INFORMATION NETWORKS: EVIDENCE FROM ILLEGAL INSIDER TRADING TIPS

KENNETH R. AHERN[†]

ABSTRACT

This paper exploits a novel hand-collected dataset to provide a comprehensive analysis of the social relationships that underlie illegal insider trading networks. I find that inside information flows through strong social ties based on family, friends, geographic proximity, and ancestry. On average, inside tips originate from corporate executives and reach buy-side investors after three links in the network. Inside traders earn prodigious returns of 35% over 21 days, with traders farther from the original source earning lower returns, but higher dollar gains. More broadly, this paper provides some of the only evidence on information networks that employs direct observations of person-to-person communication.

This version: 5 February 2015

JEL Classification: D83, D85, G14, K42, Z13

Keywords: Illegal insider trading, information networks, social networks

[†] University of Southern California, Marshall School of Business, 3670 Trousdale Parkway, Los Angeles, CA 90089. E-mail: kenneth.ahern@marshall.usc.edu.

For useful comments, I am grateful to Harry DeAngelo, Diane Del Guercio, Chad Kendall, Paul Tetlock, and seminar participants at Erasmus University, McGill University, Tilburg University, University of Oregon, and University of Southern California.

Information Networks: Evidence from Illegal Insider Trading Tips

ABSTRACT

This paper exploits a novel hand-collected dataset to provide a comprehensive analysis of the social relationships that underlie illegal insider trading networks. I find that inside information flows through strong social ties based on family, friends, geographic proximity, and ancestry. On average, inside tips originate from corporate executives and reach buy-side investors after three links in the network. Inside traders earn prodigious returns of 35% over 21 days, with traders farther from the original source earning lower returns, but higher dollar gains. More broadly, this paper provides some of the only evidence on information networks that employs direct observations of person-to-person communication.

In March 2007, a credit analyst at UBS learned through his job that the private equity firm, Hellman & Friedman, would acquire the software company, Kronos. On March 14, the UBS analyst tipped this information to his friend, Deep Shah, an analyst at Moody's. On the same day, Shah tipped the information to his roommate's cousin, Roomy Khan. Khan then tipped two former business associates: Jeffrey Yokuty and his boss, Robert Feinblatt; and two friends: Shammara Hussain and Thomas Hardin. On March 19th, Hardin tipped his friend, Gautham Shankar, who tipped Zvi Goffer, David Plate, and unidentified traders at the investment firm Schottenfeld Group. Plate subsequently tipped others at Schottenfeld and Goffer tipped his long-time friend, Joseph Mancuso. After the acquisition was officially announced on March 23, the group of inside traders had realized ill-gotten gains of $2.9 million in nine days.

Figure 1 shows that these insiders are a small part of a larger network of 50 inside traders centered around Raj Rajaratnam, the former hedge fund manager of the Galleon Group. In turn, this network is just one of many networks of inside traders. Who are inside traders? How do they know each other? What type of information do they share, and how much money do they make? Existing research provides few answers to these basic questions. Yet, illegal insider trading is an important component of the stock market. Augustin, Brenner, and Subrahmanyam (2014) suggests that 25% of M&A announcements are preceded by illegal insider trading. Similarly, the U.S. Attorney for the Southern District of New York believes that insider trading is "rampant" (Frontline, 2014). More broadly, a better understanding of illegal insider trading might provide insight into how social relations influence stock trading in general (Hong and Stein, 1999).

To provide answers to these fundamental questions, this paper provides a comprehensive analysis of 183 insider trading networks. I identify networks using hand-collected data from all of the insider trading cases filed by the Securities and Exchange Commission (SEC) and the Department of Justice (DOJ) between 2009 and 2013. The case documents are highly detailed. They include biographical information on the insiders and descriptions of their social relationships, such as family, friends, and business associates. They also include the specific information that is shared, the date the information is shared, the amount and timing of insider trades, and the types of securities traded. To address selection bias, I collect a sample of counterfactual observations from the massive LexisNexis Public Records Database (LNPRD). This sample includes insiders' broader social networks of family

members, neighbors, and associates, including people not named as insiders in the case documents. The data cover 1,139 insider tips shared by 622 insiders who made an aggregated $928 million in illegal profits. In sum, the data assembled for this paper provide an unprecedented view of how investors share material, nonpublic information through word-of-mouth communication.

This paper's main objective is to present a series of facts about illegal insider trading. First, I present a profile of individual traders, the events on which they trade, the firms that are the subject of the information, and the traders' investment returns. Second, I investigate the social relationships that connect inside traders. Third, I analyze the flow of information from the original source to the final tippee. Fourth, I analyze the network structure of insider trading rings. Finally, I test for selection bias.

First, the data show that insiders share information about specific corporate events that have large effects on stock prices. Merger-related events account for 51% of the sample, followed by earnings-related events, accounting for 26%. The remaining events include clinical trial and regulatory announcements, sale of new securities, and operational news such as CEO turnovers. The firms in which insiders trade tend to be large high-tech firms with a median market equity of $1 billion. The original sources of inside information are varied, including both internal employees and external sources, such as employees of law firms and investor relations firms. In addition, a significant fraction of leaks are secretly misappropriated from a friend or family member.

Trading in advance of these events yields large returns. Across all types of events, the average stock return from the date of the original leak to the official announcement of the event is 34.9% over an average holding period of 21 trading days. Clinical trials generate the largest average gains at 101% in 9 trading days. M&As generate average returns of 43% in 31 trading days. Earnings generate relatively smaller returns of 14% in 11 days.

Insider trading networks involve a wide array of people. The average inside trader is 43 years old and about 10% of insiders in the sample are women. The most common occupation among insiders is top executive, including CEOs and directors, accounting for 17% of known occupations. There are a significant number of buy side investment managers (10%) and analysts (11%), as well as sell side professionals, such as lawyers, accountants, and consultants (10%). The sample also includes non-"Wall Street" types, such as small business owners, doctors, engineers, and teachers.

The median inside trader invests about $200,000 per tip, though some invest as little as a few thousand dollars, and others invest hundreds of millions. For these investments, traders earn about $72,000 per tip at the median. Using home prices as a proxy for wealth, insiders are among the wealthiest people in the nation, with median home values three times the national median. Compared to next-door-neighbors identified in the LNPRD, insiders have fewer family members, more non-family associates, and are much more likely to have prior criminal records.

The second set of results documents that inside traders share strong social connections. Of the 461 pairs of tippers and tippees in the sample, 23% are family members, 35% are friends, and 35% are business associates, including pairs that are both family members and business associates. Sibling and parental relations are the most common type of family connections. Of business associates, about half of the relationships are between a boss and a subordinate or client and agent. Across the whole sample, 74% of pairs of insiders met before college and 19% met during college. Excluding family members, about 43% of pairs met during college. These results provide direct evidence on the validity of using a common educational background as a proxy for current information flows, as in Cohen, Frazzini, and Malloy (2010).

Inside traders are connected in other ways, too. First, insiders live close to each other. The median distance between a tipper and his tippee is 26 miles. This finding validates the use of geographic proximity as an instrument for social interaction (e.g., Brown, Ivković, Smith, and Weisbenner, 2008). Second, women are more likely to tip and be tipped by other women. Third, insiders are more likely to share tips with people who share a common surname ancestry. Finally, information tends to flow from subordinates to bosses, from younger tippers to older tippees, and from children to parents. These patterns suggest that social hierarchies may influence how information flows among market participants.

To mitigate selection bias in the connections between insiders, I create a sample of counterfactual tippees. For each tipper in the sample, I use the LNPRD to identify a broader set of insiders' family members and associates, beyond those listed as insiders in the SEC and DOJ documents. This counterfactual sample allows me to investigate why some people receive tips and others do not. Controlling for fixed effects of the tipper, I find that insiders tend to share information with people that are closer in age and of the same gender. Insiders are less likely to tip family members

compared to non-family associates, though there is variation among family relationships: insiders are more likely to tip fathers and brothers than mothers and sisters, regardless of the gender of the tipper. To my knowledge, this is the first paper that identifies a counterfactual social network among traders, and is the first paper to exploit the social network data in the LNPRD.

The third set of results in this paper exposes how information flows across a network of traders. I find that as information diffuses away from the source, top executives and mid-level managers are less likely to send or receive tips. Instead, after three degrees of separation from the original source, buy side managers and analysts account for the majority of information sharing. The first tippees are more likely to be friends and family, but as the information diffuses further from the source, business links become more prevalent. People further from the source invest larger amounts, make smaller percentage returns, and earn larger dollar gains. The speed of information also increases as it moves further from the source.

Finally, the last set of facts documents the structure of the networks of inside traders. Of 183 insider networks in the sample, 59 contain only one person. These are original sources who trade on inside information, but do not tip anyone else. On the other end of the spectrum, the network surrounding SAC Capital has 63 members. In the cross-section of networks, larger networks are less dense with fewer clusters of links. Information networks sprawl outward, like a tree, rather than through one central node, as in a star network. Larger networks have younger members and fewer women who are more likely connected through business relationships.

Studying information networks using data from illegal insider trading cases offers both advantages and limitations. The primary advantage is the credibility and level of detail provided in the case documentation. To support an accusation of illegal insider trading, the SEC and DOJ must provide convincing evidence that information is transmitted and that insiders traded based on the information. This means I don't need to rely on an instrumental variable for social relations — the data are direct observations of word-of-mouth communication. A secondary advantage is that the case documents provide the identities of the insiders. This allows me to trace the information from person to person. It also allows me to match individuals to outside data sources, unlike most data on individual traders, such as that of Barber and Odean (2000).

The primary limitation of studying illegal activity is selection bias. Because the sample only includes traders who were caught by regulators, the sample might not be representative of the average inside trader. I address selection bias in a number of novel ways. First, as mentioned above, I use counterfactual observations of relatives and associates not identified by regulators as potential tippees. Second, I proxy for selection bias using variation in the degree of prosecution among traders that are caught. Some insiders named in the legal filings are not charged as defendants, others are charged only in a civil case, and others are charged in both civil and criminal cases. Under reasonable assumptions, I show that the degree of prosecution is correlated with the likelihood of getting caught by regulators. Using the degree of prosecution as a proxy, I find that the sample tends to omit infrequent, opportunistic traders who make smaller investments, share information with family and friends, and are older and more likely female. These results highlight that the sample comprises traders that are more likely to actually impact stock prices: wealthy CEOs and fund managers who are likely to be in larger networks and invest larger sums.

The central contribution of this paper is to provide the most detailed description of illegal insider trading to date. The most closely related paper is Meulbroek (1992), which uses SEC cases from the 1980s to show that insider trading affects takeover prices, which is further tested in subsequent research (Meulbroek and Hart, 1997; Chakravarty and McConnell, 1999; Fishe and Robe, 2004). Other papers study whether enforcement of insider trading laws affects financial markets (Bhattacharya and Daouk, 2002; Bushman, Piotroski, and Smith, 2005; Del Guercio, Odders-White, and Ready, 2013). In contrast, this paper focuses on the flow of information through the social connections of traders. Bhattacharya (2014) provides an overview of the literature on insider trading.

By documenting the flow of information through social connections, this paper also contributes to a broader research agenda on social interactions in finance. Most directly, this paper contributes to the field of information networks. Theoretical models predict that the structure of information networks affect price informativeness, liquidity, and trading strategies (Colla and Mele, 2010; Ozsoylev and Walden, 2011; Walden, 2013; Han and Hirshleifer, 2013; Han and Yang, 2013). However, apart from Ozsoylev et al. (2014), which uses correlated trades to infer social connections, there is little existing empirical evidence on information networks.

More generally, this paper sheds light on theories of private information sharing through non-market social interactions (Hong and Stein, 1999). While supporting empirical evidence relies on imperfect proxies for social interaction, such as geographic proximity (Hong, Kubik, and Stein, 2005; Brown, Ivković, Smith, and Weisbenner, 2008) and common educational backgrounds (Cohen, Frazzini, and Malloy, 2010), this paper provides some of the first direct evidence of actual communication between individual investors.

I. Legal Environment

According to the Securities and Exchange Commission (SEC), insider trading refers to "buying or selling a security, in breach of a fiduciary duty or other relationship of trust and confidence, while in possession of material, nonpublic information about the security."[1] Under U.S. law, insider trading is both a crime punishable by monetary penalties and imprisonment and a civil offense requiring disgorgement of illegal profits and payment of civil penalties. Criminal offenses are charged by the DOJ and civil offenses are charged by the SEC. Civil and criminal charges can be made at the same time for the same offense. Criminal charges are less common, because criminal law requires evidence of guilt beyond a reasonable doubt in order to convict someone of a crime. In contrast, civil cases only require evidence of guilt based on the preponderance of the evidence.

Prosecution of illegal insider trading usually falls under Rule 10b-5 of the Securities Act of 1934. Whether a trade is covered by Rule 10b-5 is based on two theories. The classical theory applies to corporate insiders that purchase or sell securities on the basis of material, nonpublic information. Insiders include both employees of the firm and others who receive temporary access to confidential information, such as externally hired lawyers and accountants. The misappropriation theory applies to anyone who uses confidential information for gain in breach of a fiduciary, contractual, or similar obligation to the rightful owner of the information (typically the firm).

Most practitioners agree that the definition of illegal insider trading is not well defined in the U.S. This paper's findings are directly related to the most recent debate over its definition. In December 2014, the a U.S. Court of Appeals overturned the convictions of Todd Newman and Anthony Chiasson, based in part on the fact that they were several steps removed from the original

[1]From the SEC's website: http://www.sec.gov/answers/insider.htm.

source. As I show in this paper, many of the insiders named in court documents are many steps removed from the original source. Moreover, those that are many steps removed tend to be buy side portfolio managers who make large investments based on inside information. Based on my evidence, the December 2014 ruling is likely to lead to appeals for many of the convictions of serial inside traders. For more detail on the legal environment of insider trading see King, Corrigan, and Dukin (2009).

II. Data Sources

A. Legal Documents

The primary sources of data in this paper are legal documents filed by the SEC and the DOJ as part of illegal insider trading cases. To identify SEC cases, I record the titles of all of the cases reported in the SEC's annual summaries of enforcement actions, "Select SEC and Market Data," for each fiscal year between 2009 and 2013, the most recent publication date. Because some cases involve multiple SEC violations, an insider trading case could be categorized in a different section of the "Select SEC and Market Data" publication. Therefore, I also search in Factiva for SEC publications that include the text "insider trading." This search finds seven cases that involve insider trading that are categorized as Investment Advisers/Investment Companies or Issuer Reporting and Disclosure cases. The Factiva search also identifies cases filed prior to fiscal year 2009 which were amended after 2009, and new cases filed during calendar year 2013, but after the end of fiscal year 2013. I include all of these cases in the sample. I drop 26 cases of fraud in which insiders release false or misleading information, such as pump and dump cases. These cases are fundamentally different because the insider wants to broadcast false information as widely as possible, whereas in the sample cases, the insider shares factual information locally. I also drop nine SEC cases that do not name specific individuals. These cases are based on suspicious trading activity, typically from an overseas trading account, in which the SEC accuses "one or more unknown purchasers of securities." All cases were filed in calendar years 2009 to 2013, except one case that was filed in December 2008. I include this case because the DOJ case was filed in 2009.

Unlike the SEC, the DOJ does not provide summary lists of all the insider trading cases it brings. Therefore, I search Factiva for all DOJ press releases with the words "insider trading" and record the name of the case from the press release. In the sample, only three cases filed by the DOJ are not also charged by the SEC.

I use a number of sources to collect the original case documents. First, I search the SEC's website for official documents, the most useful of which is a civil complaint. Complaint filings typically include a detailed narrative history of the allegations, including biographies of defendants, trading records, and descriptions of the relationships between tippers and tippees to justify the allegations. Some cases are not available on the SEC web page. For these cases and for all DOJ cases, I search for the case documents using Public Access to Court Electronic Records (PACER). The most useful DOJ documents are the criminal complaints and "information" documents. These are similar to civil complaints, but contain less information. Transcripts of hearings, while potentially informative, are typically not available on PACER.

This search procedure yields 336 primary source documents comprising 5,440 pages. Since the documents provide data in narrative histories, the data must be read and recorded by hand. Reading the individual cases in detail is also necessary to sort out the identities of all of the insiders named in the cases. In particular, in many DOJ documents, co-conspirators remain anonymous. Some co-conspirators' identities are revealed in future cases, but other co-conspirators' identities are never revealed by the DOJ. However, these same people are often named in the SEC documents. Therefore, it is necessary to read all of the cases and their amendments in order to piece together the identities of as many people as possible. For instance, in many cases it is easy to infer who the co-conspirator is by the description of their job and relationship to the defendant in connection with another DOJ case in which the co-conspirator is the named defendant. In some instances, the identities of certain insiders are never revealed. In these cases, I rely on investigative journalism in media reports that uncovers the identities of people that the SEC and DOJ do not name explicitly.

From the primary source documents, I record five key types of information. First, I record the names, locations, employers, and ages of all people named in the document. In many cases, the documents refer to people who are not officially charged by the SEC or DOJ, but who the government alleges were privy to inside information. I include these people in the data as well.

Second, I record the social relationships between tippers and tippees, including family relations, friendships, and co-worker relationships. Third, I record the original source of the information and how the source received the information. For instance, the documents might explain that a lawyer was assigned to work on an upcoming acquisition. Fourth, I record the timing of information flows. These include the days when tippers and tippees communicated in person, by phone, or electronically. In some cases, the documents record the timing of phone calls to the minute. Finally, I record detailed records of trading behavior, including the dates of purchases and sales, the amount purchased, the types of securities purchased (e.g., shares or options), and the profits from the sales.

B. LexisNexis Public Records Database

The second major source of data in this paper is the LexisNexis Public Records Database (LNPRD). The LNPRD includes a wide array of biographical information, including age, address, real property ownership, employers, licenses, liens, bankruptcies, and criminal records for over 300 million people who reside in the US, whether living or dead. Because the case documents include name, age, and location, I am able to identify people named in the filing documents with a high degree of accuracy. For instance, in one case, the SEC complaint only states, "Richard Vlasich is a friend of Michael Jobe and resides in Fort Worth." I find the entry for Vlasich on LNPRD using his name, city, and approximate age. The LNPRD data states that his employer is Vlasich Associates. Using this information, I find his resume on an online professional networking site, which describes his role in Vlasich Associates as the owner of a small real estate business.

I also record data on family members and person associates from the LNPRD. The specific type of familial relation is not identified in the LNPRD, though the data do indicate first or second degree connections and through whom the connections run. For example, using married and maiden names, I identify wives as women who are roughly the same age as the insider that also share a history of common addresses and own property jointly. The second degree connections through a wife could include the wife's parents or siblings, which I can identify by age, surname, and address of the second degree relatives. If a familial relationship can not be identified, I record it as unknown. Person associates are non-family members with whom an insider shares a relationship. The exact algorithm for identifying person associates is proprietary to LexisNexis using their vast public and

private records database. In general, these are people with whom an insider may have shared an address, had business dealings, or is connected in some other way through primary source records. The Internet Appendix provides more detail on the LNPRD.

C. Additional Sources of Data

Not all documents contain all information. In particular, the job titles of many people are not listed. To find occupations, I search online professional social networking sites. Using the reported employer and location helps to identify the particular person on these sites. However, because people charged with insider trading often wish to hide their connection with the illegal trading charges, they may not list their old employer on online resumes. In these cases I use the employment records in the LNPRD. However, the LNPRD doesn't report job titles. To overcome this obstacle, I use the Internet Archive's "Wayback Machine" to search company websites on dates before the insider trading charges were filed to identify job titles and other biographical information. Finally, I use web searches to try to find any remaining data.

In addition to the case documents and the LNPRD, I estimate home values as a proxy for wealth using estimates from the online real estate website, Zillow.com. I also record the gender of every person in the dataset based on the person's first name. For unfamiliar names, I rely on two databases: namepedia.org and genderchecker.com. Finally, I record the ancestry of every insider's surname using the Onomap database.

III. Illegal Insider Trading Events and Firms

A. Events

Table I presents the time series of the 465 events in the sample. The earliest event is 1996, and the most recent is 2013, though 89% of the events occur within 2005 to 2012. The cases that involve insider trading in the earlier periods typically concern a defendant that is charged with a long-running insider trading scheme. There are 25 events for which I cannot identify an announcement date because the SEC and DOJ documents do not specify a specific event.

Table II presents statistics on the frequency of different types of events, stock returns, and holding periods surrounding insider trading. Panel A provides the frequency of six types of events. A detailed breakdown of the types of events is presented in Internet Appendix Table I. The most common type of event with 239 instances, or 51.4% of all events, is a merger or acquisition (M&A). The large majority of these events are acquisitions, though I also include 12 joint ventures, licensing agreements, strategic alliances, and restructuring events, plus eight events related to developments in merger negotiations, such as the collapse of a deal. Of the 219 acquisition events, informed investors traded in the target's stock in 216 cases, and the acquirer's stock in just three cases.

The next most common type of event in the sample is earnings-related events, with 123 events, or 26.4% of the sample. The large majority of these events (112 events) are regularly scheduled earnings announcements. The rest of the earnings-related events are announcements of earnings restatements and earnings guidance.

The remaining 22% of the sample comprises drug clinical trial and regulatory announcements (8.0%), the sale of securities (7.5%), general business operations, such as the resignation or appointment of a senior officer, employee layoffs, and announcements of new customer-supplier contracts (2.8%), and other announcements (3.9%), such as analysts reports, dividend increases, and the addition to a stock index. The other events category also contains events that are not specified in the SEC and DOJ documents. All but two of the sale of securities events involve private investment in public equity transactions, with the vast majority for Chinese firms traded in the United States.

Table II also distinguishes whether the inside information contains positive or negative news at the time the information is tipped. I base this distinction on the trading patterns of tippees: long positions and call options indicate positive news and short positions and put options indicate negative news. Almost all M&A announcements are positive news events (234 vs. 5). Earnings events are more evenly split between positive and negative news, with 66 positive events and 54 negative events. Clinical trials tend to be positive news events with 24 positive events compared to 13 negative events. Sale of securities is overwhelming bad news in the sample with 34 negative events and one positive event. Overall, there are 335 positive events and 112 negative events. In 18 cases, the details provided in the SEC complaints do not provide enough detail to classify an event as positive or negative, including the 12 cases in which the events themselves are not specified.

B. Stock Returns from Insider Trading

Panel B of Table II presents average stock returns for each event type. The stock returns are calculated as the return from buying stock on the first trading date after the original tipper first receives the information through the date of the corporate event. If the date that the original source receives the information is not available in the filings documents, I use the first date that the original source tips the information. Panel C presents averages of this holding period. Not all tippees earn returns equal to those in Table II because many tippees do not receive the information until closer to the event date and some do not trade stock. The final column of Table II aggregates stock returns by taking the average of the returns for a long position in positive events and a short position in negative events.

Stock returns from insider trading are large by any measure: on average, trading on inside information earns returns of 34.9% over 21.3 trading days. Because the returns are based on idiosyncratic inside information, there is virtually no financial risk to the strategy. There is of course, legal risk. Also, median trading periods are small: 16 days for M&A events and 10 days for all types of events.

Clinical trial and drug regulatory announcements generate the largest returns, on average, with gains of 101.2% for the 24 positive events and −38.6% for the negative events, with an average holding period of just 9.2 days. Next, M&As generate average returns of 43.1% in 30.5 days. Insider trading based on earnings announcements generates relatively smaller returns of 13.5% in 11.3 days.

The holding period averages reported in Panel C of Table II reveal that the time between learning about a future event and the actual event date is shorter for negative events than positive events. In untabulated t-tests, I find that the average holding period is longer for positive news than negative news in clinical trial events (p-value= 0.013), M&As (p-value= 0.002), operations (p-value= 0.087), and events overall (p-value$<$ 0.001). The difference for earnings announcements is not statistically significant. This could imply that firms delay the announcement of good news compared to bad news, or it could imply that good news travels faster than bad news.

C. Firms

Table III presents summary statistics of the firms at the event level. The sample includes 351 firms whose stocks are traded by inside traders and whose data are available on the CRSP-Compustat merged database. The missing observations tend to be small firms or foreign firms, primarily Chinese, which are traded on OTC markets. Because there are 465 events, many of the firms have information tipped about multiple events. For instance, Best Buy Co. has five different earnings announcements in the sample.

The firms in the sample are relatively large firms, though there is wide variation across the sample. The average firm has market equity of $10 billion and the median firm's market equity is $1 billion. As a comparison, the median firm listed on the New York Stock Exchange in December 2011 has a market equity of $1.2 billion. The 25th percentile on the NYSE is $0.41 billion compared to $0.30 billion for sample firms. At the the 75th percentile, NYSE firms have market equity worth $3.45 billion compared to $3.56 billion for sample firms. These distributions suggests that the sample firms include a representative sample of NYSE firms, plus a number of smaller firms in the left tail of the distribution.

The dollar trading volume of larger firms may be attractive for illegal inside traders because they are less likely to affect the stock price through their trades. Table III provides summary statistics of the trading volume of the firms in the sample. The median firm has a daily trading volume of about 680,000 shares and a daily dollar trading volume of $13.08 million. There is wide variation, with some firms having very small dollar trading volumes, and others, some of the largest on the stock market.

To compare the normal trading volume of the sample firms to the illegal trading volume, Table III presents statistics on the total dollar volume of illegal trades per event. I first aggregate the total dollars invested by inside traders over all the days in the period between when the original source receives the information and the event date (21.3 days, on average; 10 days at the median). The total amount traded by tippees is $370,000 for the median event and $4.06 million for the average. I next calculate the total dollar amount of illegal trades divided by the firm's average daily dollar volume during a non-event period. The ratio is 6.57% at the median. Across 10 trading days, this is only 0.66% of the normal trading volume. However, at the 75th percentile, the figures are

34.56% using aggregate numbers and 3.5% at a daily level. This represents a significant fraction of the daily volume for the upper tail of insider trading activity.

Internet Appendix Table II presents the fraction of events by firms' industries. Compared to all firms in the CRSP database, the sample of industries represented in the illegal trading database overweights high-tech industries. Based on the size of the insider trading sample, the distribution of firms across industries in CRSP predicts that there would be about 33 insider trading events for firms in the chemical manufacturing industry, including pharmaceutical firms. In contrast, the sample includes 90 events for chemical manufacturers. Similarly, the CRSP distribution of industries predicts that firms in the computer and electronics manufacturing industry would account for 43 events. In comparison there are 91 in the sample. In contrast, the sample underweights credit intermediaries and bank holding companies.

Using industry distributions for firms in CRSP may be an inaccurate benchmark because a large fraction of the sample involves trading around mergers, which tend to cluster by industry. Industries with many mergers could have more information, not necessarily more information leakage. However, a comparison of the industry distribution of the insider trading sample to a sample of public, US merger targets from 2004 to 2012 from SDC provides a similar result. The SDC data predict that there will be 17 merger events in the sample that involve a firm in the chemicals industry, compared to 33 actual M&A events in the sample. Similarly, software publishers are expected to account for 11 M&A events, compared to 21 in the sample. Finally, credit intermediaries account for 15 M&A events in the insider trading data, but are expected to account for 33 events based on the frequency of actual mergers of credit intermediaries.

IV. The Characteristics of Inside Traders

A. Demographics

There are 622 people in the data set. Of these, 162 people are tippers only, 249 are tippees only, 152 are both tippees and tippers, and 59 are original information sources who do not tip anyone else. Table IV presents summary statistics of the people. Across the entire sample, the average

age is 44.1 years and 9.8% of the people are women. The youngest person is 19 years old and the oldest is 80. The large majority of insiders (92%) are married.

I use an insider's home value as an imperfect proxy for wealth. Because identifying the specific timing of real estate purchases is difficult in the LNPRD, for each insider, I compute his median house value across all of the real estate he owned at any time. I restrict real estate holdings to those that are listed both as real property and as a residence in the LNPRD to filter out investment properties. The average insider's home is worth an estimated $1.1 million in September 2014. The median insider's home is worth $656,300. The national median sales price of single-family homes at the same time is $212,400, which is less than the 25th percentile of insiders' home values.[2] The median insider's home value is roughly equivalent to the median home price in Honolulu, which is the fourth most expensive metropolitan area in the country, out of 175 areas. Thus, the inside traders in the sample tend to be among the nation's wealthiest people.

The average person gives 1.5 tips, which also equals to the number of tips received by the average person since the network of tippers and tippees is closed and every tip is received by someone else in the sample. In untabulated statistics, of the 162 people who only share tips, the average number of tips shared is 2.36. Of the 249 people who only receive tips, the average number of tips received is 1.96. Of the 152 who both give and receive tips, the average person receives 2.99 tips and gives 3.68 tips.[3]

Figure 2 presents the geographic location of the people in the sample who are located in the United States. Insiders are located all across the country, including both urban and rural locations. However, based on either state-level population, aggregate income, or the population of people invested in the stock market (using participation rates from the Health and Retirement Study), the sample of insiders is overweighted in California, New York, and Florida, and underweighted in Texas, Ohio, and Virginia. The sample also includes insiders located in other countries, including 16 people in China, 9 in Canada, 23 in Europe, and others from various countries including Australia, Brazil, Israel, and Thailand.

[2]Data are from the National Association of Realtors.

[3]Internet Appendix Table IV shows that younger insiders with more family members. Internet Appendix Table V presents Poisson regressions which show that that younger tippers having significantly more tippees.

The total amount invested per tippee ranges from a minimum of $4,400 up to a maximum of $375 million. The average total amount invested is $4.3 million and the median amount is $226,000. The median amount invested per event is $200,000, with an average of $1.7 million. These large amounts are explained in part by the composition of the types of people that receive inside information, including portfolio managers, and also because people invest more money in insider trades than they normally would when the investment's return is risky. Many of the SEC complaints document that inside traders sell all of the existing assets in their individual portfolios and borrow money to concentrate their holdings in the insider trading firm. As evidence, the median inside trader invests an amount worth 39% of his median home value.

As the stock returns documented above show, trading on inside information is highly profitable. Across all tips, the median investor realizes a total gain of $133,000 in ill-gotten profits and losses avoided. The average investor realizes gains of $2.3 million. Per tip, the median investor gains $72,000. The average percentage return for inside traders is 63.4% and the median is 26.4%. These returns are higher than the average event returns presented in Table II because some insiders trade stock options in addition to common stock.

B. Inside Traders' Occupations

Table V presents summary statistics of inside traders by nine types of occupations. Internet Appendix Table III provides a detailed breakdown of the frequency of each occupation in the sample. The most common occupation among inside traders is top executive with 107 people. Of these, 24 are board members and the rest are officers. There are 55 mid-level corporate managers and 59 lower-level employees in the sample, including 8 secretaries, 11 information technology specialists, and a few nurses, waiters, and a kindergarten teacher. There are 61 people who work in the "sell side" of Wall Street including 13 accountants, 24 attorneys, 4 investment bankers, and 3 sell side analysts. I divide the "buy side" into two groups by rank in investment firms: there are 60 portfolio and hedge fund managers and 65 lower-level buy side analysts and traders. Small business owners and real estate professionals account for 39 people in the sample and 38 people have specialized occupations, including 16 consultants, 13 doctors, and 9 engineers. There are 135 people for which I cannot identify an occupation.

Age and gender follow well-known patterns across the occupations in the sample. At 50.9 years old, top executives are among the oldest group in the sample, compared to 41.3 for mid-level managers and 41.5 for lower-level employees. Buy side managers tend to be younger than corporate executives at 42.6 years, but older than buy side analysts at 35.5, which are the youngest group in the sample, on average. Women are predominately found among lower-level employees at 25.5% and in the unknown occupation group. Using home values as a proxy for wealth, there are three distinct wealth strata. At the top are buy side managers, followed in the middle stratum by top executives, sell side individuals, and buy side analysts. The remaining occupations account for the lowest wealth stratum.

Buy side managers invest the largest amount per tip (median of $6 million), followed by people who work in the sell side ($3.8 million) and buy side analysts ($2 million). Small business owners invest the least with a median investment of $203,900 per tip. It is interesting to note that the median mid-level manager invests more than the median top executive ($2 million compared to $376,800). This likely represents the higher scrutiny on the investments of top executives. Buy side analysts have the highest median return of 117.7%. Buy side managers have median returns of 37%, among the lowest returns across occupations. However, buy side managers earn the highest dollar gains for their trades at $5.8 million per tip, at the median.

The occupations that give the most insider tips are sell side employees, such as lawyers and accountants, with three tips given, on average, and buy side analysts and traders. Buy side analysts also receive the most tips of any occupation (3.1) on average, and top executives and corporate managers receive the fewest (0.5 and 0.6). I provide more detail on the flow of information by occupation below.

C. Inside Traders Compared to their Neighbors

As a final set of summary statistics, I compare inside traders to their neighbors. For the 448 inside traders that I can identify in the LNPRD, I randomly pick a neighbor of the same gender as the insider. Neighbors are literally next-door-neighbors, as I choose the person that lives on the same street as the insider with the street number as close to the insider's street number as possible. I require the neighbor to have a social security number and date of birth recorded in

the LNPRD. Choosing a comparison sample from the same neighborhood and of the same gender helps to control for wealth, age, and occupation, and highlights the remaining differences between insiders and non-insiders.[4]

The first three columns of Table VI present average characteristics of insiders and their neighbors. Insiders are statistically different than their neighbors in many ways. Insiders have a higher likelihood of owning residential real estate, are more likely to be accountants and attorneys, and significantly less likely to be registered as a Democrat, compared to their neighbors.

I also record financial dealings of insiders and their neighbors, including a dummy for bankruptcy filings, judgments, or liens against a person. For each of these variables, I only include records if they occurred prior to the last inside trade in the sample. This mitigates concerns about reverse causation from prosecution to judgments or bankruptcies. Table VI reveals that inside traders are less likely to declare bankruptcy, but are about twice as likely to have liens and judgments filed against them, compared to their neighbors.

Finally, I record prior criminal records of insiders. As for bankruptcy, I only record a criminal record if it is filed before the last insider trading date in the sample. Criminal records in the LNPRD are at the state level and do not include any instances of insider trading. Of the 246 instances of criminal records, 122 are unspecified crimes, 103 are traffic-related crimes, and 21 are other crimes, including drug charges, assaults, and issuing bad checks. Thus, the overwhelming majority of identifiable criminal charges are for traffic violations.

Table VI shows that insiders are considerably more likely to have a criminal record compared to their neighbors (53.7% versus 12.8%). This supports the notion that illegal insider trading is consistent with other patterns of behavior. Though the greater prevalence of criminal records among inside traders could reflect that they are just more likely to get caught breaking the law than are their neighbors (whether speeding or illegal trading), it seems more likely that insiders have less respect for the rule of law and are more brazen in their illegal activities than their neighbors.

[4]Because many fields of LNPRD data is provided at the state level, missing data (i.e., licensing, criminal records) will be missing at the state level, too. Therefore, using neighbors mitigates any concerns about selectively missing observations.

Finally, using information on family members and associates from the LNPRD, Table VI shows that insiders tend to have the same number of listed family members, but significantly more associates. This could reflect that insiders have large social networks or just have more of the type of business dealings that would be identified in public records and included in the LNPRD. I investigate this issue in greater detail below.

Since many of these characteristics are correlated, columns 4 and 5 of Table VI presents logit regression coefficients on the likelihood that a person is an inside trader, using neighbors as the counterfactual. Column 5 includes insider-neighbor pair fixed effects to account for any fixed characteristics of the neighborhood that could drive the results. The regressions show that insiders tend to be younger, have liens against them, have fewer family members and more associates, and have a criminal record.

D. Summary Description of Insiders

Though there is significant variation across multiple dimensions, in general, the inside traders in the sample are predominately men between the ages of 35 and 50 who work either as top executives or as buy side investors. They tend to be among the wealthiest people in the nation, with residences all across the globe, but predominately in New York, California, and Florida. Compared to their neighbors, they are younger with more social connections and a much higher likelihood of having a prior criminal record. Younger insiders are more likely to tip others, compared to older insiders. Insiders tend to trade on insider knowledge in advance of significant corporate events, such as mergers and earnings announcements, for firms in high tech industries. Their trades earn very large returns of about 35% over a month. These investments make large profits of $1.3 million on average per tip.

V. The Relationships Among Inside Traders

Having established the characteristics of inside traders, I next analyze the connections between them. Understanding how insiders are connected through social relations, geographic proximity, and shared backgrounds can shed light on how private information transmits across insiders.

A. Social Relationships

Table VII presents the prevalence of information flows by the type of relationship between the 461 pairs of tippers and tippees in the sample. Of these relationships, 22.6% are familial, 34.7% are business-related, 35.1% are friendships, and 21.3% do not have any clear relationships. Multiple types of relationships are allowed. In untabulated numbers, 11 pairs of insiders have both familial and business relationships, 56 pairs of relationships are both business-related and friendships, and 4 relationships are both family and friends (typically in-laws and distant family members).

The family relationships are led by siblings (24.0% of familial relationships) and parent-child (19.2%). About 14% of the family relations are between married couples and 11.5% are through in-law relationships. The 'other' category, which includes cousins, uncles, aunt, etc, accounts for 8.7% of family relations. The 'unspecified' category (15.4% of family relationships) includes observations where the SEC and DOJ filings indicate that people in a pair are relatives, but doesn't specify the exact relationship.

Among the business-related relationships, 54.5% are among associates. Business associates are people that work together or know each other through their profession. In comparison, boss-subordinate relationships account for 25.6% of business-related ties and client-provider relationships account for 20%. This means that slightly more than half of business-related relationships are between people of equal status, and half are relationships where one person holds a supervisory role over the other.

For friendship relations, the filings commonly describe relationships as either friends or close friends, and occasionally as acquaintances. In the sample, there are just three pairs described as acquaintances, compared to 115 described as friends, and 44 described as close friends. In untabulated statistics, I verify that the distinction between friends and close friends is meaningful. In some cases, the filings identify when the insider relationship began. Among the 13 pairs of close friends with available data, 77% met before college and 23% met in college. In comparison, among 21 friendship pairs, 29% met before college, 24% met in college, 29% met in graduate school, and 19% met after they had completed their education. This suggests that close friendships were formed during childhood, as might be expected.

The 98 pairs in Table VII in which no social relationship is listed comprises a sizable number of expert networking firm relationships, where insiders are paid consultants to clients of the expert networking firm. Of the 98 pairs, 22 are related to the expert network firm Primary Global Research LLC (PGR). Thus, these are not just missing observations. Instead, these pairs actually have no social relationships other than through sharing inside information.

The data allow me to provide a rough estimate of when relationships formed. In 50 pairs, the filings provide the time when the individuals first met. Among all friendship relations, 44% met before college, 46% met in college or graduate school, and 10% met after completing school. Among business-related pairs, 33% met before college, 33% met during college or graduate school, and 33% met after finishing school. There is little information on when in-laws, engaged, and married couples met, so I exclude them from the calculation, leaving 70 familial pairs, all assumed to have met before college. I also ignore the 98 pairs where no social relation is listed. Across the 124 pairs where I can estimate when the relationship began, I find that 74% met before college, 19% met in college or graduate school, and 7% met after completing their education. Since family relationships account for much of these figures, most people met during childhood. Excluding family relationships, 41% met before college, 43% met in college, and 17% met after completing school. Overall, these patterns imply that people who share inside information have long-standing and close relations with each other, on average. It also implies that the presence of school-ties are related to actual social interactions, as assumed in a number of papers (Cohen, Frazzini, and Malloy, 2010), though information is mostly shared between people who met before college.

B. Geographic Proximity in Relationships

Table VIII presents summary statistics of the geographic distance between insiders. Distance is calculated as the great circle distance in miles between the cities of the people in a pair. Longitude and latitude for each city are taken from Google Maps. Therefore, if two people live in the same city, they have a distance of zero miles.

Across all pairs of insiders, the median distance is 26.2 miles, with an average of 581.1 miles. The maximum distance is from Hong Kong to Schwenksville, Pennsylvania at 8,065 miles. At the median, the geographically closest relationships are familial relations, with a median of 14.3 miles,

followed by business-related relationships with a median of 18.9 miles, and friendship relationships at 28.4 miles. If no social relation is listed, the members of a pair are located substantially farther from each other, at 80.9 miles at the median. In unreported tests, I find no statistically significant difference in the medians of family, business, or friendship relationships. Siblings and children live slightly further away (28 and 26.2 miles at the medians) than do business associates (16.8 miles) and clients (19.0 miles). These statistics suggest that relatively small geographic zones are reasonable proxies for a large fraction of information flows, across all types of interpersonal relationships.

Information flows are not always local, however. The 75th percentile of distance in all insider pairs is 739 miles. From New York City, this would include cities as distant as Chicago and Atlanta. At the 75th percentile for business ties, which are the lowest across the three types of relationships at 220 miles, business relations in New York City would include people in Washington D.C., Philadelphia, and Boston. This means more that a substantial fraction of relationships are distant. The distant geographic relations follow particular patterns. Figure 3 shows the information flows between New York City and the Miami area, between New York and the San Francisco Bay area, and between Southern and Northern California.

C. Directionality in Tipping Relationships

I next test how age, ancestry, occupation, and gender are related to the flow of inside information across tippers. First, Figure 4 provides a heat map of the connections between tippers and tippees by occupation. These relations are based on binary connections between people, not the total number of tips, where a cell entry reports the total number of pairs where the tipper occupation is listed on the row heading and the tippee occupation is listed on the column heading. Unknown occupations are not detailed in the figure, but they are included in the totals for each row and column.

The figure shows that top executives are by far, the most frequent tippers. Their tippees are spread over all occupations. Top executives tip other top executives (15% of top executive pairs), specialized occupations like doctors and engineers (13%), buy side analysts (10%), and managers (11%), and all other occupation categories in the sample. In contrast, buy-side analysts are the

next most common tippers with 88 pairs, but their tippees are concentrated among other buy side analysts (34% of their tippees) and buy side managers (21%).

The heat map provides an indication of the direction of information flow. First, the strong diagonal pattern in the figure shows that tipping relationships tend to concentrate among people in the same occupations. Buy side managers and analysts tend to tip other buy side managers and analysts. Sell side employees tend to tip other sell side employees, and top executives tend to tip other top executives. Comparing the number of pairs in which a particular occupation is a tippee compared to a tipper, shows that top executives are 2.85 times more likely to be a tipper than a tippee. In contrast, buy side managers are tippees roughly twice as often as they are tippers.

Figure 5 presents a similar analysis by age of tipper and tippee. At the aggregate level, there is a strong negative relation between age and the number of tipping relationships for both tippers and tippees. Tippers between the ages of 30 to 34 have the most tipping relations, which then decline as tipper age increases. Similarly, tippees who are between 35 to 39 have the most relations which then decline as tippees get older. The strong diagonal component of the figure reveals that tippers and their tippees tend to be close in age. Of the 64 tippee-tipper age groups in Figure 5, 35% of all pairs are found in the eight groups where tippers and tippees are the same age. In the off-diagonal regions, there are more relationships where younger tippers are sharing information with older tippees than vice versa (108 pairs versus 79 pairs).

Figure 6 presents the prevalence of pairs by surname ancestry. As before, the strong diagonal in the figure reflects that people of the same ancestry tend to tip each other. For example, of the sample of tippers with South Asian surnames, one-third of the tippees also have South Asian surnames, out of nine possible ancestries. Similarly, of tippers with Muslim surnames, about half of their tippees also have Muslim surnames. In unreported tests that exclude family relationships, a similar though weaker pattern emerges where tippers and tippees share a common surname ancestry.

Finally, in untabulated statistics, I find that men and women tend to form tipping relationships within genders. When the tipper is a man, 9% of relationships are with female tippees. Instead, when the tipper is a woman, 13% of relations are with female tippees. I also find that subordinates are more likely to tip supervisors than vice versa. In 63% of boss-subordinate relationships the

tipper is the subordinate, a significant difference from 50%. In 55% of client-provider relationships, the tipper is the provider, which is not significantly different than 50%. These results imply that information flows in both directions equally between clients and their agents, but disproportionately from subordinates to bosses. Similarly, in parent-child relationships, the tipper is the parent in 30% of cases, significantly less than 50% (p-value= 0.072). This means that information tends to flow from child to parent.

D. Counterfactual Observations to Analyze Who Gets Tipped and Who Doesn't

The above results show that within the sample of inside traders, insiders have strong social relationships. However, because all of the people in the sample received inside information at some point, there is no counterfactual observation of someone who could have received a tip, but did not. In this section, I construct a counterfactual sample of potential tippees to better understand how the characteristics of actual tippees differ from potential tippees.

For each insider in the data that I can identify in the LNPRD, I record the family and personal associates listed in the LNPRD, as described above. Then I append to this set any insiders in the SEC and DOJ documents that do not appear in the list of a tipper's family and associates. From this broader set of social relations, I create a dummy variable equal to one if a tippee receives inside information from the tipper.

Table IX provides logit regressions tests on the likelihood of receiving a tip. Columns 3 and 4 include tipper fixed effects to account for all observable and unobservable characteristics that could influence with whom a tipper chooses to share information. Across all specifications, I find similar results. Potential tippees that are close in age and of the same gender as the tipper are more likely to receive inside information. Family members are less likely to be tipped than are associates. However, compared to non-family associates, husbands are more likely to be tipped, while wives, sisters, and in-laws are less likely to be tipped. Brothers and fathers are equally likely to be tipped as associates.

To my knowledge, in the literature on social networks in finance, this analysis is the first to provide a counterfactual sample of potential tippees. However, it is not without limitations. First, the list of potential tippees that I record from LNPRD is certainly not exhaustive. Family members

are capped at 10 people and associates are only identified through records available to LexisNexis. Second, it is likely that a tipper shares information with other people than just those named in the legal documents. These people may appear on the list of potential tippees from the LNPRD, but I would not correctly identify them as an actual tippee. I believe these limitations will only provide noise in the estimation, not bias, and that these tests provide important evidence on the nature of insider relationships.

E. Summary of Pair-Level Relationships

The statistics in this section reveal a detailed picture of tipper-tippee relations. Tippers and tippees are not randomly matched. Instead, they tend to have strong social connections through family, friends, and professional contacts, many of which are formed during childhood. When parents and children tip each other, information is more likely to flow from children to parents. When bosses and subordinates share information, the information typically flows from the subordinate to the boss. Similarly, tippers and their tippees tend to be close in age, but tippers tend to be younger than their tippees. People also tend to tip other people of similar ancestry, gender, and occupation. These relationships are supported in tests that include a counterfactual set of potential tippees. Finally, tippers and tippees live close to each other, suggesting that face-to-face social interaction is associated with information sharing.

VI. The Diffusion of Information Across Insider Networks

In this section of the paper, I trace out the path through which information flows from the original source to the final tippee.

A. Original Sources

The original source of an insider trading event differs based on the type of event. Table X presents statistics on the original sources of M&A events.

The original source in an M&A leak obtains the information through connections with the acquirer, the target, or a third party. Of the total original sources in M&As, 99 (35%) original

sources are associated with the acquirer, 119 (42%) with the target, and the remaining 67 (23%) with a third party or an unknown party. Sources can also be classified as internal or external. Internal sources include employees of the acquirer or target. External sources include employees of firms that contract with the acquirer or bidder, such as accounting firms, investor relations firms, investment banks, and law firms. In M&As, 117 of the 285 original sources (41%) are internal sources, compared to 164 external sources (58%). Internal acquirer leaks are evenly split between officers and lower-ranked employees, with no leaks from acquirer directors. In contrast, target officers are responsible for 48% of leaks, target directors for 29%, and lower-ranked employees for 23% of internal leaks from the target.

The most common type of firm where external leaks occur in M&As is a law firm, accounting for 40% of all external sources. Investment banks are the next most common (23%), followed by accounting firms (13%). For many external sources, there is an even balance between sources associated with the target and those associated with bidder. However, accounting firms working for an acquirer account for 2.5 times as many sources as accounting firms working for the target. In contrast, target law firms account for twice as many leaks as acquirer law firms. Finally, a substantial number of leaks are stolen (about 18% of all external sources). These are cases where a friend or family member secretly accesses information. For instance, a merger attorney's father misappropriated confidential information from his daughter's work documents while she visited him during holidays (see SEC v. Dean A. Goetz). These statistics reveal that insider news about M&A come from a wide range of sources, both internal and external.

Tips about earnings announcements are either internal or external to the announcing firm. These data are presented in the last column of Table X. In contrast to M&As, lower ranked employees of the firm are the most common internal source of the information. These employees account for 52% of all internal sources, compared to directors, which account for 5%. Another difference between M&A sources and earnings sources is that earning sources are predominately internal (67%) compared to external (31%). Not surprisingly, the most common external source is an accounting firm, followed by investor relations firms. There are five tips in the sample where an employee at the Market Intelligence Desk of the NASDAQ received advanced notice of earnings releases for NASDAQ listed firms.

Internet Appendix Table VI presents the original source of information in the remaining event types. Most of these sources are external to the firm. In particular, in 34 events, an employee of a potential investor or bidder in a failed merger is the original source. In 27 cases, a regulatory agency employee receives inside information, typically about drug approvals or clinical trial updates.

B. Tip Chains

Insider information follows a path from the original source to other tippees. I denote a "tip chain" as the ordering of people who are tipped as the information flows from the original source to others. The order in the tip chain is the number of links from the original source. If a tipper tips multiple tippees, than each of these tippees is in the same order in the tip chain. The first order in the tip chain is the connection between the original source and his tippees. By identifying these orderings, we can better understand the path through which information flows between insiders.

Table XI presents the occupations of tippers and tippees by their order in the tip chain. Tippers in the first link are the original sources. Top executives (34%) make up the largest component of original sources followed by sell side employees (26.3%) and then lower-level employees (16.1%) and corporate managers (13.7%). The tippees in the first link from the original source are most commonly buy side analysts (19.3%), followed by buy side managers (12.5%). The rest of the tippees in the first link are evenly spread across all occupations. These results are consistent with the heat maps of the inter-occupation information flows shown previously.

As information progresses across the tip chain, a number of patterns emerge. Officers become less common tippees, going from 9.2% of all tippees in the first link to 0% of tippees in fourth and subsequent links. Corporate managers, lower-level employees, and people in specialized occupations follow the same decreasing pattern. Insiders in these roles are much more likely to receive the information from an original source than an intermediary. In contrast, buy side managers and analysts are increasingly the tippees as the information travels further from the source, accounting for 34.5% and 25.5% of all tippees in the fourth and subsequent links. On the other side of the tipping relationship, Table XI shows that as information flows further from the original source, buy side managers and traders are tipping other buy side traders. Executives, corporate managers,

lower-level employees, small business owners, and specialized workers are all becoming less likely to be tippers as the tip moves further from the original source.

Internet Appendix Table VII presents a similar analysis of occupations across the tip chain, for long versus short tip chains. The distribution of original sources is highly similar in long and short tip chains. In longer tip chains, the first tippees are more likely to be buy side traders than corporate executives and managers. In shorter tip chains, corporate executives and managers are more likely to be the first tippee compared to buy side traders. Compared to the same link in the tip chain, the later tips in short chains tend to go to executives and sell side employees, whereas they go to buy side traders in longer chains. These results suggest that there is path-dependence based on the earliest tippees. When early tippees are buy side traders, the tip chain is longer and few corporate employees and sell side workers receive the information. When the early tippees are corporate and sell side workers, the tip chain is shorter and buy side traders are less likely to receive the information.

These patterns show that information travels from corporate insiders through various intermediaries and eventually reaches professional traders who then share the information among themselves. The intermediary links include a wide variety of occupations, such as managers and officers, lawyers, specialized occupations, and investment bankers and analysts. It takes at least three links before professional buy side traders dominate the information exchange.

Table XII presents additional information about the characteristics of tippers and tippees along the tip chain. Tippees are women in 8.4% of first tips, declining to 2.3% in fourth and later links in the tip chain.[5] This probably reflects that compared to other professions, women are underrepresented among buy side managers and analysts. Tippers and tippees become younger as the information flows further from the original source. In the first link, tippees are 42.4 years old on average, declining to 38.7 by the fourth and subsequent links. The average tipper's age declines from 42.8 years to 34.4 years old. Tippees and tippers further from the original source are wealthier than those earlier in the tip chain.

There are also clear patterns of social connections over the tip chain. Tippers and tippees are primarily friends (42.4%) and family members (24.6%) in the first link of the tip chain and then

[5]The jump to 16.8% for women in the third link in the tip chain is driven by one set of serial inside traders.

steadily decline as the tip moves further from the source to 18.6% for friendship connections and 11.9% for family connections by the fourth and later links. In contrast, business connections grow in prevalence from 28.4% in the first link to 66.1% by the fourth and later links. Geographic distance follows a hump-shaped pattern over the tip chain. Tippers and tippees live further away from each other in the second link than they do in the first, but by the fourth and later links, distance has declined to zero at the median, which indicates that the median tippers and tippees live in the same city. Greater similarity in tipper and tippees' house values and surname ancestry also decline over the tip chain, though common ancestry remains high at 42.5% even in fourth and later tips.

C. Trading Characteristics over the Tip Chain

Table XII next documents trading behavior of the tippees by their position in the tip chain. The median amount invested rises monotonically from $200,400 for the first tippee to $492,700 for the fourth and subsequent tippees. Median profits rise as well from $17,600 to $86,000 per tip. However, trading returns decline over the tip chain, indicating that the information is moving stock prices. The initial tippee earns returns of 46.0% on average and 25.2% at the median. By the fourth link, the returns have dropped to 23.0% for the average and 18.8% for the median. The drop in returns is partially caused by an increase in the use of shares rather than options over the tip chain. As the trade sizes get larger, the fraction of total trading volume accounted for by insiders increases. These patterns are consistent with information flowing to professional traders over the tip chain. These traders invest large amounts of money using shares, rather than options. In return, they earn lower percentage returns, but greater dollar returns.

Next, Table XII presents evidence on the speed of the information flow over the tip chain. The average time between receiving information and sharing it with others decreases over the tip chain. The original source waits 12.1 days, on average, before tipping the information. At the second link in the chain, the delay is 9.2 days, followed by 5.0 days at the third link, and then 0.4 days for the fourth and higher links. The fraction of tippers who tip the same day that they receive information is 46.5% for the original source, increasing to 92.1% in the fourth and higher links. This delay means that the holding period between when the tippee receives the information and the event

date declines over time, from an average of 13.9 trading days in the first link to 9.1 days in the later links.

Finally, Table XII shows the network centrality of the tipper and tippee over the tip chain. Centrality is measured as the number of tipping links a person has to all other insiders. More central people have more information connections with others. The table shows a quick increase in the centrality of the tipper over the tip chain. In the first link, the average tipper has 2.9 information connections. In the second link, the tippers' average degree centrality increases to 4.3. By the fourth link, the tipper has a degree centrality of 4.6. In contrast, the tippee's centrality decreases from 2.9 in the first link to 1.8 by the fourth and subsequent links. This pattern implies that tippers in later links are more central figures that are spreading the information to peripheral people in the network. I explore these relations in more detail in the next section of the paper.

Because many of the characteristics of inside traders are correlated, Table XIII presents regression results where the dependent variable is the position in a tip chain and the independent variables are tipper and tippee characteristics. Columns 1 and 2 present OLS regression coefficients. Columns 3 and 4 present ordered logit regression coefficients. The results indicate that age is negatively related to a tipper's position in the tip chain, even after controlling for occupation, gender, and distance. Family and friends relationships decline with order in the tip chain, even after controlling for occupation. Likewise, the presence of buy side tippers are positively related with distance from the original source in a tip chain, compared to top executives, after controlling for age, gender, and the type of relationship.

D. Summary of the Diffusion of Information

The results in this section of the paper reveal distinct patterns in the diffusion of inside information. Original sources of M&A information are most commonly top executives of firms, and the original source is equally likely to be an employee of the acquirer, the target, or an outside firm. In contrast, lower level employees are more likely to leak earnings news than are top executives. As information diffuses from the original source, top executives and mid-level corporate managers are less likely to share information. Instead, buy side investors become dominant after three degrees of separation from the original source, with family and friend relationships declining over the tip

chain and professional relations increasing. Investors invest larger sums, earn lower returns, share information faster, and have shorter holding periods as information moves away from the original source.

VII. The Networks of Inside Traders

Within the entire sample there are 183 insider networks. There are 59 networks that contain only one person. These are people who obtained inside information and did not tip it to anyone else. The remainder of the size distribution of the networks is as follows: 60 networks with 2 members, 18 networks with 3 members, 18 networks with 4 members, 11 networks with 5 or 6 members, 11 networks with 7 to 10 members, and 6 networks with more than 10 members.

The largest network, shown in Figure 7, centers on traders affiliated with the expert networking firm, Primary Global Research (PGR), and the hedge funds owned by SAC Capital. The SAC-PGR network has 63 members with a few key players. In the bottom left of the figure are Danny Kuo and Jesse Tortora, a buy side manager and analyst. In the top right is James Fleishman, a PGR employee and Mark Longoria, a mid-level manager in Advanced Micro Devices. Close to the middle of the network are Ronald Dennis, a buy side technology analyst at CR Intrinsic, and Noah Freeman, a portfolio manager at SAC Capital.

Following the direction of information flow in the figure reveals that the center of the network is Steven Cohen, who oversees a number of large hedge funds owned by SAC Capital. Cohen receives information from five different sources, three of which have multiple sources of information themselves. While Cohen receives information from many sources, he does not tip anyone else. Other large networks are presented as examples in Internet Appendix Figures 1, 2, and 3.

Table XIV presents summary statistics of personal characteristics, social connections, trading activity, timing of information flows, and trading behavior by the size of insider networks. The density of the network is the proportion of all possible connections that actually exist. As insider networks grow in size, they become less connected. For networks with six or more members, only 20% of possible information connections actually exist. The diameter of the network is defined as the longest of all shortest paths between any two members. In the sample of networks, the diameter of a network increases as more members are added. This provides additional evidence that networks

become more dispersed and sprawling, instead of compact and closely tied to a central hub. The average clustering coefficient is a measure of how closely connected is the average node in a network. A node's clustering coefficient is the fraction of a node's links that are also linked to each other. In the sample of insider networks, the average clustering coefficient is also decreasing. These network statistics reveal that as a network gets larger it is spreading out from its outer members. The Rajaratnam and SAC-PGR networks illustrate this phenomenon.

Next, as networks increase in size, they have a smaller fraction of female tippers, which means that women tend to be tippers in small networks. The age of tippers and tippees also decrease with network size. Younger insiders tend to be included in large networks. Also, as networks grow, there exist fewer family connections and more business connections. Friendship connections remains roughly the same across network size. Given that family sizes are limited, this is not surprising. The median geographic distance increases as networks get larger.

As networks increase in size, the average and median amount invested per tippee increase. The median profit increases as well, though percentage returns decrease. Finally, the time lapse between tips is lowest in the larger networks. These results suggest that large networks include professional traders who trade larger stakes and have lower percentage returns. Since the larger networks still experience time delays between tips, it is likely that the returns are lower because they are receiving the tip after insider trading has already begun to move stock prices.

VIII. Generalizability and Selection Bias

The results presented so far paint a detailed picture of illegal insider trading networks. However, the results are based only on traders who are identified in SEC and DOJ case documents. This means that the sample is not randomly selected and is unlikely to present an unbiased depiction of the average illegal inside trader. To understand how the average inside trader differs from the traders in the sample, I would need to know what drives the likelihood of getting caught. This requires observations of both insiders that are caught and insiders that are not caught, which, for obvious reasons, I cannot observe.

Though I cannot observe inside traders that are not caught, I can observe something that may help to address selection bias: the extent of prosecution for the traders that are caught. In my

sample, there are 199 traders that are identified in a case against someone else, but who are not formally charged by the SEC or DOJ themselves. In addition, there are 289 traders that are charged with a civil violation by the SEC, but not a criminal violation by the DOJ. Finally, there are 134 traders that are charged both by the SEC and the DOJ. Thus, there is variation in the extent of prosecution. As I discuss next, under two key assumptions, this variation will help to explain the likelihood of getting caught, and hence sample selection bias.

The first assumption is that the likelihood of getting caught is correlated with the extent of prosecution. I believe this is a reasonable assumption. First, there are two forces that likely influence whether an insider is caught: sophistication and the extent of illegal activity. It is reasonable to assume that sophisticated insiders are less likely to get caught than unsophisticated insiders. Some insiders are oblivious to the regular monitoring of financial markets by regulators. In one case, a trader's purchases accounted for 100% of the volume of out-of-the-money call options in the days before a merger announcement, sending up red flags to regulators. Sophisticated traders may also be more careful to hide incriminating evidence, such as emails and text messages. The second force that likely influences whether an insider is caught is the extent of an insider's activity. A serial insider who trades in seemingly unrelated companies that can be traced to a common source (e.g., an M&A lawyer) is easier to detect than an opportunistic insider who makes one illegal trade based on a once-in-a-lifetime tip (e.g., a brother's company is acquired). Plus, regulators have a greater incentive to identify insiders who are investing larger sums and making more trades.

Second, though the U.S. government does not have explicit guidelines for the extent of prosecution, practitioners believe the government decides whether to bring no case, a civil case, or a criminal case based primarily on two factors: the extent of the illegal activity and the credibility of the evidence (Pavlo, 2013). The extent of the activity refers to the size of investments, the number of different stocks that are traded, and the size of the insider network. The credibility of the evidence is important because civil cases only require guilt by a preponderance of the evidence, in contrast to criminal cases which require guilt beyond a reasonable doubt. Therefore, prosecutors often depend upon the testimony of cooperating witnesses in criminal cases (Newkirk and Robertson, 1998).

As this discussion shows, an insider's extent of illegal trading drives both the likelihood of getting caught and the extent of prosecution. Large trades in many companies provide more chance to be caught and also make a criminal case more likely. Insiders in large networks face the risk that they are easier to catch and that once they are caught, the DOJ can more easily find a cooperating witness in a criminal case. In contrast, sophisticated insiders are less likely to get caught, and if they are caught, they are less likely to face criminal charges because they are careful to avoid leaving incriminating evidence. Therefore, it is reasonable to believe that the likelihood of getting caught and the extent of prosecution are correlated.

The second assumption required for this approach is that selection bias does not affect the estimation of the extent of prosecution. This would be the case if the undetected insiders would face greater prosecution if they were caught than do the insiders who are actually caught. The greater is the variation across the traders that are caught, the more reasonable is this second assumption. Consider Figure 8, which presents a stylized view of the population of inside traders across two dimensions: sophistication and trading extent. Triangles represent insiders that are caught by regulators and circles represent insiders that are not caught. Insiders that make large trades and that are unsophisticated are more likely to get caught, and hence be in the sample. However, it is reasonable to assume that in the population of insiders, sophistication is positively correlated with the extent of trading, as shown in the figure. Because the likelihood of getting caught is orthogonal to the principal dimension of variation in the population, the sample includes a wide range of individuals, not just unsophisticated insiders making large trades. The empirical evidence supports these assumptions. The sample includes insiders who trade very small amounts and sophisticated traders who invest millions of dollars, such as successful hedge fund managers. These findings support the argument that the sample observations provide enough variation for the extent of prosecution to proxy for selection bias.

A. Tests of the Extent of Prosecution

Panel A of Table XV presents means of inside traders' characteristics by the extent of prosecution. Entries in the first column ("None") are means over insiders who were not charged by the SEC or DOJ. Entries in the second column include insiders charged by the SEC, but not the DOJ, and

the third column includes insiders charged by the DOJ. The numbers in parentheses in the second column are p-values from t-tests of the difference between the "None" column and the "SEC" column. Numbers in parentheses in the third column are p-values from t-tests of the difference between the "SEC" and "DOJ" columns.

Younger, unmarried traders are more likely to face more severe prosecutions. Women are much less likely to be charged by the SEC or DOJ, but there is no significant difference in gender between insiders charged by the SEC or DOJ. The average home value of insiders charged by the DOJ is significantly higher than insiders charged by the SEC. Insiders charged by the DOJ have significantly more connections and send and receive more tips than insiders charged by the SEC. Finally, insiders charged by the DOJ invest much larger sums and receive greater gains than insiders charged by the SEC, though their percentage returns are not significantly different.

Internet Appendix Table VIII shows that the occupations of insiders varies across the extent of prosecution. Buy-side analysts account for 24% of insiders charged by the DOJ, but only 9% of insiders charged by the SEC. In contrast, lower-level employees account for 5% of DOJ cases, but 15% of SEC case. Top executives and sell-side employees are the most likely to not be charged by the SEC or DOJ (25% and 20%). As a consequence, only 16% of DOJ cases list top executives as defendants, and 12% list sell-side employees. Thus, buy-side analysts are much more likely to face criminal charges, whereas lower-level employees, sell-side employees, and top executives are much more likely to avoid prosecution entirely.

Similar results are found in a multivariate setting. Internet Appendix Table IX reports the results from logit and ordered logit regressions on the extent of prosecution. The results show that young, male, unmarried insiders who give more tips are more likely to face greater prosecutions. Greater total investment and returns are related to a greater likelihood of prosecution, though they are less significant than the wealth of the insider. Relative to buy-side analysts, lower-level employees, corporate managers, and sell-side employees face a lower chance of criminal prosecution. Top executives, buy-side managers, specialized occupations, and small business owners all face a similar risk as buy-side analysts.

Panel B of Table XV presents the likelihood of civil and criminal charges by the type of relationship between pairs of tippers and tippees. The first two columns present averages for cases in

which at least the tipper or tippee was charged by the SEC (column 1) or the DOJ (column 2).[6] The difference between the averages and their statistical significance are presented in column 3. Columns 4 and 5 present coefficient estimates from logit tests where the dependent variable equals one if at least one of the tipper or tippee is charged by the DOJ and zero if neither is charged by the DOJ.

The multivariate results are consistent with the univariate results. Insiders who are family members are the least likely to be charged by the DOJ. Friends are also less likely to be charged by the DOJ, though business associates are equally likely to face criminal and civil charges. When insiders live closer to each other and share the same ancestry they are less likely to be charged criminally.

B. Implications for Generalizability

Under the assumptions discussed above, these results shed light on the generalizability of the main findings in this paper. At the individual level, the main results overweight young, wealthy male insiders, who have many connections, invest large sums, and typically work as buy-side analysts. This means that the general population of insiders includes older insiders who work in various occupations, share less inside information with fewer people, and make smaller trades. These people likely are opportunistic traders who receive information infrequently. In contrast, the traders in my sample are more likely to be serial inside traders.

The pair-level results also provide an interesting view of both the likelihood of detection and the generalizability of my results. The results show that insiders who do not get caught are those that share information with family and friends who live closer and are of the same ancestry. This could reflect that insiders in stronger relationships with greater trust are less likely to testify against their friends and family in a criminal case. These results also mean that in the general population of inside traders, information is more likely to spread through closer relationships than indicated in my sample. Given that close relationships already account for a large fraction of my sample, this implies that close relationships likely dominate the large majority of insider trading relations.

[6]There are very few observations where neither the tipper nor the tippee is charged by either the SEC or the DOJ, so I exclude them from the table.

C. Selection on Guilt

A final selection issue is whether the people accused of insider trading are found guilty. First, the track record of the DOJ is impressive: between 2009 and 2014, the DOJ has won 85 cases and lost just once. Therefore, the facts reported in the cases are likely to be true. Second, in SEC cases that are subsequently dropped, the facts presented in the case are typically accurate. They are usually dropped based on technical issues about what constitutes insider trading. For example, as discussed above, some defendants do not dispute the facts, but argue that they didn't violate insider trading rules because they were three to four links removed and didn't receive monetary compensation for the tip.

IX. Conclusion

This paper provides an in-depth description of the spread of private information within illegal insider trading networks. In particular, the paper provides new answers to basic questions such as who shares information with whom, what type of information is shared, who is the original source, how are these people related, and how fast does information travel?

The results show that original sources of illegal inside information are varied, but top executives are the most common. They tend to share information with other top executives first. The information then typically proceeds through a number of links and ends up with buy side analysts and managers. Information tends to flow from younger people to older people, from children to parents, and from subordinates to bosses. Tippers and tippees are more commonly friends and family in the early links of an information chain and more commonly business associates in later links. Tippers and tippees tend to live close to one another geographically, but there is wide variation, as insiders reside all over the globe. Networks of inside traders are sprawling, rather than centralized, where larger networks are more likely based on business relationships than family or friends.

The results of this paper validate existing research and suggests new directions for future investigation. First, the results provide validation for proxies of social interaction used in existing papers:

geographic proximity (e.g., Brown, Ivković, Smith, and Weisbenner, 2008) and common educational backgrounds (e.g., Cohen, Frazzini, and Malloy, 2010). Second, the paper shows that social connections are an important mechanism in the diffusion of important information across market participants, as suggested by Hong and Stein (1999). More broadly, this paper contributes to a burgeoning field of finance concerned with the role of social interactions in financial decision-making, which Hirshleifer (2014) calls "social finance."

REFERENCES

Augustin, Patrick, Menachem Brenner, and Marti G. Subrahmanyam, 2014, Informed options trading prior to M&A announcements: Insider trading? *McGill University and New York University Working Paper*.

Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Bhattacharya, Utpal, 2014, Insider trading controversies: A literature review, *Annual Review of Financial Economics* 6, 385–403.

Bhattacharya, Utpal, and Hazem Daouk, 2002, The world price of insider trading, *Journal of Finance* 57, 75–108.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott Weisbenner, 2008, Neighbors matter: Causal community effects and stock market participation, *Journal of Finance* 63, 1509–1531.

Bushman, Robert M., Joseph D. Piotroski, and Abbie J. Smith, 2005, Insider trading restrictions and analysts' incentives to follow firms, *Journal of Finance* 60, 35–66.

Chakravarty, Sugato, and John H. McConnell, 1999, Does insider trading really move stock prices? *Journal of Financial and Quantitative Analysis* 34, 191–209.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2010, Sell side school ties, *Journal of Finance* 65, 1409–1437.

Colla, Paolo, and Antonio Mele, 2010, Information linkages and correlated trading, *Review of Financial Studies* 23, 203–246.

Del Guercio, Diane, Elizabeth R. Odders-White, and Mark J. Ready, 2013, The deterrence effect of SEC enforcement intensity on illegal insider trading, Unpublished working paper, University of Oregon and University of Wisconsin–Madison.

Fishe, Raymond P.H., and Michael A. Robe, 2004, The impact of illegal insider trading in dealer and specialist markets: evidence from a natural experiment, *Journal of Financial Economics* 71, 461–488.

Frontline, 2014, Preet Bharara: Insider trading is "rampant" on Wall Street, Online transcript of interview.

Han, Bing, and David Hirshleifer, 2013, Self-enhancing transmission bias and active investing, Unpublished working paper, University of Texas-Austin and University of California-Irvine.

Han, Bing, and Liyan Yang, 2013, Social networks, information acquisition, and asset prices, *Management Science* 59, 1444–1457.

Hirshleifer, David A., 2014, Behavioral finance, Unpublished working paper, University of California – Irvine.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *Journal of Finance* 60, 2801–2824.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

King, Matthew R., Elizabeth Corrigan, and Craig Francis Dukin, 2009, Securities fraud, *American Criminal Law Review* 46, 1027–1098.

Meulbroek, Lisa K., 1992, An empirical analysis of illegal insider trading, *Journal of Finance* 47, 1661–1699.

Meulbroek, Lisa K., and Carolyn Hart, 1997, The effect of illegal insider trading on takeover premia, *European Finance Review* 1, 51–80.

Newkirk, Thomas C., and Melissa A. Robertson, 1998, Speech by SEC staff: Insider trading — a U.S. perspective, *16th International Symposium on Economic Crime, Jesus College, Cambridge, England*.

Ozsoylev, Han N., and Johan Walden, 2011, Asset pricing in large information networks, *Journal of Economic Theory* 146, 2252–2280.

Ozsoylev, Han N., Johan Walden, M. Deniz Yavuz, and Recep Bildak, 2014, Investor networks in the stock market, *Review of Financial Studies* 27, 1323–1366.

Pavlo, Walter, 2013, Insider trading: Civil or criminal crime? *Forbes* October 24.

Walden, Johan, 2013, Trading, profits, and volatility in a dynamic information network model, Unpublished working paper, University of California – Berkeley.

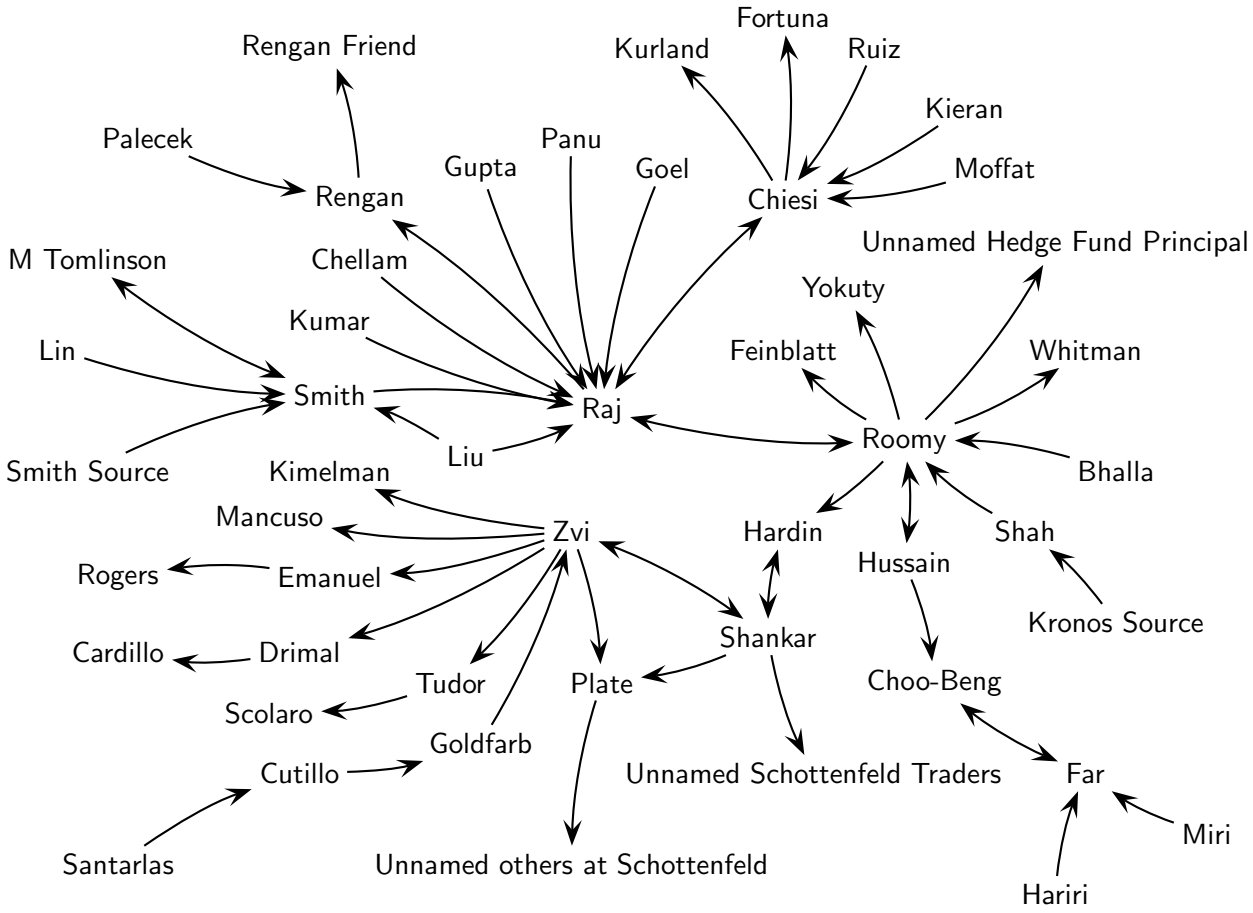


Figure 1
Raj Rajaratnam Network
This figure represents the illegal insider trading network centered on Raj Rajaratnam. Arrows represent the direction of information flow. Data are from SEC and DOJ case documents, plus additional source documents to identify individuals not named in the SEC and DOJ documents.

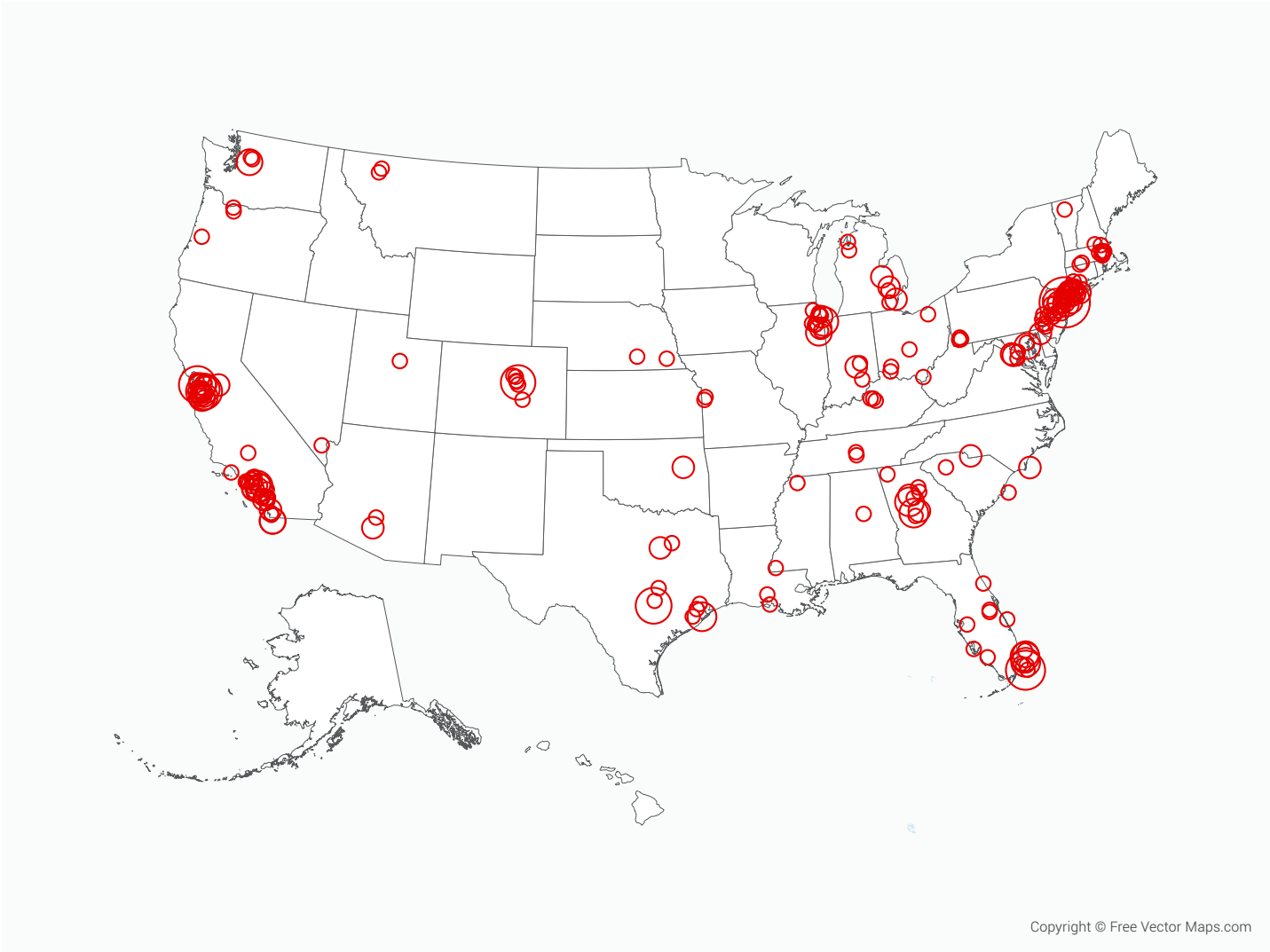


Figure 2
Location of Tippers and Tippees
Larger circles represent cities with more tippers and tippees.

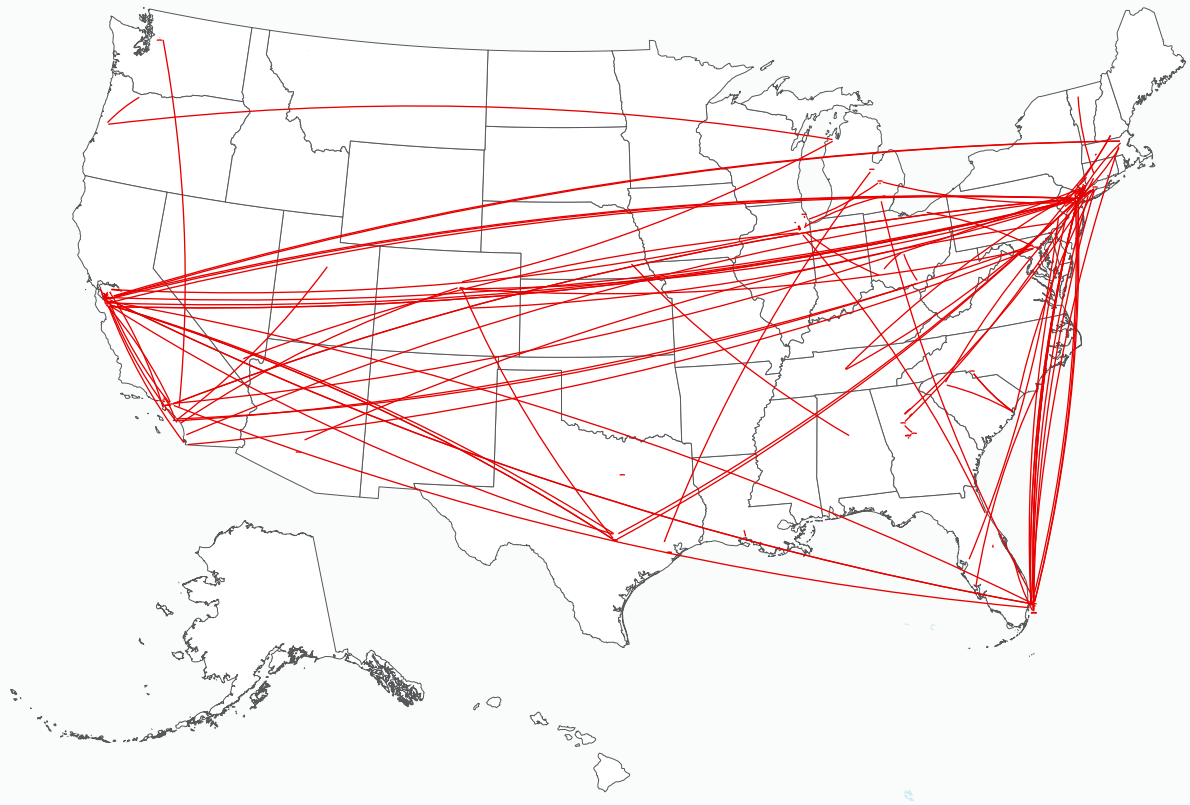


Figure 3
Tipper-Tippee Links

This figure shows the links between tipper and tippees by city. Intra-city links are not shown.

Tippers \ Tippees	Top executives	Corporate managers	Lower-level employees	Sell side	Buy side: managers	Buy side: analysts	Small business owners	Specialized occupations	Total
Top executives	15	6	7	5	11	10	9	13	100
Corporate managers	6	5	2	2	10	9	3	1	44
Lower-level employees	3	2	8	1	5	2	10	5	45
Sell side	1	1	4	13	3	14	8	1	51
Buy side: managers	1	2	3	0	24	8	0	0	44
Buy side: analysts	4	0	6	0	19	31	2	2	90
Small business owners	1	0	3	4	1	2	3	1	29
Specialized occupations	1	0	0	2	8	1	2	2	21
Total	35	17	38	39	85	86	41	29	473

Figure 4
Number of Connections between Tipper-Tippee by Occupation
This figure shows the number of tipper-tippee pairs sorted by the occupation of the tipper and the tippee. Darker colors represent more pairs. Totals include connections to people with unknown occupations, so row and column sums do not add up to the totals. Occupation definitions are detailed in Table III.

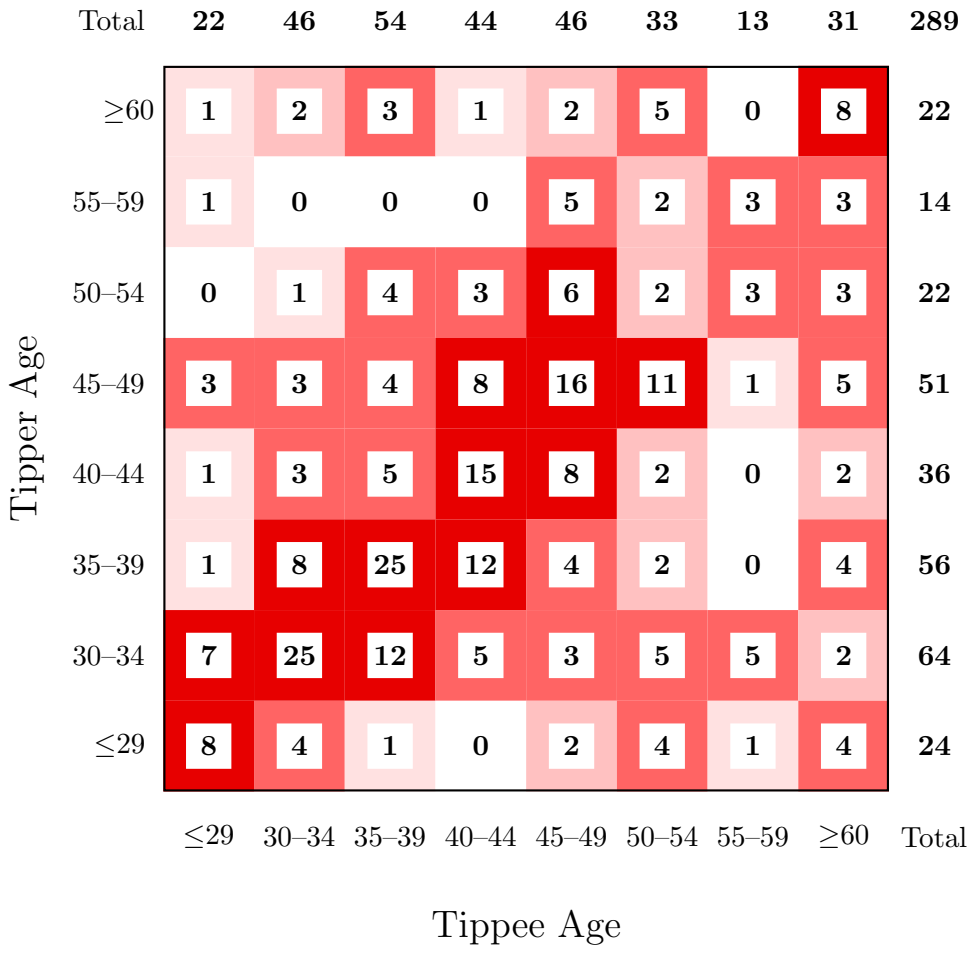


Figure 5
Number of Connections between Tipper-Tippee by Age
This figure shows the number of tipper-tippee pairs sorted by the ages of the tipper and the tippee. Darker colors represent more pairs.

Tippee Surname Ancestry

Tipper Surname Ancestry	Western European	Celtic	South Asian	Muslim	East Asian & Pacific	Eastern European	Hispanic	Jewish	Nordic	Total
Western European	77	10	9	1	5	10	2	7	1	128
Celtic	24	14	0	1	0	1	0	1	1	42
South Asian	8	0	10	8	2	0	0	2	0	31
Muslim	4	0	4	16	1	1	1	1	0	29
East Asian & Pacific	8	0	5	0	14	1	0	0	0	28
Eastern European	9	1	1	0	1	3	0	2	0	18
Hispanic	9	0	0	1	1	1	4	0	0	17
Jewish	9	1	1	0	0	0	0	5	0	16
Nordic	0	0	0	0	0	1	0	0	4	5
Total	156	29	30	28	24	20	8	18	6	329

Figure 6
Number of Connections between Tipper-Tippee by Ancestry of Surname
This figure shows the number of tipper-tippee pairs sorted by the ancestries of the tipper and the tippee. Darker colors represent more pairs. Totals include connections to people with unknown ancestries, so row and column sums do not add up to the totals. Ancestry of surnames is from the Onomap database.

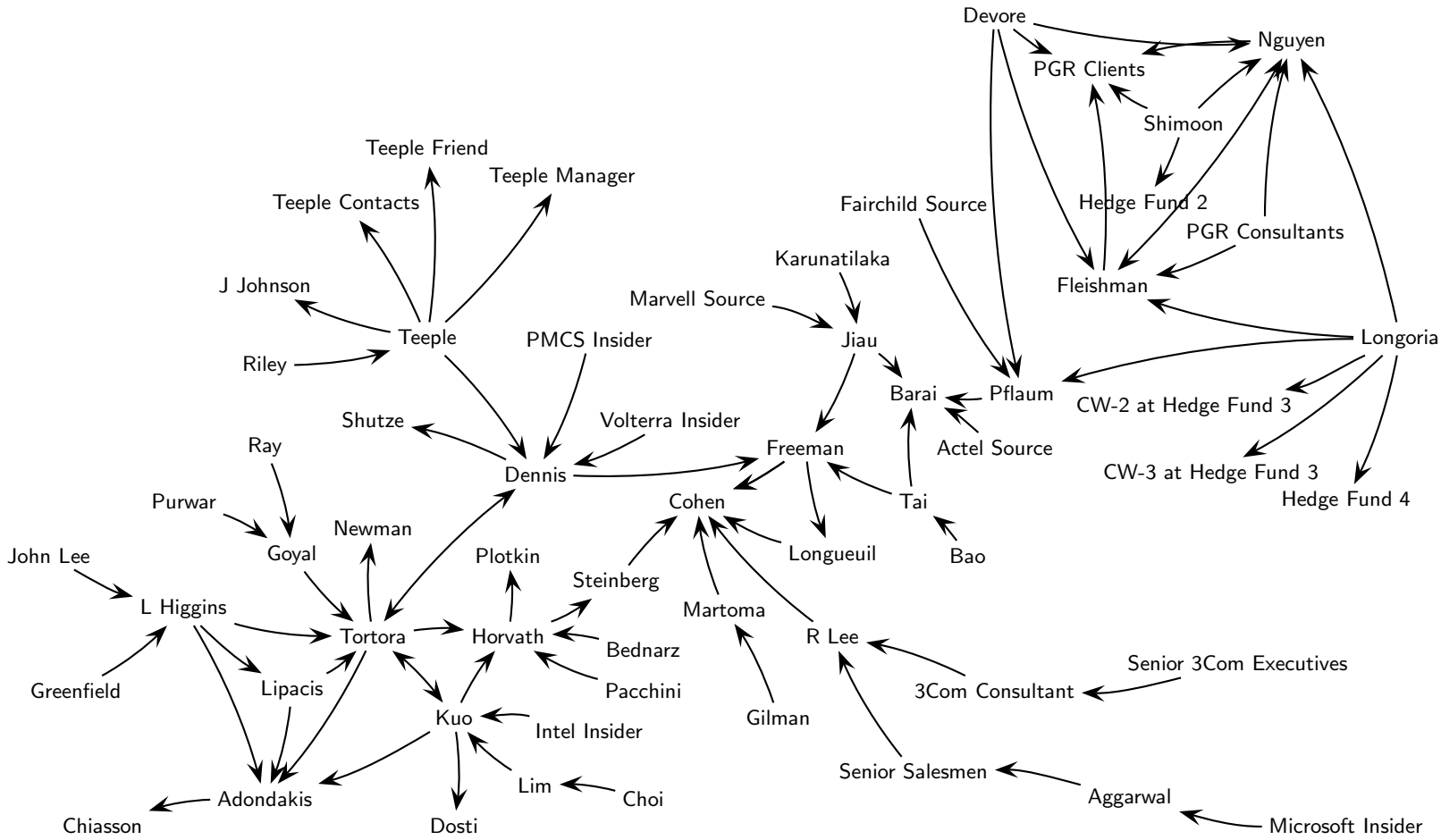


Figure 7
SAC-PGR Network
This figure represents the illegal insider trading network centered on the hedge funds controlled by SAC Capital and the expert networking firm, Primary Global Research. Arrows represent the direction of information flow. Data are from SEC and DOJ case documents, plus additional source documents to identify individuals not named in the SEC and DOJ documents.

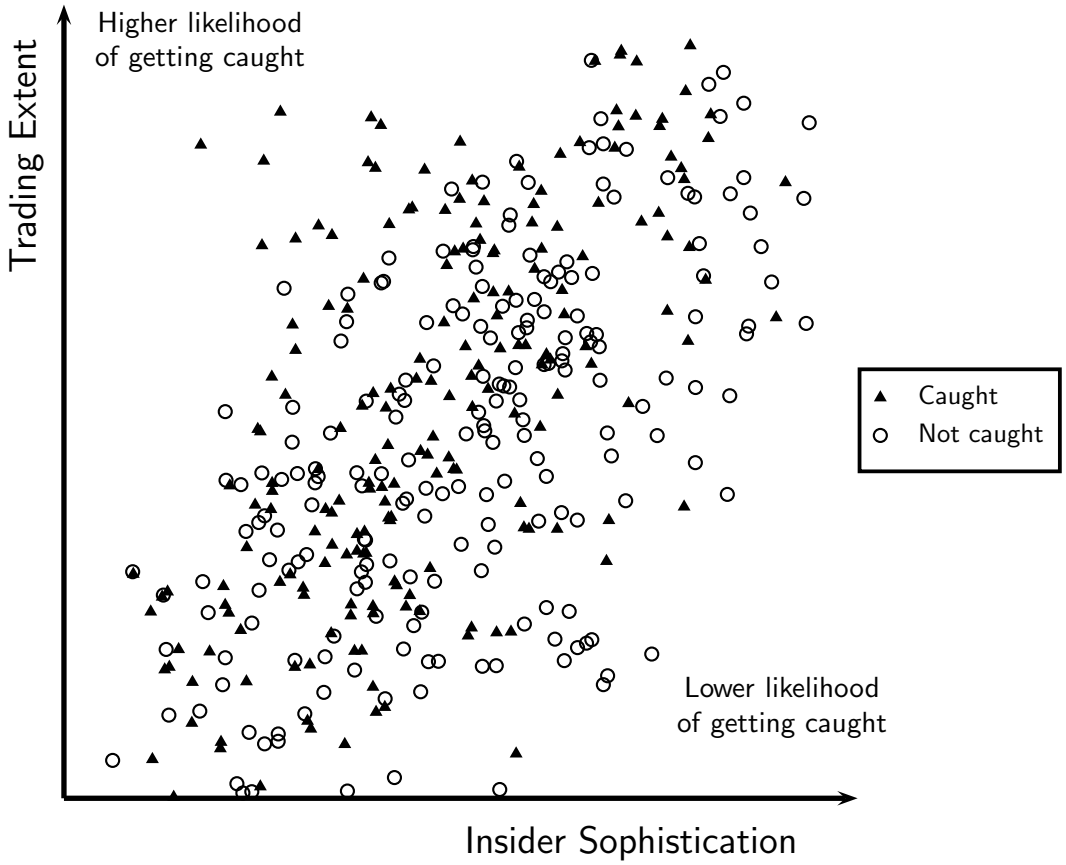


Figure 8
Stylized Illustration of Selection Bias
This figure represents a stylized illustration of potential selection bias in the sample of illegal insider trading cases. The circles and triangles represent the total population of inside traders. Triangles represent insiders that are caught by regulators and in the sample. Circles represent insiders that are not caught. The likelihood of getting caught increases as insiders trade more (y-axis) and are less sophisticated (x-axis). However, insiders are not randomly distributed across the two dimensions. Instead, sophistication and trading extent are positively correlated. The positive correlation between sophistication and trading extent and the orthogonal direction of the likelihood of getting caught implies that the sample will include a wide range of insiders, not just unsophisticated insiders that trade extensively.

Table I
Events by Year
This table provides a breakdown of the years of the corporate events in the sample. The data include 465 corporate events over the years 1996 to 2013 from SEC and DOJ case documents.

Year	Number of Events	Fraction
1996	1	0.2
1997	6	1.3
1998	8	1.7
1999	0	0.0
2000	0	0.0
2001	1	0.2
2002	1	0.2
2003	0	0.0
2004	4	0.9
2005	18	3.9
2006	47	10.1
2007	69	14.8
2008	81	17.4
2009	76	16.3
2010	66	14.2
2011	40	8.6
2012	17	3.7
2013	5	1.1
Unknown	25	5.4
Total	465	

Table II
Corporate Events and Returns
This table presents the frequency (panel A), stock returns (panel B), and timing (panel C) of illegal insider trading events. The data include 465 corporate events over the years 1996 to 2013 from SEC and DOJ case documents. Frequency of events in panel A are the number of occurrences of an event type. Stock returns in panel B are the average raw stock returns from the date that the original tipper receives the information through the date of the corporate event. Panel C presents the average number of trading days between the date that the original tipper receives the information through the date of the corporate event. "Negative", "Positive", and "Unknown" columns indicate the expected effect on stock prices, based on the trading behavior of tippees. Stock returns in the "All" column are calculated as long positive events and short negative events. Detailed descriptions of each event type are presented in the appendix.

	Outcome			
Event Type	Negative	Positive	Unknown	All
Panel A: Frequency of events				
Clinical Trial/Drug Regulation	13	24	0	37
Earnings	54	66	3	123
M&A	5	234	0	239
Operations	3	8	2	13
Sale of Securities	34	1	0	35
Other	3	2	13	18
Total	112	335	18	465
Panel B: Information to event returns (%)				
Clinical Trial/Drug Regulation	−38.6	101.2		79.7
Earnings	−12.2	14.7		13.5
M&A	−20.3	43.7		43.1
Operations	−29.7	22.9		24.9
Sale of Securities	−12.8	0.0		12.8
Other	−28.9	−4.8		8.9
Total	−16.8	42.4		34.9
Panel C: Information to event time (trading days)				
Clinical Trial/Drug Regulation	5.3	11.2		9.2
Earnings	10.2	12.3		11.3
M&A	7.8	31.1		30.5
Operations	2.0	7.9		6.1
Sale of Securities	11.1	5.0		10.9
Other	101.1	3.0	1.3	44.3
Total	12.2	25.0	1.3	21.3

Table III
Event Firms Summary Statistics
This table presents summary statistics of the firms whose stocks are illegally traded based on inside information, using data from 465 corporate events over the years 1996 to 2013. "Total dollars invested by tippees" is the total dollar amount of a company's stock purchased or sold across all trades by all inside traders in the data. "Invested/Daily dollar volume" is total dollars invested by tippees divided by daily dollar volume. This means the amount invested is the aggregate trading by insiders over many trading days relative to an average day of normal trading activity.

				Percentiles				
	Mean	S.D.	Min	25th	50th	75th	Max	Observations
Market equity (billions)	10.09	37.39	0.01	0.30	1.01	3.56	422.64	391
Employees (thousands)	13.71	41.63	0.00	0.37	1.70	6.55	398.46	387
Tobin's Q	2.54	2.41	0.35	1.31	1.87	3.06	36.24	391
Daily trading volume (millions)	3.18	8.85	0.00	0.21	0.68	1.85	71.98	393
Daily dollar trading volume (millions)	114.32	372.91	0.02	2.64	13.08	47.19	3456.96	393
Total dollars invested by tippees (millions)	4.06	12.99	0.01	0.08	0.37	1.43	132.64	269
Invested/Daily dollar volume (%)	45.96	153.11	0.02	1.31	6.57	34.56	2011.43	259

Table IV
Summary Statistics of People Involved in Insider Trading

This table presents summary statistics for the 622 people in the sample of insider trading from 1996 to 2013. Data are from SEC and DOJ case documents, plus additional sources detailed in the paper. "Median house value" ("Median house size") is the median estimated value (square footage) over all properties owned by a person that are also listed as a person's residence in the LexisNexis data. Home values and square footage are as of September 2014 from Zillow. "Tips given" is the number of insider trading tips given to others. "Tips received" is the number of insider tips received. "Total invested" is the aggregated dollar amount of all trading positions in absolute value. Thus, this includes the size of short positions. "Total gains" is the aggregated dollar amount received by a person across all events and trades. "Average return" is the average of a person's trades, based on actual buy and sell dates from the SEC and DOJ documents.

				Percentiles				
	Mean	S.D.	Min	25th	50th	75th	Max	Observations
Age	44.1	11.5	19.0	35.8	42.7	51.5	80.0	454
Female (%)	9.8	29.8	0.0	0.0	0.0	0.0	100.0	498
Married (%)	92.0	68.2	0.0	100.0	100.0	100.0	100.0	425
Median house value ($1,000s)	1,114.3	1,846.5	49.6	390.0	656.3	1,170.5	25,600.0	365
Median house size (100 ft.^{2})	30.0	21.3	7.5	18.6	26.5	35.4	316.4	351
Tips given	1.5	3.2	0.0	0.0	1.0	1.0	29.0	622
Tips received	1.5	2.5	0.0	0.0	1.0	1.0	24.0	622
Total invested ($1,000s)	4,288.0	25,334.9	4.4	74.5	226.0	1,116.2	375,317.3	255
Average invested per tip ($1,000s)	1,690.3	6,088.6	4.4	65.0	200.0	701.4	72,427.5	255
Total invested/median house value (%)	581.6	3,609.5	0.6	13.2	38.9	140.3	44,183.1	159
Total gains ($1,000s)	2,331.6	13,207.1	0.9	34.2	136.0	606.0	139,500.0	399
Average gains per tip ($1,000s)	1,289.9	10,036.0	0.1	20.6	72.4	285.0	139,000.0	399
Average return (%)	63.4	231.8	0.0	14.0	26.4	46.5	3347.3	255

Table V
Summary Statistics by Occupation
This table presents averages of age, gender, tipping activity, investment, and returns by occupation for the 622 people in the sample of insider trading from 1996 to 2013. Data are from SEC and DOJ case documents, plus additional sources detailed in the paper. Table IV defines the variables. Table III presents a breakdown of occupations into narrower categories.

Occupation	Count	Percent	Age	Female	House value ($1,000s)	Tips Given	Tips Received	Median invested per tip ($1,000s)	Median gains per tip ($1,000s)	Median Return
Top executive	107	17.3	50.9	6.1	1,152	1.5	0.5	377	2,903	33.7
Corporate manager	55	8.9	41.3	5.6	624	1.6	0.6	1,973	286	71.8
Lower-level employee	59	9.5	41.5	25.5	621	1.7	1.5	558	149	47.5
Sell side/lawyer/accountants	61	9.9	41.7	11.1	1,164	3.0	1.9	3,756	255	58.8
Buy side: manager	60	9.7	42.6	6.4	3,006	1.4	2.6	5,972	5,768	37.0
Buy side: analyst/trader	65	10.5	35.5	5.2	1,061	2.7	3.1	2,051	442	117.7
Small business owner	39	6.3	47.4	5.3	678	0.9	2.3	204	196	62.0
Specialized occupation	38	6.1	52.0	2.8	729	1.2	1.6	599	234	32.8
Unknown	135	21.8	41.5	20.4	613	0.5	1.1	584	355	77.0

Table VI
Inside Traders Compared to their Neighbors

This table presents averages and t-tests of characteristics of inside traders compared to their neighbors (columns 1–3) and logit regression coefficients where the dependent variable equals one if a person is in the insider trading data and zero if the person is a neighbor (columns 4 and 5). Column 5 includes insider-neighbor pair fixed effects, which accounts for all time invariant characteristics of the neighborhood. Neighbors are identified as people of the same gender who have a residence on the same street as the inside trader. "Age is calculated for insiders and neighbors based on the insider trading dates. "Owns property" is a dummy variable equal to one if a person has any real assets, as recorded in the Lexis Nexis database. "Bankruptcy" equals one if a person filed for backruptcy, as recorded in the Lexis Nexis database. "Judgments and liens" is a dummy variable. For insiders, bankruptcies, judgments, and liens are recorded as one only if the event occurred before the end of the insider trading dates in the sample. This means that these events are not the direct outcome of any prosecution for insider trading. For neighbors, these variables are recorded as a one if the event happens at any time. 'UCC liens," "Licenses," "Criminal record," "Democrat," and "Republican" are dummy variables if a person has any UCC liens filed against him, has a license, any criminal record (non-Federal crimes, mostly traffic violations), registered as a Democrat or a Republican. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	Averages			Logit Regressions	
	Insiders	Neighbors	Difference	Likelihood of insider	
	(1)	(2)	(3)	(4)	(5)
Age	43.1	41.7	1.5	−0.015**	−0.039***
			(0.814)	(0.017)	(< 0.001)
Owns property (%)	85.3	77.7	7.6***	−0.010***	−0.170
			(0.001)	(0.000)	(0.988)
Bankruptcy (%)	2.2	4.3	−2.0*	−0.005	−0.012
			(0.072)	(0.253)	(0.146)
Judgments and liens (%)	23.0	14.3	8.7***	−0.001	−0.002
			(0.001)	(0.729)	(0.538)
UCC liens (%)	21.0	11.0	10.1***	0.008***	0.010***
			(< 0.001)	(0.002)	(0.009)
Family members	7.3	7.5	−0.1	−0.017	−0.112**
			(0.386)	(0.604)	(0.048)
Person associates	6.5	5.1	1.4***	0.047***	0.043**
			(0.002)	(0.001)	(0.044)
Licenses (%)					
Insurance	0.9	0.7	0.2	0.001	0.002
			(0.655)	(0.950)	(0.844)
Accounting	3.3	1.1	2.2**	0.012*	0.012
			(0.018)	(0.091)	(0.156)
Engineering	0.7	1.8	−1.1	−0.007	−0.016
			(0.132)	(0.326)	(0.188)

	Averages			Logit Regressions	
	Insiders	Neighbors	Difference	Likelihood of insider	
	(1)	(2)	(3)	(4)	(5)
Attorney	2.7	0.9	1.8**	0.010	0.010
			(0.045)	(0.237)	(0.272)
Real estate	1.8	2.9	−1.1	−0.010**	−0.013
			(0.276)	(0.018)	(0.125)
Medicine	3.1	2.9	0.2	−0.003	0.002
			(0.842)	(0.451)	(0.713)
Other	3.6	3.8	−0.2	0.000	0.003
			(0.862)	(0.983)	(0.586)
Criminal Record (%)	53.7	12.8	40.9***	0.005***	0.010***
			(< 0.001)	(< 0.001)	(< 0.001)
Republican (%)	46.3	38.8	7.5		
			(0.173)		
Democrat (%)	21.8	40.8	−19.0***		
			(0.000)		
Insider-neighbor pair fixed effects				No	Yes
Pseudo R^2				0.096	0.329
Observations				674	498

Table VII
Tippers and Tippees Social Relationships
This table reports the frequency of different types of social relationships among the 461 pairs of people in the insider trading data. When one person tips insider information to another person, they are are considered a pair. Pairs can have more than one type of social relationship, so the sum of relationship types is greater than 445. The type of relationship is defined based on the text in SEC and DOJ case documents. "Business Associates" are people who work together or know each other through business relationships where neither person has a supervisory role over the other. "Boss" refers to business relationships where one person is subordinate to another. "Client" is a relationship where one person is business client of the other.

Type of relationship	Count	Fraction of All Pairs	Fraction of Relationship Type
Family			
Dating/Engaged	7	1.5	6.7
Married	15	3.3	14.4
Parent-child	20	4.3	19.2
Siblings	25	5.4	24.0
In-laws	12	2.6	11.5
Other	9	2.0	8.7
Unspecified	16	3.5	15.4
All Family Relationships	104	22.6	
Business			
Business associates	87	18.9	54.4
Boss	41	8.9	25.6
Client	32	6.9	20.0
All Business Relationships	160	34.7	
Friends			
Acquaintances	3	0.7	1.9
Friends	115	24.9	71.0
Close Friends	44	9.5	27.2
All Friendship Relationships	162	35.1	
No social relation listed	98	21.3	

Table VIII
Geographic Distance by Relationship
Geographic distance is measured in miles using the great circle distance between the cities of residence of the people in a pair. Residences are from the SEC and DOJ case documents. Acquaintances and unknown family relations are omitted because they have very few observations.

				Percentiles				
	Mean	S.D.	Min	25th	50th	75th	Max	Observations
All	581.1	1,190.8	0	0.0	26.2	739.3	8,065.9	229
Family	466.3	1,080.1	0	0.0	14.3	322.5	5,350.4	59
Married/Dating	47.3	183.4	0	0.0	0.0	0.0	710.2	15
Parent-child	326.8	603.6	0	14.3	26.2	160.1	1,986.5	13
Siblings	783.3	1,490.8	0	6.2	28.0	1,063.6	5,350.4	14
In-laws	1,110.5	1,751.5	0	9.6	216.5	1,626.3	5,350.4	9
Other	88.0	128.3	0	8.1	11.5	237.3	307.6	7
Business ties	327.5	709.5	0	3.5	18.9	219.9	4,452.0	98
Associates	334.5	751.6	0	0.0	16.8	236.8	4,452.0	57
Boss	198.4	395.1	0	3.9	24.8	53.4	1,235.4	22
Client	455.9	857.2	0	10.7	19.0	625.6	2,666.9	19
Friendship	715.1	1,352.4	0	4.0	28.4	1,045.8	8,065.9	106
Friends	708.2	1,247.7	0	4.5	32.1	1,077.6	6,622.1	72
Close friends	694.6	1,581.8	0	0.0	26.2	835.3	8,065.9	32
No social relation listed	962.2	1,680.0	0	15.0	80.9	1,117.4	5,358.7	14

Table IX
The Likelihood of Receiving a Tip Using Potential Tippees
This table presents logit regression coefficients where the dependent variable equals one if a person receives an insider trading tip and zero otherwise. Observations include a tipper's family members and person associates. Family members and person associates are identified using the Lexis Nexis database. Person associates are non-family members who have a connection to the tipper. Columns 3 and 4 include tipper fixed effects. In columns 2 and 4, the omitted benchmark category is "Person associates," which means the coefficients on the family members are relative to the likelihood that a person associate receives a tip. "Other" are family members other than immediate family and in laws, such as uncles and grandparents. Standard errors are clustered at the tipper level. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	Likelihood of Receiving a Tip			
	(1)	(2)	(3)	(4)
Same gender	1.447***	1.351***	1.223***	1.155***
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Age difference	−0.049***	−0.063***	−0.049***	−0.062***
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Family member?	−1.097***		−0.618***	
	(< 0.001)		(< 0.001)	
Omitted baseline category: Person associates				
Husband		1.651***		1.862***
		(0.007)		(< 0.001)
Wife		−1.862***		−1.513**
		(0.003)		(0.017)
Brother		−0.435		−0.111
		(0.128)		(0.704)
Sister		−2.820***		−2.407**
		(0.007)		(0.018)
Father		−0.024		0.652
		(0.957)		(0.157)
Mother		−1.057		−0.387
		(0.171)		(0.615)
Son		0.468		0.698
		(0.433)		(0.183)
Daughter		−0.300		−0.223
		(0.781)		(0.831)
Other		0.644		0.987**
		(0.205)		(0.042)
In law		−1.588***		−1.070***
		(< 0.001)		(0.003)

	Likelihood of Receiving a Tip			
	(1)	(2)	(3)	(4)
Unknown		-3.907^{***}		-3.569^{***}
		(< 0.001)		(< 0.001)
Tipper fixed effects	No	No	Yes	Yes
Pseudo R^2	0.187	0.252	0.099	0.150
Observations	2,604	2,604	2,604	2,604

Table X
Original Source in M&A and Earnings Leaks
This table presents the frequency of different original sources of inside information in M&As and earnings announcements. For information leaks related to M&As, internal sources are those that are employed by the acquirer, target, or other firms, or are unknown. Other firms include potential and failed bidders. External sources include people that are employed by firms other than the merging firms. Consulting firms include human relations firms, investor relations firms, medical consultants, credit rating agencies, etc. For information leaks related to earnings, internal sources are people that are employed by the firm releasing the earnings statements.

	M&A				Earnings
	Acquirer	Target	Other/ Unknown	Total	
Internal sources					
Director	0	19	1	20	4
Officer	22	31	7	60	36
Employee	21	15	1	37	44
Total internal	43	65	9	117	84
External sources					
Accounting firm	15	6	0	21	20
Consulting firm (HR, IR, etc.)	4	3	0	7	13
Investment bank	13	10	14	37	0
Law firm	11	20	35	66	0
Stock market employee	0	0	0	0	5
Stolen	11	14	5	30	0
Other	2	1	0	3	1
Total external sources	56	54	54	164	39
Unknown sources	0	0	4	4	3
Total	99	119	67	285	126

Table XI
Tipper and Tippee Occupations by Order in Tip Chain
This table presents the distribution of occupations for each link in a tip chain. The position in a tip chain is the distance from the original source of the inside information, where distance is measured as information links between people. "1" indicates the first link in the chain from an original source to all the people to whom the source shared the information. "2" indicates the links from the first people to receive the tip from the original source to all of the people with whom they share the information. "3" and "≥ 4" are defined analagously. Entries in the table present the fraction of all people at a given position in the tip chain that are employed in each of the listed occupations. Thus, columns sum to 100%. "Number" indicates the number of relationships for each position in the tip chain.

	Position in Tip Chain			
	1	2	3	≥ 4
Tippee Occupation				
Top executive	9.2	2.4	2.3	0.0
Corporate manager	5.5	2.1	0.8	0.0
Lower-level employee	9.6	12.8	9.9	0.0
Sell side/lawyer/accountants	10.8	14.1	13.7	7.3
Buy side: manager	12.5	22.1	16.0	25.5
Buy side: analyst/trader	19.3	14.5	28.2	34.5
Small business owner	10.4	10.7	10.7	3.6
Specialized occupation	10.4	4.8	3.1	0.0
Unknown	12.3	16.6	15.3	29.1
Number	415	290	131	55
Tipper Occupation				
Top executive	34.0	6.6	0.8	0.0
Corporate manager	13.7	8.6	3.1	0.0
Lower-level employee	16.1	5.2	11.5	3.6
Sell side/lawyer/accountants	26.3	6.9	29.8	7.3
Buy side: manager	1.0	15.2	17.6	12.7
Buy side: analyst/trader	1.9	27.9	27.5	70.9
Small business owner	0.7	6.9	7.6	1.8
Specialized occupation	1.4	13.4	1.5	0.0
Unknown	4.8	9.3	0.8	3.6
Number	415	290	131	55

Table XII
Tipper and Tippee Characteristics by Order in Tip Chain
This table presents the characteristics of tippers and tippes for each link in a tip chain. The position in a tip chain is the distance from the original source of the inside information, where distance is measured as information links between people. "1" indicates the first link in the chain from an original source to all the people to whom the source shared the information. "2" indicates the links from the first people to receive the tip from the original source to all of the people with whom they share the information. "3" and "≥ 4" are defined analagously. "Degree centrality" is the number of connections of the person in a connected component.

	Order in Tip Chain			
	1	2	3	≥ 4
Characteristics				
Tippee female	8.4	4.8	16.8	2.3
Tipper female	11.5	14.8	4.5	10.5
Tippee age	42.4	41.6	40.7	38.7
Tipper age	42.8	41.4	35.7	34.4
Tippee house value - median ($1,000s)	668.7	724.5	833.7	1,072.0
Tipper house value - median ($1,000s)	811.7	840.1	758.5	1,072.0
Connections				
Family connection	24.6	15.5	28.4	11.9
Friendship connection	42.4	35.2	36.6	18.6
Business connection	28.4	47.0	34.3	66.1
No connection	20.9	19.7	15.7	11.9
Geographic distance - mean (miles)	505.0	731.2	459.8	217.4
Geographic distance - median (miles)	15.8	40.2	46.9	0.0
Same house value quintile	49.6	36.2	15.2	20.0
Same surname ancestry	51.9	33.8	42.1	42.5
Trading				
Amount invested - average ($1,000s)	4,853.0	2,639.6	1,618.6	1,726.1
Amount invested - median ($1,000s)	200.4	250.1	280.1	492.7
Gross profit - average ($1,000s)	795.9	1,028.9	230.7	1,538.1
Gross profit - median ($1,000s)	17.6	36.3	39.5	86.0
Tip return - average (%)	46.0	43.5	29.2	23.0
Tip return - median (%)	25.2	27.9	28.2	18.8
Use shares dummy (%)	50.8	56.2	77.1	76.4
Use options dummy (%)	27.0	23.4	16.8	21.8
Insider volume/total volume (%)	2.8	4.7	2.9	5.4
Timing				
Time lapse from information to tip (days)	12.1	9.2	5.0	0.4
Tipped passed on same day as received (%)	46.5	62.7	49.5	92.1
Holding period - average (days)	13.9	16.8	11.3	9.1
Holding period - median (days)	5.2	7.0	4.0	5.0
Network Position				
Tippee degree centrality	2.9	2.3	2.0	1.8
Tipper degree centrality	1.8	4.3	5.0	4.6

Table XIII
Tip Chain Position Regressions
This table presents OLS regressions (columns 1 and 2) and order logit regressions (columns 3 and 4) where the dependent variable is the position in the tip chain, where the variable equals one if the observed tipping relationship is the first link from the original source, two if it is the second link from the original source, and so on. The omitted benchmark occupation for tippers and tippees is "Top executive". Heteroskedaticity-robust p-values are reported in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	OLS		Ordered logit	
	Dependent Variable: Position in Tip Chain			
	(1)	(2)	(3)	(4)
Tippee is female	0.206	0.249*	0.519	0.929**
	(0.129)	(0.084)	(0.128)	(0.050)
Tipper is female	−0.176*	−0.369***	−0.201	−0.587
	(0.069)	(< 0.001)	(0.421)	(0.126)
log(Tippee age)	−0.232	0.359*	−0.838*	0.452
	(0.237)	(0.083)	(0.092)	(0.554)
log(Tipper age)	−0.776***	−0.609***	−1.731***	−2.485***
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Family relationship	−0.435***	−0.099	−1.076***	−0.703*
	(0.002)	(0.413)	(0.004)	(0.097)
Friend relationship	−0.478***	−0.221**	−1.016***	−0.702**
	(< 0.001)	(0.020)	(< 0.001)	(0.029)
Business relationship	0.080	−0.029	0.217	−0.257
	(0.368)	(0.732)	(0.370)	(0.396)
No social relationship	−0.505**	−0.142	−1.135*	−0.750
	(0.019)	(0.458)	(0.051)	(0.315)
log(1+geographic distance)	−0.014	−0.017	−0.030	−0.055
	(0.302)	(0.202)	(0.389)	(0.248)
Same surname ancestry	−0.202**	−0.104	−0.636***	−0.428
	(0.019)	(0.163)	(0.006)	(0.107)
Tipper job: Buy side: analyst		1.214***		3.993***
		(< 0.001)		(< 0.001)
Tipper job: Buy side: manager		1.261***		4.150***
		(< 0.001)		(< 0.001)
Tipper job: Low-level employee		0.398***		1.672**
		(0.004)		(0.015)
Tipper job: Corporate manager		0.193*		1.561***
		(0.057)		(0.005)

	OLS		Ordered logit	
	Dependent Variable: Position in Tip Chain			
	(1)	(2)	(3)	(4)
Tipper job: Sell side		0.306***		1.457***
		(0.006)		(0.007)
Tipper job: Small business owner		1.089***		4.290***
		(< 0.001)		(< 0.001)
Tipper job: Specialized		0.624***		3.087***
		(< 0.001)		(< 0.001)
Tipper job: Unknown occupation		0.416***		2.271***
		(< 0.001)		(< 0.001)
Tippee job: Buy side: analyst		0.145		0.563
		(0.298)		(0.403)
Tippee job: Buy side: manager		0.100		0.690
		(0.484)		(0.329)
Tippee job: Low-level employee		−0.062		0.376
		(0.650)		(0.588)
Tippee job: Corporate manager		−0.194		−0.343
		(0.171)		(0.658)
Tippee job: Sell side		0.258		0.837
		(0.143)		(0.273)
Tippee job: Small business owner		0.215*		1.373**
		(0.094)		(0.042)
Tippee job: Specialized		−0.303***		−0.891
		(0.004)		(0.201)
Tippee job: Unknown occupation		−0.072		0.050
		(0.692)		(0.952)
Adjusted R^2	0.168	0.398		
Pseudo R^2			0.089	0.253
Observations	515	515	515	515

Table XIV
Tipper and Tippee Characteristics by Size of Insider Network
This table reports characteristics of people in information networks of increasing size. "Size of Network" refers to the number of people in an insider network (where every member can be reached by every other member through at least one path). Density is the proportion of all possible connections that actually exist. The diameter of a network is the longest of all shortest paths between any two members. Average cluster is the average node's clustering, where clustering is the fraction of a node's links that are also linked to each other.

	Size of Component			
	1	2	3-5	≥ 6
Component Characteristics				
Density		1.0	0.6	0.2
Diameter		1.0	2.3	4.4
Average cluster		2.0	0.2	0.1
Personal Characteristics				
Female tipper (%)	3.4	22.7	11.7	7.2
Female tippee (%)		8.6	10.6	8.5
Tipper age	47.8	45.4	43.6	42.1
Tippee age		44.9	44.5	42.3
Tippee house value - mean ($1,000s)		856.8	893.2	1,311.6
Tippee house value - median ($1,000s)		594.8	476.2	768.0
Tipper house value - mean ($1,000s)	849.6	884.2	796.2	1,548.4
Tipper house value - median ($1,000s)	572.7	560.7	620.6	960.8
Social Connections				
Family connections (%)		43.2	28.2	23.8
Business connections (%)		20.5	34.0	36.4
Friendship connections (%)		36.4	37.8	39.7
Geographic distance - mean (miles)		523.8	326.0	583.6
Geographic distance - median (miles)		7.2	33.6	36.6
Trading				
Amount invested - mean ($1,000s)	725.3	1,054.0	1,227.7	2,457.3
Amount invested - median ($1,000s)	131.6	294.8	154.9	303.1
Gross profit - mean ($1,000s)	2,666.1	1,248.1	255.7	526.6
Gross profit - median ($1,000s)	67.4	154.3	41.7	216.1
Tip return - mean (%)	137.7	82.0	51.0	37.3
Tip return - median (%)	32.4	34.8	34.2	28.0
Timing				
Time lapse from information to tip (days)		11.3	15.4	9.3

Table XV
Selection Bias: Test of the Extent of Prosecution
Panel A of this table presents averages and t-tests of characteristics of individual inside traders based on the extent to which they are prosecuted. The "None" column includes insiders that are identified in a civil case (SEC) or criminal case (DOJ), but not charged as a defendant. "SEC" includes insiders that are defendants in civil cases by the SEC, but not criminal cases by the DOJ. "DOJ" includes insiders that are defendants in criminal cases by the DOJ. Variables are as defined in prior tables. Numbers in parentheses are p-values from t-tests of the difference in means. Panel B presents averages (columns 1 and 2), their difference (column 3), and logit coefficient estimates (columns 4 and 5) where observations are at the pair-level. Logit coefficients reflect the change in the odds-ratios for a one-unit change in the explanatory variable. The "SEC" column includes pairs of insiders in which at least one insider of the pair is a defendant in an SEC case and neither insider is a defendant in a DOJ case. "DOJ" includes pairs in which at least one insider is a defendant in a DOJ case. Coefficient estimates in columns 4 and 5 are from logit regressions where the dependent variable equals one if at least one insider in a pair is a defendant in a DOJ case and zero if at least one insider is a defendant in an SEC case, but neither is a defendant in a DOJ case. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Panel A: Individual Level

	Averages			Differences	
	None	SEC	DOJ	SEC−None	DOJ−SEC
	(1)	(2)	(3)	(4)	(5)
Age	46.85	45.60	39.68	−1.25	−5.92***
				(0.52)	(< 0.01)
Female (%)	29.87	6.97	4.48	−22.90***	−2.49
				(< 0.01)	(0.29)
Married (%)	97.49	81.66	67.91	−15.83***	−13.75***
				(< 0.01)	(< 0.01)
Median house value ($ millions)	1.71	0.88	1.36	−0.84	0.48**
				(0.17)	(0.01)
Median house size (100 ft.^{2})	39.42	28.62	29.31	−10.80	0.69
				(0.16)	(0.71)
Outward connection	0.56	0.62	1.40	0.05	0.78***
				(0.46)	(< 0.01)
Inward connections	0.65	0.65	1.19	0.00	0.54***
				(0.96)	(< 0.01)
Tips given	0.87	1.01	3.55	0.14	2.54***
				(0.48)	(< 0.01)
Tips received	0.98	1.25	2.89	0.27	1.64***
				(0.12)	(< 0.01)
Total invested ($ millions)	2.74	1.18	13.98	−1.56	12.80*
				(0.34)	(0.07)

Average invested per tip ($ millions)	2.31	0.83	3.47	−1.48	2.64***
				(0.35)	(< 0.01)
Total invested/ median house value (%)	34.74	153.31	1,786.00	118.58**	1,632.68
				(0.05)	(0.13)
Total gains ($1,000s)	2.38	1.19	5.37	−1.19	4.18**
				(0.53)	(0.05)
Average gains per tip ($ millions)	0.57	0.98	2.79	0.41	1.81
				(0.54)	(0.30)
Average return (%)	34.75	75.81	54.69	41.06*	−21.12
				(0.08)	(0.52)

Panel B: Pair Level

	Averages		Difference	Logit Regressions	
	SEC	DOJ	DOJ−SEC	Likelihood of DOJ	
	(1)	(2)	(3)	(4)	(5)
Family (%)	31.48	12.30	19.18***	0.200***	0.254***
			(< 0.001)	(< 0.001)	(0.002)
Friendship (%)	37.96	30.16	7.80*	0.440***	0.393**
			(0.077)	(< 0.001)	(0.014)
Business associates (%)	32.9	33.7	−0.9	0.733	0.831
			(0.844)	(0.150)	(0.590)
Geographic distance (miles)	340.09	801.68	−461.59***		1.001*
			(0.001)		(0.060)
Same gender (%)	81.63	83.77	−2.14		0.700
			(0.609)		(0.431)
Age difference	8.57	7.94	0.64		0.986
			(0.588)		(0.434)
Same occupation (%)	18.31	27.50	−9.19**		2.466**
			(0.020)		(0.018)
Same ancestry (%)	50.75	40.11	10.64*		1.008
			(0.060)		(0.981)
Constant				2.415***	2.343
				(0.000)	(0.127)
Observations				468	212
Pseudo R^2				0.065	0.116

Internet Appendix for

"Information Networks: Evidence from Illegal Insider Trading Tips"

I. Data Collection

A. SEC and DOJ Data

The SEC brings civil cases in two ways: civil complaints and administrative proceedings. In many cases, both a complaint and an administrative proceeding are filed. For the purposes of this paper, the legal distinction is unimportant. However, complaints typically include a detailed narrative history of the allegations, including biographies of defendants, trading records, and descriptions of the relationships between tippers and tippees to justify the allegations. In contrast, administrative proceedings typically include far fewer details, providing summaries instead. Therefore, with one exception, I only include cases that have a complaint. The exception is the well publicized case against SAC Capital. The SEC does not explicitly name Steven Cohen in the complaint document, but the related administrative proceeding does.

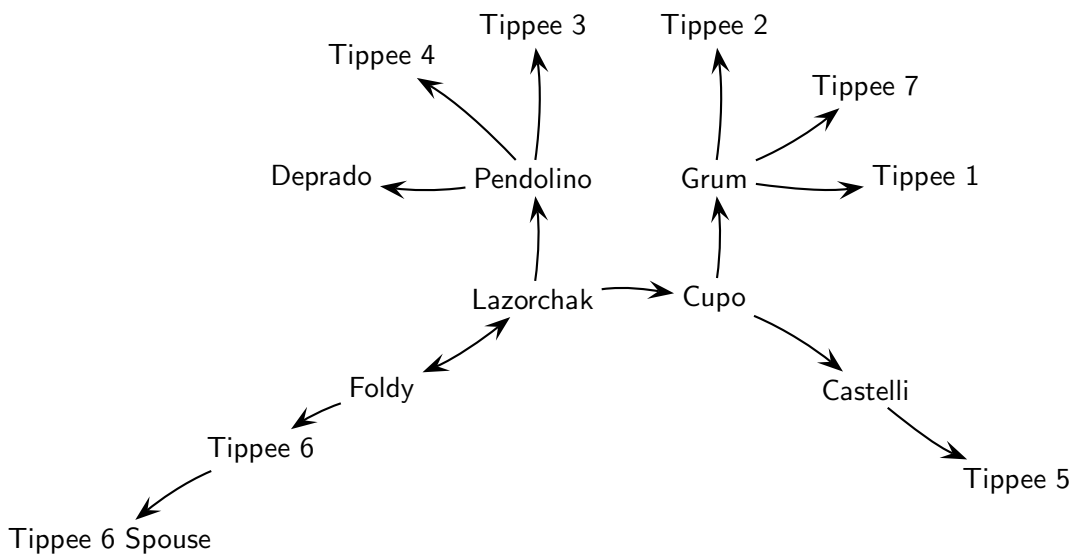
The DOJ cases include a number of different documents. The most useful are the criminal complaints, and "information" documents. These are similar to civil complaints, but contain less information. Transcripts of hearings, while potentially informative, are typically not available on PACER. DOJ cases are usually filed against a single individual. SEC complaints are often filed against multiple people. This means in the DOJ documents, the co-conspirators remain anonymous, but they are named in the SEC documents. This makes the DOJ documents less useful. In some cases, the SEC documents discuss people anonymously too. Often cooperating witnesses are not named until the SEC brings a future case against this person. Therefore, it is necessary to read all of the cases and their amendments in order to piece together the identities of as many people as possible. For instance, in many cases it is easy to infer who the co-conspirator is by the description of their job and relationship to the defendant in connection with another DOJ case in which the co-conspirator is the named defendant. I also rely on investigative journalism in media reports that uncovers the identities of people that the SEC and DOJ does not name explicitly.

B. LexisNexis Public Records Database

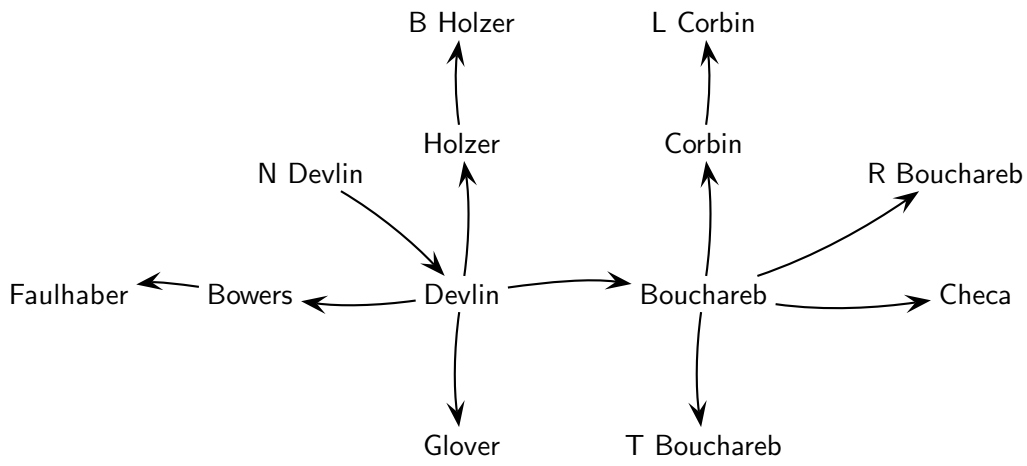
The LexisNexis Public Records Database (LNPRD) provides biographical details on over 300 million individuals, living and dead, that have resided in the United States. These data are taken from state licensing authorities, utility records, criminal files, and property records, among others. The data include month and year of birth, prior addresses, real estate owned, liens and judgments, bankruptcy filings, criminal filings, hunting and sport licenses, and professional licenses.

The LNPRD also includes family members and person associates. Family members are limited to 10 people, but person associates are not limited. Family members are identified as first or second degree relations. Second degree relations are in-laws. The LNPRD also identifies through which relative the in-laws are related.

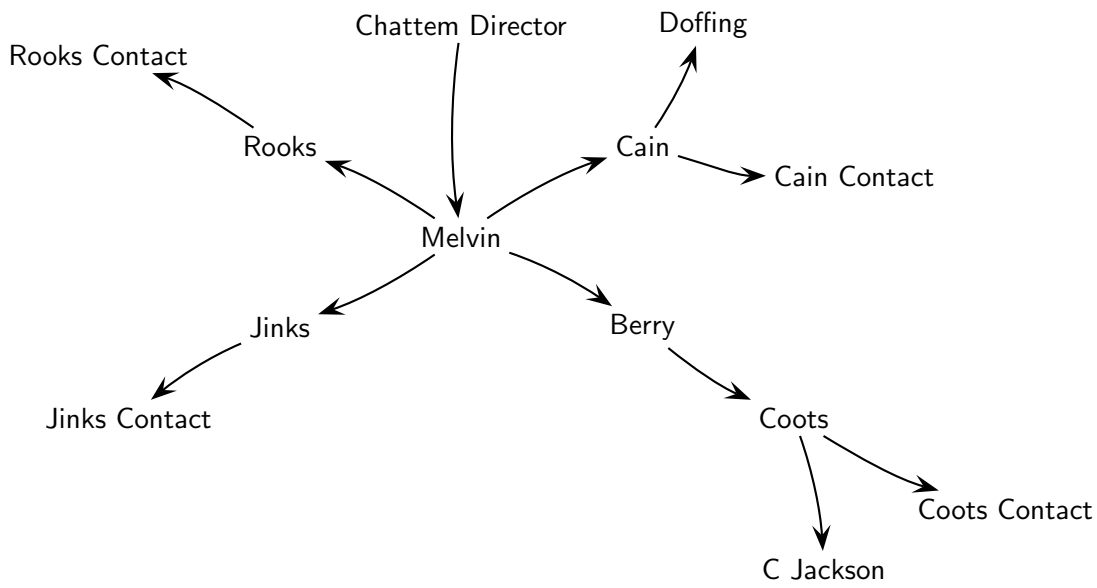
To identify the specific familial relationship of family members listed in the LNPRD, I use the age, addresses, and real estate ownership records to identify relationships. For example, for a male insider, if the LNPRD lists a first degree female family member with the same last name, plus a maiden name, of about the same age as the male insider, the family member could be a wife or a sister. If the male insider and the female family member own real estate in which they both reside, I code the female family member as a wife. Otherwise, I could the relative as a sister. In this case, if a second degree male relative of the female relative is listed who has the same last name as the female relative and shares a residence, I assume this is the husband or the female relative, which implies the female relative is a sister. As another example, children are typically 20 to 40 years younger than the observed person in the LNPRD and when young enough, they have a history of sharing the same address. I also use external sources, such as online obituaries, to assess family relationships, when available.



Internet Appendix Figure 1
Lazorchak-Foldy Network
This figure represents the illegal insider trading network centered on John Lazorchak and Mark Foldy who shared inside information with others. Arrows represent the direction of information flow. Data are from SEC and DOJ case documents, plus additional source documents to identify individuals not named in the SEC and DOJ documents.



Internet Appendix Figure 2
Devlin Network
This figure represents the illegal insider trading network centered on Matthew Devlin who received inside information from his wife, Nina Devlin. Arrows represent the direction of information flow. Data are from SEC and DOJ case documents, plus additional source documents to identify individuals not named in the SEC and DOJ documents.



Internet Appendix Figure 3
Melvin Network
This figure represents the illegal insider trading network centered on Thomas Melvin who shared inside information from an unnamed Chattem Director with others. Arrows represent the direction of information flow. Data are from SEC and DOJ case documents, plus additional source documents to identify individuals not named in the SEC and DOJ documents.

Internet Appendix Table I
Event Types
This table provides a breakdown of the types of corporate events in the sample. The data include 465 corporate events over the years 1996 to 2013 from SEC and DOJ case documents.

Clinical Trial/Drug Regulation	37
Clinical trial results	7
Regulatory approvals and rejections	30
Earnings	123
Earnings announcements	112
Earnings guidance	7
Earnings restatements	4
M&A	239
Acquisition announcement (trading in target stock)	216
Acquisition announcement (trading in acquirer stock)	3
Merger negotiation developments	3
Failed acquisitions	4
Joint ventures	3
Licensing	2
Restructuring	3
Target seeking buyer	1
Share repurchase	1
Strategic alliances/investments	3
Operations	13
Appointment/resignation of CEO	2
Business agreement/contract	9
Layoff	1
Other operations	1
Sale of Securities	35
Other	18
Analyst report	1
Dividend increase	1
Financial distress	2
Fund liquidation	1
Addition to stock index	1
Unspecified	12
Total	465

Internet Appendix Table II
Events by Industry
This table presents the frequency of illegal insider trading events by two-digit NAICS level industry definitions, using data from 465 corporate events over the years 1996 to 2013. "Example firms in sample" presents non-exhaustive lists of sample firms for each industry definition.

Industry	Number of events	Fraction of total	Example firms in sample
Computers and electronics	91	23.1	Dell, Intel, Nvidia, Sun Microsystems
Chemical manufacturing	90	22.8	Celgene, MedImmune, Pharmasset, Sepracor
Food and apparel manufacturing	40	10.2	Carter's, Green Mountain Coffee Roasters, Smithfield Foods
Information	39	9.9	Akamai, Autodesk, Clearwire, Google, Microsoft
Finance and insurance	27	6.9	East West Bancorp, Goldman Sachs, Mercer Insurance Group
Professional and technical services	20	5.1	Alliance Data Systems, F5 Networks, Perot Systems
Wholesale trade	18	4.6	Herbalife, InVentiv Health, World Fuel Services
Mining and oil extraction	13	3.3	Delta Petroluem, Mariner Energy, Puda Coal
Retail trade	12	3.0	Albertson's, Best Buy, J. Crew, Walgreen Company
Health care	8	2.0	Humana, LCA Vision, Option Care
Real estate	8	2.0	Bluegreen Corporation, Mitcham Industries
Administrative support	8	2.0	Comsys IT Partners, DynCorp International, PeopleSupport
General merchandise stores	6	1.5	Barnes & Noble, Office Depot, Sears
Transportation	4	1.0	Airtran Holdings, K-Sea Transportation Partners
Accomodation and food services	4	1.0	CKE Restaurants, Hilton, Rubio's Restaurants
Utilities	2	0.5	SouthWest Water Company, TXU Corp.
Construction	2	0.5	Complete Production Services, Warrior Energy Services
Education	1	0.3	Global Education and Technology Group Limited
Other	1	0.3	Berkshire Hathaway Inc.

Internet Appendix Table III
Job Titles of Insiders
This table provides a breakdown of the job titles in the sample. The data include 611 people involved in insider trading over the years 1996 to 2013 from SEC and DOJ case documents.

Top executive	107
Chairman	11
Board member	14
CEO	12
CFO	7
Officer	64
Mid-level corporate manager	55
Lower-level employee	59
Unspecified	40
Secretary	8
IT	11
Sell side: lawyer, accountant, investment banker	59
Investment banker	4
Sell-side analyst	5
Attorney	24
Accountant	13
Principal	15
Buy side: manager	60
Hedge fund manager	30
Portfolio manager	29
Investor	2
Venture capitalist	1
Buy side: analyst, trader	65
Trader	43
Buy-side analyst	22
Small business owner	39
Small business owner	35
Real estate broker	4
Specialized occupation	38
Consultant	16
Doctor	13
Engineer	9
Unknown	135
Total	622

Internet Appendix Table IV
The Likelihood that an Insider Tips Someone Else
This table presents coefficient estimates from logit tests of the likelihood that an insider tips anyone else. Observations are the 622 insiders in the sample, reduced by data availability. Heteroskedaticity-robust p-values are reported in parentheses. P-values from robust standard errors are presented in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	(1)	(2)	(3)	(4)
Female	0.302	0.759*	0.589	0.581
	(0.378)	(0.061)	(0.221)	(0.231)
Log(age)	−1.365***	−1.894***	−2.166***	−2.246***
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Number of family members		0.089**	0.065	0.065
		(0.027)	(0.208)	(0.210)
Number of person associates		0.044***	0.023	0.026
		(0.005)	(0.176)	(0.145)
Buy side: analyst			−0.043	0.049
			(0.931)	(0.921)
Buy side: manager			−1.393***	−1.436***
			(0.003)	(0.003)
Low-level employee			−0.556	−0.500
			(0.204)	(0.263)
Corporate manager			−0.428	−0.472
			(0.335)	(0.294)
Sell side			−0.460	−0.474
			(0.288)	(0.287)
Small business owner			−0.766*	−0.672
			(0.092)	(0.137)
Specialized			−1.226**	−1.227**
			(0.013)	(0.014)
Unknown occupation			−1.141*	−1.049*
			(0.056)	(0.083)
Married			0.075	0.103
			(0.676)	(0.578)
Log(median house value)			0.196	0.186
			(0.166)	(0.196)
Criminal record				−0.437
				(0.131)

	(1)	(2)	(3)	(4)
Bankruptcy				−0.436
				(0.518)
Liens and judgments				−0.236
				(0.484)
UCC liens				−0.014
				(0.962)
Constant	5.128***	6.243***	5.463*	6.003**
	(< 0.001)	(< 0.001)	(0.059)	(0.043)
Pseudo R^2	0.024	0.064	0.089	0.096
Observations	447	407	334	334

Internet Appendix Table V
Determinants of a Tipper's Number of Tippees
This table presents coefficient estimates from poisson regressions on the number of tipping relationships in which an insider gives inside information to another person. The dependent variable is the number of tippees that are connected to the tipper. Observations are the 622 insiders in the sample, reduced by data availability. Heteroskedaticity-robust p-values are reported in parentheses. P-values from robust standard errors are presented in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	(1)	(2)	(3)	(4)
Female	0.118	0.252	0.189	0.186
	(0.623)	(0.275)	(0.445)	(0.457)
Log(age)	−0.774***	−0.962***	−0.701**	−0.716**
	(0.005)	(< 0.001)	(0.045)	(0.035)
Number of family members		0.026	0.004	0.003
		(0.343)	(0.903)	(0.926)
Number of person associates		0.014*	0.006	0.006
		(0.076)	(0.468)	(0.501)
Buy side: analyst			0.331	0.380*
			(0.133)	(0.089)
Buy side: manager			−0.303	−0.305
			(0.378)	(0.396)
Low-level employee			−0.202	−0.158
			(0.379)	(0.484)
Corporate manager			−0.239	−0.235
			(0.315)	(0.328)
Sell side			−0.254	−0.234
			(0.288)	(0.338)
Small business owner			−0.114	−0.101
			(0.725)	(0.759)
Specialized			−0.495	−0.477
			(0.197)	(0.217)
Unknown occupation			−0.740**	−0.742**
			(0.033)	(0.034)
Married			−0.134	−0.126
			(0.183)	(0.232)
Log(median house value)			0.153	0.135
			(0.121)	(0.163)
Criminal record				−0.117
				(0.515)
Bankruptcy				−0.496

	(1)	(2)	(3)	(4)
				(0.182)
Liens and judgments				−0.115
				(0.600)
UCC liens				0.188
				(0.247)
Constant	2.689***	3.140***	0.609	0.897
	(0.008)	(0.002)	(0.743)	(0.632)
Pseudo R^2	0.014	0.024	0.047	0.052
Observations	447	407	334	334

Internet Appendix Table VI
Original Source in Other Events

This table presents the frequency of different original sources of inside information in information events other than M&As and earnings announcements. These events include announcements about clinical trial results, operations, sale of securities, and others. "Tip Firm"' refers to the firm for which the information is relevant. Internal sources are those that are employed by the tip firm. External sources include people that are employed by firms other than the merging firms. Service providers include human relations firms, investor relations firms, medical consultants, law firms, credit rating agencies, and others.

	Tip Firm	Failed/ Potential bidder	Regulatory Agency	Other/ Unknown	Total
Internal sources					
Director	2	0	0	0	2
Officer	9	0	0	0	9
Employee	12	0	0	3	15
Total internal sources	23	0	0	3	26
External sources					
Potential investor	0	34	0	0	34
Regulatory agency	0	0	27	0	27
Service providers	10	0	4	2	16
Total external sources	10	34	31	2	77
Unknown	0	0	0	3	3
Total	33	34	31	8	106

Internet Appendix Table VII
Tipper and Tippee Occupations by Order in Tip Chain: Long Versus Short Chains
This table presents the distribution of occupations for each link in a tip chain. The position in a tip chain is the distance from the original source of the inside information, where distance is measured as information links between people. "1" indicates the first link in the chain from an original source to all the people to whom the source shared the information. "2" indicates the links from the first people to receive the tip from the original source to all of the people with whom they share the information. "3" and "≥ 4" are defined analagously. Entries in the table present the fraction of all people at a given position in the tip chain that are employed in each of the listed occupations. Thus, columns sum to 100%. "Number" indicates the number of relationships for each position in the tip chain. "Short Tip Chains" have four or fewer members, thus, they have a maximum of three links in the tip chain. "Long Tip Chains" have more than four members.

	Short Tip Chains			Long Tip Chains			
Position in Tip Chain	1	2	3	1	2	3	≥ 4
Tippee Occupation							
Top executive	10.0	4.6	14.3	6.4	0.6	0.9	0.0
Corporate manager	4.4	1.5	0.0	9.6	2.5	0.9	0.0
Lower-level employee	11.2	14.5	0.0	4.3	11.3	11.1	0.0
Sell side/lawyer/accountants	11.5	16.0	28.6	8.5	12.6	12.0	7.3
Buy side: manager	12.5	22.9	21.4	12.8	21.4	15.4	25.5
Buy side: analyst/trader	16.5	13.0	14.3	28.7	15.7	29.9	34.5
Small business owner	10.9	9.2	0.0	8.5	11.9	12.0	3.6
Specialized occupation	10.9	2.3	0.0	8.5	6.9	3.4	0.0
Unknown	12.1	16.0	21.4	12.8	17.0	14.5	29.1
Number	321	131	14	94	159	117	55
Tipper Occupation							
Top executive	34.3	4.6	7.1	33.0	8.2	0.0	0.0
Corporate manager	12.5	9.2	0.0	18.1	8.2	3.4	0.0
Lower-level employee	16.5	6.1	14.3	14.9	4.4	11.1	3.6
Sell side/lawyer/accountants	25.5	6.9	42.9	28.7	6.9	28.2	7.3
Buy side: manager	0.3	8.4	14.3	3.2	20.8	17.9	12.7
Buy side: analyst/trader	2.5	22.1	0.0	0.0	32.7	30.8	70.9
Small business owner	0.9	4.6	7.1	0.0	8.8	7.7	1.8
Specialized occupation	1.9	20.6	14.3	0.0	7.5	0.0	0.0
Unknown	5.6	17.6	0.0	2.1	2.5	0.9	3.6
Number	321	131	14	94	159	117	55

Internet Appendix Table VIII
Selection Bias: The Extent of Prosecution and Occupation
This table presents counts and percentages of inside traders' occupations based on the extent to which they are prosecuted. The "None" column includes insiders that are identified in a civil case (SEC) or criminal case (DOJ), but not charged as a defendant. "SEC" includes insiders that are defendants in civil cases by the SEC, but not criminal cases by the DOJ. "DOJ" includes insiders that are defendants in criminal cases by the DOJ. The top entry is the count of insiders by occupation. The bottome entry in brackets is the percentage of insiders by occupation, within each column, not including unknown occupations.

	None	SEC	DOJ
	(1)	(2)	(3)
Top executive	24	62	21
	[25.3]	[24.0]	[16.0]
Corporate manager	7	35	13
	[7.4]	[13.6]	[9.9]
Lower-level employee	13	39	7
	[13.7]	[15.1]	[5.3]
Sell side/lawyer/accountants	19	26	16
	[20.0]	[10.1]	[12.2]
Buy side: manager	18	19	23
	[18.9]	[7.4]	[17.6]
Buy side: analyst/trader	9	24	32
	[9.5]	[9.3]	[24.4]
Small business owner	1	30	8
	[1.1]	[11.6]	[6.1]
Specialized occupation	4	23	11
	[4.2]	[8.9]	[8.4]
Unknown	101	31	3

Internet Appendix Table IX
Selection Bias: Logit and Ordered Logit Tests

This table presents coefficient estimates (in odds-ratios) of logit and ordered logit regressions. Columns 1–3 are ordered logit regressions where the dependent variable takes on three values: zero for insiders that are identified in a civil case (SEC) or criminal case (DOJ), but not charged as a defendant; one for insiders that are defendants in civil cases by the SEC, but not criminal cases by the DOJ; and two for insiders that are defendants in criminal cases by the DOJ. Variables are as defined in prior tables. Columns 4–6 present coefficient estimates from logit regressions where the dependent variable equals one if an insider is charged by the DOJ and zero if the insider is charged by the SEC, but not the DOJ. Numbers in parentheses are p-values from robust standard errors. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

	Ordered Logit Extent of Prosecution			Logit Likelihood of DOJ		
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.966***	0.973	0.960***	0.961***	0.965	0.953***
	(0.000)	(0.253)	(0.000)	(0.001)	(0.187)	(0.000)
Female	0.105***	0.043***	0.123***	0.370**	0.000**	0.470
	(0.000)	(0.003)	(0.000)	(0.048)	(0.019)	(0.217)
Married	0.530**	0.542	0.549**	0.674	0.906	0.735
	(0.014)	(0.265)	(0.020)	(0.172)	(0.878)	(0.311)
Tips given	1.181***	1.316**	1.181***	1.207***	1.279**	1.209***
	(0.002)	(0.011)	(0.001)	(0.002)	(0.013)	(0.002)
Tips received	1.111**	1.014	1.087*	1.132**	1.119	1.109
	(0.044)	(0.905)	(0.083)	(0.049)	(0.284)	(0.124)
Median house value ($1,000,000s)		1.436***			1.767**	
		(0.008)			(0.048)	
Total invested ($1,000,000s)		1.106*			1.118	
		(0.063)			(0.128)	
Average return (%)		1.240*			1.227*	
		(0.058)			(0.083)	
Buy side: manager			1.366			1.734
			(0.518)			(0.238)
Lower-level employee			0.261***			0.208***
			(0.002)			(0.005)
Corporate manager			0.460*			0.457
			(0.098)			(0.172)
Sell side/lawyer/accountants			0.408*			0.666
			(0.056)			(0.395)
Small business owner			0.640			0.538
			(0.307)			(0.241)
Specialized occupation			1.042			1.068
			(0.931)			(0.904)
Top executive			0.655			0.854
			(0.309)			(0.723)
Unknown			0.261***			0.142**
			(0.004)			(0.031)

Constant				1.860	0.446	3.821**
				(0.247)	(0.561)	(0.032)
Cut 1	0.011***	0.003***	0.004***			
	(0.000)	(0.000)	(0.000)			
Cut 2	0.542	1.595	0.258***			
	(0.170)	(0.710)	(0.009)			
Observations	447	157	444	406	154	406
Pseudo R^2	0.138	0.252	0.175	0.152	0.295	0.207