCO DISTRIBUTORS, INC.	960	1	0.10%	24	2.50%	160
188	**7560**	**PERSHING LLC**	**1,077**	**1**	**0.09%**	**18**	**1.67%**	**186**
189	**15794**	**BNP PARIBAS SECURITIES CORP.**	**1,173**	**1**	**0.09%**	**11**	**0.94%**	**207**
190	**18353**	**PRUDENTIAL INVESTM MANAGEM**	**1,276**	**1**	**0.08%**	**23**	**1.80%**	**182**
191	**13041**	**NATIONAL FINANCIAL SERVICES**	**1,345**	**1**	**0.07%**	**28**	**2.08%**	**173**
192	**38642**	**BLACKROCK INVESTMENTS, LLC**	**1,478**	**1**	**0.07%**	**19**	**1.29%**	**199**
193	**8209**	**MORGAN STANLEY & CO. LLC**	**4,028**	**2**	**0.05%**	**39**	**0.97%**	**205**
194	8099	W&S BROKERAGE SERVICES, INC.	445	0	0.00%	26	5.84%	106
195	5633	TD AMERITRADE CLEARING, INC.	566	0	0.00%	16	2.83%	144
196	40638	GUGGENHEIM SECURITIES, LLC	406	0	0.00%	11	2.71%	153
197	**8348**	**T. ROWE PRICE INVESTMENT SERV**	**1,801**	**0**	**0.00%**	**41**	**2.28%**	**166**
198	37404	MERCER ALLIED COMPANY, L.P.	754	0	0.00%	17	2.25%	168
199	5249	JOHN HANCOCK DISTRIBUTORS LLC	883	0	0.00%	18	2.04%	174
200	17344	FIRST CLEARING, LLC	605	0	0.00%	10	1.65%	188
201	7834	OPPENHEIMERFUNDS DISTRIBUTOR	843	0	0.00%	13	1.54%	191
202	17437	AMERICAN CENTURY INVESTM SERV	460	0	0.00%	7	1.52%	192
203	102920	J.P. MORGAN INST INVESTM	842	0	0.00%	12	1.43%	195
204	18476	TD SECURITIES (USA) LLC	437	0	0.00%	6	1.37%	196
205	19647	MIZUHO SECURITIES USA INC.	557	0	0.00%	7	1.26%	200
206	**6271**	**SUNTRUST ROBINSON HUMPHREY**	**1,125**	**0**	**0.00%**	**14**	**1.24%**	**201**
207	17507	FIDELITY INVESTM INST SERV COMP	512	0	0.00%	6	1.17%	202
208	17708	HOULIHAN LOKEY CAPITAL, INC.	682	0	0.00%	7	1.03%	204
209	19899	SECU BROKERAGE SERVICES, INC.	535	0	0.00%	5	0.93%	208
210	**4297**	**NOMURA SECURITIES INTERNAT**	**1,000**	**0**	**0.00%**	**7**	**0.70%**	**210**