University of California
Los Angeles

Essays Concerning the Network Structure of Mutual Fund Holdings and the Behavior of Institutional Investors

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Management

by

Phillip Stephen Wool

2013

Abstract of the Dissertation

Essays Concerning the Network Structure of Mutual Fund Holdings and the Behavior of Institutional Investors

by

Phillip Stephen Wool

Doctor of Philosophy in Management

University of California, Los Angeles, 2013

Professor Richard W. Roll, Chair

In the first chapter of this dissertation, I describe a method for representing institutional investors' portfolio holdings as a graph, in which funds connect to stocks through patterns of common ownership. I then demonstrate that changes to a firm's position within this network are closely related to future stock market performance. Specifically, stocks moving toward the center of the holdings network outperform those drifting toward the periphery by approximately 4.1%, annually, adjusting for standard risk factors, consistent with a model in which short-sale constraints combined with increasing dispersion in investors' beliefs signal potential overvaluation. After controlling for a number of additional variables, including the "breadth of ownership" measure proposed by Chen, Hong, and Stein (2002)—a local indicator of a firm's network importance—stocks with the largest decrease in holdings network centrality still underperform by 2.2% per year.

In the second chapter, using a novel data set consisting of Schedule 13D filings and amendments over a seven-year period, from 2003 to 2010, I present evidence

that managers of large investment portfolios exploit periods of perceived investor distraction to minimize the adverse impact of the disclosure of large sales on future transactions. Specifically, managers reporting substantial decreases in holdings favor Friday disclosure over disclosure on other weekdays, and prefer to release the news in the hours after markets close. Moreover, investors who go on to make future sales are significantly more likely to pursue an opportunistic filing strategy. Employing event study methodology, I test for underreaction to Friday filings, but find no support for investor inattention to Friday 13D disclosures. Investors seem to rapidly incorporate available information from regulatory disclosures into stock prices, correctly attributing heavy selling to liquidations and not informed trading.

The dissertation of Phillip Stephen Wool is approved.

Antonio E. Bernardo

Mark S. Grinblatt

Mark S. Handcock

Richard W. Roll, Committee Chair

University of California, Los Angeles

2013

To Mom, Dad, Josh,

Phyllis, Dale,

Helen, and Sammy.

Contents

List of Tables

List of Figures

This page is intentionally blank.

Preface

The chapters of this dissertation document my investigation into a number of problems that have occupied my mind in the course of the last six years. The research efforts underlying each essay took place over distinct spans of time and employ markedly different methods, although it would be wrong to regard the two chapters as independent studies of separate phenomena. Rather, one should view both chapters as reflections of a core assertion that our understanding of the means by which financial markets assign prices to assets will never be complete without due consideration of the manifold ways in which the characteristics and behaviors of institutional investors—including those measured along psychological, sociological, and political dimensions—affect the process of price formation.

Scholarly research is, like most worthwhile human endeavors, the work of an entire community, and I am compelled at the outset to recognize the contribution of so many other people who helped me to satisfy the very personal intellectual curiosities that motivated the explorations I describe below. I am forever grateful for the lifelong love and support of my parents, Kim and Elliot Wool, who have always encouraged me to pursue my passions, and who continue to provide the environment, resources, and confidence that make such ambitions and achievements possible. My brother and best friend, Josh Wool, is as much a scientist as anyone I know, and he has served as an invaluable adviser and collaborator throughout my journey as a doctoral student.

His experiences as a practitioner were also the direct inspiration for my research into patterns in the timing of Schedule 13D amendment disclosures, which makes up the second chapter of this dissertation. I also thank my grandparents—Phyllis and Dale Clements, and Helen and Sam Wool—for supporting me throughout this process. Einstein supposedly remarked that one does not truly understand a subject until one is able to explain it to one's grandmother. I have confirmed this claim in almost daily conversations with Phyllis about everything to do with my work: from issues of computation to the deepest economic intuitions behind my doctoral research. I would be remiss if I did not acknowledge the support of my great friend, Patrick Keefe, who risked his life helping me move on the highways from St. Louis to Los Angeles, and then made annual trips with me to the sports books and poker rooms of Las Vegas. Finally, I thank Ann Kwak, for tolerating my long phone calls to Josh, and for making visits to my brother in Dallas all the more fun.

It was an immense pleasure completing this work under the supervision of my committee chair, Richard Roll, who always had an open door, and whose diverse intellectual interests never cease to impress and inspire me. Tony Bernardo is one of the most incisive and creative minds I have encountered in my years as a student, and Mark Grinblatt's financial intuition and practical knowledge of markets are sources of amazement. Moreover, each man exhibits a selfless dedication to the success of his students. I thank my outside committee member, Mark Handcock, for graciously coming aboard on short notice and providing thoughtful insights on both essays. Avanidhar (Subra) Subrahmanyam served as an early mentor in the program and I continue to look up to him as a researcher, teacher, and enthusiastic promoter of good science. I am also fortunate, over the last three years, to have worked with and learned from Jason Hsu, who has made me a better instructor and significantly influenced my understanding of investment practice. Rick Grannis assisted in my introduction to the rigorous quantitative study of network structures, which forms the basis for the methods used in the first chapter of this dissertation.

Of course, many other scholars—from UCLA Anderson Finance faculty, to researchers in different departments and at other institutions—have positively influenced my work. In particular, my research benefits from feedback by participants in Anderson's Third-Year Doctoral Student Seminar, and from insightful questions and comments posed during and after my UCLA Finance Brown Bag talk. In addition, I am thankful for a number of helpful conversations with those attending presentations of my work at George Mason University and the University at Buffalo.

My life at UCLA was made considerably more efficient and enjoyable by the tireless efforts of Lydia Heyman, Director of Anderson's Doctoral Program Office, and Delores Rhaburn, the Finance Department Administrative Coordinator, who was particularly helpful during my year on the job market. I also acknowledge the generous financial support for my doctoral studies provided by the Anderson School of Management and UCLA's Graduate Division.

Lastly, I recognize the contributions of my friends and colleagues in the Finance PhD program, particularly Shaun Davies, Matthias Fleckenstein, Andrew Iannaccone, Patrick Kiefer, Yaron Levi, Jiasun Li, Kyle Matoba (who shares too many of my intellectual interests to count and frequently consulted on my statistical and computational questions), Michael Nowotny, Aurélien Philippot (the greatest cube-mate of all time), Rob Richmond, Brian Waters (fellow sports economics enthusiast and my favorite critic of Finance seminar presentations), and Mindy Zhang. The stimulating conversations we had while "working" in the PhD student office are some of my fondest memories of the last six years.

This page is intentionally blank.

Vita

2003	A.B., Economics, Washington University in St. Louis Honors in Economics, *magna cum laude*
2003	B.S.B.A., Finance and Accounting, Washington University in St. Louis Honors in Management, *magna cum laude*
2003–2005	Research Analyst, Alternative Investments Hammond Associates Institutional Fund Consultants
2005–2006	Non-degree study, Mathematics, Washington University in St. Louis
2007	Chancellor's Prize, University of California at Los Angeles
2007-2010	University Fellowship, University of California at Los Angeles
2012	Dissertation Year Fellowship, University of California at Los Angeles

This page is intentionally blank.

Chapter 1

The Network of Mutual Fund Holdings: Stock Centrality and Future Returns

1.1 Introduction

This study begins with the observation that networks are everywhere: from the patchwork of interpersonal connections that form the worldwide social network, to linkages among the billions of hardware components that allow communication over the Internet. Even in the absence of physical networks, many human activities trace out implicit relational structures which give rise to useful graphical representations and network-theoretic modes of analysis. In this chapter, I investigate an evolving holdings network comprised of overlapping equity ownership ties among active U.S. mutual fund managers. Applying appropriate measures of network structure to the graph of mutual fund portfolios, I show that changes to a firm's position within the network of institutional investors' holdings are closely related to future stock market performance, suggesting an enhanced role for relational models of investor behavior, and highlighting the importance of the network perspective for gaining a better understanding of complex and highly connected financial markets.

As a means of demonstrating that patterns in the market-wide configuration of fund holdings encode information relevant to the pricing of financial assets, I apply my network-based approach to a traditional model built on the intuition, expressed by Miller (1977), that in the presence of binding short-sale constraints, the trading of optimists will exert a disproportionate influence on stock prices, leading to lower future returns. Along these lines, Chen, Hong, and Stein (2002) present a formal model in which short-sale restrictions combined with increased dispersion in investors' beliefs serve as an indication of possible overvaluation, since heightened disagreement suggests that more investors would sell, were it not for a reluctance or inability to establish short positions. As a proxy for the degree of consensus among investors, the authors propose using a firm's breadth of ownership, a simple count of the fraction of mutual funds with long positions in the stock, and confirm that a stock sold by many mutual funds in a given quarter significantly underperforms over the subsequent year. The authors benchmark this proxy against a similar measure, introduced by Chen, Jegadeesh, and Wermers (2000), who sort stocks on the fraction of each firm's shares owned by mutual funds, and conclude that breadth of ownership dominates as a predictor of future stock returns.

In fact, the Chen, Hong, and Stein (2002) measure has a natural interpretation in the context of a network comprised of ownership ties between funds and stocks, as an indicator of a stock's immediate ties to mutual fund portfolios. Intuitively, when a stock has links to most funds, one might reasonably conclude that managers tend to agree on its investment merits. An important weakness of local network measures, however, is that they discard the information bound up in connections spanning more than one link—despite the fact that the existence of subtle, network-wide dependencies is a principal reason for turning to relational data, in the first place. Although a node's immediate ties say something about its function in the network, they cannot tell us the whole story. While breadth of ownership identifies a stock as "popular" to the extent that it appears in the portfolios of many fund managers, it fails to rate as prominent those stocks that show up in just a few portfolios, but always alongside other highly popular stocks. Moreover, large company stocks attain high breadth not as a result of consensus in investors' beliefs, but simply because a high market capitalization supports a broader base of investors.

Recognizing these limitations, I propose a global network measure of stock importance that incorporates information from the entire set of interlocking mutual fund portfolio holdings to provide a clearer picture of each stock's position in the tacit rankings of active fund managers. I begin by defining a connection between a pair of stocks if that pair appears more often in investors' portfolios than one would expect by chance, given the number of funds holding each firm's shares. Intuitively, such connections reflect a latent preference on the part of managers to hold these stocks together, and permit the transmission of investors' information or sentiment from one stock to another. Based on this set of relationships, my holdings network centrality score identifies a stock as popular to the extent that it connects to many other stocks in the network, particularly if its neighbors are, themselves, very popular. This definition of prominence acknowledges the value of indirect chains of ownership and mitigates the mechanical effect of firm size on a stock's status with investors.

This graph-theoretic notion of a stock's prominence within the network of overlapping mutual fund portfolios, along with the model presented in Chen, Hong, and Stein (2002), gives rise to a number of testable hypotheses. First, the model implies that a stock's declining popularity with mutual fund managers signals increased disagreement as to its valuation which, because of binding short-sale constraints, indicates potential overpricing and predicts lower future performance. If holdings network centrality is thought to capture the relative importance of stocks in the portfolios of mutual fund managers, then one would expect changes in centrality to have similar implications for future stock returns. Moreover, while conceptual similarities should lead to correlation between changes in stock centrality and innovations in the measures of stock breadth and mutual fund ownership (MFO) suggested by Chen, Hong, and Stein (2002) and Chen, Jegadeesh, and Wermers (2000), if features of the network's global structure are important for adequately gauging a stock's prominence—and, hence, according to the model, the consensus among mutual fund managers—then holdings network centrality should provide predictive value above and beyond that of local indicators like breadth and MFO. Accordingly, much of this chapter will be concerned with tests of the following propositions:

Hypothesis 1.1. *A large decrease in a stock's holdings network centrality during the previous quarter should predict lower returns in subsequent months, controlling for other variables known to predict future returns.*

Hypothesis 1.2. *Controlling for changes in the fraction of funds holding a stock and the fraction of a stock's shares held by mutual funds should diminish, but not eliminate, the predictive power of changes in holdings network centrality.*

My work draws on concepts and methodology common to the field of social network analysis, although I am certainly not the first to apply such thinking to financial questions. A number of authors in recent years have studied the relationship between social network ties and investor behavior. Hong, Kubik, and Stein (2004) present evidence that an individual's level of social interaction increases the likelihood of that individual's participating in the stock market. Ivković and Weisbenner (2007) use discount brokerage data covering the trades of roughly 36,000 retail investors to demonstrate that word-of-mouth effects among neighbors lead to significant commonalities in portfolio holdings, particularly with respect to investments in locally headquartered firms. In the domain of institutional investors, Hong, Kubik, and Stein (2005) find that mutual fund companies based in the same city have similar holdings and trade individual stocks in the same direction, on average, attributing the effect to social contact among the funds' managers. Cohen, Frazzini, and Malloy (2008) report that mutual fund managers disproportionately invest in firms run by individuals who attended the same university, that the funds earn more on these trades, and that most of the profits result from news events, supporting the hypothesis that these investments exploit information flow over social network connections.

This work presented in this chapter differs from such research in several important ways. First, each group of authors cited above begins with evidence of a social connection (e.g., geographic proximity, school ties), and proceeds to draw conclusions about the linked individuals' investing behavior. I work in the opposite direction, using holdings data to infer potential ties among investors, then feeding these associations into network-based measures that apply to a range of empirical questions. Moreover, I do not restrict myself to analyzing connections among funds and stocks based on direct social relations. It is quite possible, for example, that ties uncovered in my analysis reflect investors' exposure to the same sources of information—as a result of reading the same news stories or attending the same conferences—similar styles of thinking and fund management, herding behavior, or common susceptibility to psychological bias, in addition to relationships based on direct social contact.

There have been other efforts to construct network representations of financial markets on the basis of portfolio holdings and transactions data. An oft-referenced example is the paper of Hochberg, Ljungqvist, and Lu (2007), who define a network of venture capitalists linked by coinvestment in VC projects, and conclude that better-networked venture capital funds have stronger performance and lead to better outcomes for portfolio companies. More recently, Ozsoylev et al. (2011) form a network using account-level data for all trades executed on the Istanbul Stock Exchange in 2005, placing an edge between two traders if they submit orders in the same direction at around the same time, and show that more central individuals systematically trade before peripheral actors, and tend to earn larger profits. In this chapter, I infer network structure from quarterly snapshots of static mutual fund portfolios, in part because the irregularity of holdings disclosures makes it relatively difficult to reliably match trades in time.

Solis (2009) represents mutual fund holdings as a bipartite graph, as I do, although he restricts his analysis to 18 Vanguard and Fidelity family funds and 99 unique stocks comprising each manager's ten largest positions, making it impossible to compute anything but local network summary statistics which, due to the small and biased sample, cannot be assumed to generalize to the broader set of stocks and funds. In a more comprehensive study, restricting attention to the indirect connections among stocks resulting from common ownership by mutual funds, Antón and Polk (2012) find that pairs of stocks held in many of the same portfolios experience higher future comovement in returns, and attribute the effect to correlated trading by mutual funds. Blocher (2013) takes a similar approach to identifying pairs of mutual funds linked by common portfolio holdings, and estimates that spillover effects associated with the fund flows of an investor's network neighbors account for roughly 2% of that fund's quarterly performance, on average.

Perhaps the closest research to my own appears in an article by Cohen, Coval, and Pástor (2005), who devise recursive measures of stock quality and manager performance based on holdings overlap in a manner strikingly similar to the calculation of eigenvector centrality in Section 1.2. Essentially, high-quality stocks are those held by high-talent managers, where a manager's talent is a function of the quality of stocks in that manager's portfolio. In contrast to my approach, the authors incorporate estimates of mutual fund alpha into the calculation of stock and fund status, and focus on assessing the ability of these scores to forecast future mutual fund performance. More importantly, despite the close correspondence between their framework and my graphical representation of holdings overlap, because Cohen, Coval, and Pástor (2005) do not take an explicit network perspective on the connections among funds and stocks, they overlook opportunities to extend this relational model in pursuit of answers to other interesting questions that might benefit from such an approach. My work constitutes a significant step in that direction.

The remainder of the chapter is organized as follows. Section 1.2 builds up a mathematical framework for representing the network of overlapping mutual fund stock holdings as a graph, and defines measures to summarize structural features of that network, including the relative positions of its constituent funds and stocks. Section 1.3 describes the data I use in subsequent analyses, and details the construction of an empirical holdings network based on quarterly portfolio disclosures by active mutual fund managers. In Section 1.4, I conduct asset pricing tests of the hypotheses stated above. Section 1.5 concludes with a summary of my findings, some remarks on the broader significance of these results, and suggestions for future research.

1.2 Holdings Networks

In order to quantify the relational structure of stock ownership, I model the network of overlapping mutual fund portfolio holdings at any point in time as a graph. This section establishes notation and provides a formal mathematical description of the network structures and measures that will facilitate the empirical tests presented in the remainder of the chapter. In the process, I show that the Chen, Hong, and

Stein (2002) breadth of ownership measure has a simple graphical interpretation, and suggest an alternative approach motivated by network considerations.

1.2.1 Holdings as a Graph

Let the collection of funds correspond to nodes in the set $F = \{f_1, f_2, \ldots, f_{n_F}\}$, and let the universe of stocks correspond to nodes in the set $S = \{s_1, s_2, \ldots, s_{n_S}\}$. An undirected edge exists between a mutual fund and a stock if the fund holds shares in that stock. These two sets of nodes and the collection of edges defined among them constitute a two-mode, bipartite graph, in which vertices from the first mode (the set of stocks) may only connect with vertices from the second mode (the set of funds). It is convenient to encode the relationships among stocks and funds in an $n_S \times n_F$ holdings network affiliation matrix, $\mathbf{A} = \{a_{ij}\}$, with elements defined by

$$a_{ij} = \begin{cases} 1 & \text{if fund } j \text{ holds shares of stock } i, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

While the affiliation matrix records direct linkages among stocks and mutual funds, it will also be useful to consider a projection of these relationships onto just the collection of stocks. Based on holdings at a given point in time, the $n_S \times n_S$ stock network co-affiliation matrix, expressed as $\mathbf{X}^S = \{x_{ij}^S\}$, has elements given by

$$x_{ij}^S = \sum_{k=1}^{n_F} a_{ik} a_{jk} \quad \Leftrightarrow \quad \mathbf{X}^S = \mathbf{A}\mathbf{A}^\top. \tag{1.2}$$

The co-affiliation matrix defined above reflects the fact that a pair of stocks is connected when those stocks appear together in the portfolio of one or more mutual fund managers. The entries for a particular pair of stocks, $x_{ij}^S = x_{ji}^S$, record the number of mutual fund portfolios in which stocks i and j coexist; an entry on the diagonal,

x^S_{ii}, counts portfolios to which stock i belongs (although, given that loops make little sense in the present context, I remove diagonal elements prior to making use of co-affiliation matrices in network measurements). Of course, it is also possible to define the $n_F \times n_F$ fund network co-affiliation matrix, $\mathbf{X}^F = \{x^F_{ij}\}$, with entries given by

$$x^F_{ij} = \sum_{k=1}^{n_S} a_{ki} a_{kj} \quad \Leftrightarrow \quad \mathbf{X}^F = \mathbf{A}^\top \mathbf{A}, \tag{1.3}$$

although, due to the nature of the hypotheses under investigation, I will focus almost exclusively on the relationships among stocks throughout this chapter.

1.2.2 Holdings Network Measures

A crucial step in the analysis of networks like those just described is the identification of the most important or "central" nodes in the network. We often regard the central actors in a social network, to take just one example, as being more visible to others, as exerting greater influence on the rest of the group, or as having greater access to resources (material or otherwise) flowing along ties in the network. Given the varied interpretations of what it means to be central, it should come as little surprise that network analysts have proposed dozens of methods to quantify a node's importance, ranging from degree centrality—a simple count of the edges incident upon each vertex, and perhaps the most basic measure of a node's network position—to complicated schemes based, for instance, on the theory of the statistical design of experiments (Stephenson and Zelen, 1989), or on the frequency with which a node falls on a random walk between other pairs of vertices (Newman, 2005).

The choice of centrality measure usually depends on the nature of the network process under investigation. Because I am primarily interested in measuring the importance of the stocks held by mutual fund portfolio managers, I will focus on the network of stocks represented by the co-affiliation matrix defined in Equation (1.2). In the projection of overlapping fund holdings onto the collection of equity securities, a connection between two stocks, $x_{ij}^S > 0$, indicates that one or more mutual funds holds stocks i and j simultaneously. Such a connection has a number of substantive implications for the stocks in question. For one thing, shocks to stock i—based on new information relevant to valuation, for example—will be immediately visible to some holders of stock j (namely, those who own both securities). This is just the sort of connection that Cohen and Frazzini (2008) have in mind when they examine mutual funds holding stock in a customer firm as well as its supplier, and find that supplier stocks react more quickly to customer shocks when held together in many funds' portfolios. A significant degree of overlap between stocks i and j also constitutes evidence of a revealed preference on the part of mutual fund managers to hold both stocks in the pair, suggesting a meaningful connection between those securities in the minds of investors. Thus, to the extent that a stock belongs to many pairs exhibiting unusually high overlap, we might reasonably conclude that the stock is both highly visible to investors, and quite prominent in the array of active associations investors make among potential equity investments.

As it turns out, a stock's breadth of ownership, as described by Chen, Hong, and Stein (2002), is identical to that stock's degree centrality in the bipartite graph defined in Equation (1.1), normalized by the cardinality of the network's second mode—the number of funds, in this context—as recommended by Borgatti and Everett (1997). Formally, using the notation introduced in Subsection 1.2.1,

$$\text{BREADTH}_i = \frac{1}{n_F} \sum_{j=1}^{n_S} a_{ij}. \tag{1.4}$$

As alluded to in the introduction, degree centrality, although convenient due to its conceptual and computational simplicity, is an inherently local measure of network structure, since it only captures information in the ties directly incident upon a given node. In the present context, BREADTH_i gives the number of funds that own stock i, but tells us nothing about the importance of those funds or, indeed, the prominence of the other stocks held in those portfolios. This obvious deficiency of degree as an indicator of nodal importance has motivated much work on the design of new techniques for assessing network position that more fully exploit the rich global architecture of many relational data sets.

In response to these considerations, Bonacich (1972a) introduces a measure of nodal importance that weights the contributions to a node's centrality score by the importance of that node's neighbors. This measure is commonly known as eigenvector centrality for reasons which will soon be apparent. Let $\mathbf{c} = \{c_1, c_2, \ldots, c_{n_S}\}$ represent an $n_S \times 1$ vector of centrality scores for each stock in the network, and define the scores according to a system of homogeneous linear equations represented by

$$c_i = x^S_{i1}c_1 + x^S_{i2}c_2 + \cdots + x^S_{in_S}c_{n_S}, \quad i = 1, \ldots, n_S. \tag{1.5}$$

Employing matrix notation and multiplying the left hand side by a constant yields

$$\lambda\mathbf{c} = \mathbf{X}^S\mathbf{c}, \tag{1.6}$$

and reduces the task of solving for the vector of centrality scores to that of identifying eigenvalues and eigenvectors associated with the stock network co-affiliation matrix, $\mathbf{X}^S$. Note that multiplying by λ preserves the relative values of centrality indices, and merely serves to ensure that the system has a nonzero solution. Indeed, Bonacich (1972a) demonstrates that this problem always has a unique solution for which the eigenvalue, λ, and all centrality scores, c_i, are positive.

Intuitively, eigenvector centrality accounts for the fact that, in many contexts, ties to nodes that are themselves important confer greater status than connections to vertices on the fringes of a network. Applied to the network of equities, Equation (1.6) implies that a stock can be important because it exhibits high overlap with many other stocks, or as a result of ties to just a few extremely prominent firms. Given the interpretation of holdings network centrality as capturing something akin to visibility or importance from the perspective of mutual fund managers, this is an attractive feature. It means, for example, that a stock is more visible to the extent that it connects with other highly visible stocks. Indeed, Borgatti (2005) classifies a variety of centrality measures on the basis of the implicit assumptions each makes as to the "flow process" operating over network ties, and recommends eigenvector centrality for networks in which links transmit changes in attitudes or beliefs, lending further support for my use of this measure in the empirical tests in Section 1.4.[1]

Up to this point, we have considered computing eigenvector centrality scores directly from the stock network co-affiliation matrix given by Equation (1.2). While this matrix provides a raw indication of the extent to which each pair of stocks overlaps in the portfolios of mutual fund managers, its entries are heavily dependent on the number of portfolios to which each stock belongs (i.e., each stock's network degree). Logically, the higher a stock's degree, the greater must be its propensity for overlap. Since stock degree is largely a function of well-studied attributes such as firm size (a $200 billion public company is certain to have more shareholders than a $50 million firm), eigenvector centrality calculated from the unadjusted stock co-affiliation matrix is likely to provide little information beyond that contained in the usual set of com-

[1]Specifically, Borgatti (2005) contends that influence processes logically allow for parallel duplication—i.e., simultaneous transmission across multiple links—and need not make use of shortest paths through the network, making eigenvector centrality, which satisfies both criteria, an obvious choice in such applications.

pany characteristics. A more interesting measure of holdings overlap accounts for the fact that expansive stocks—those with higher network degree—mechanically assume greater connectedness with the rest of the market, weighting a stock's connections according to the likelihood that those ties constitute meaningful relationships, and are not, for example, mere artifacts of the firm's market capitalization.

To control for degree when preparing affiliation data for network analysis, Borgatti and Halgin (2011) recommend a normalization first proposed by Bonacich (1972b) for use in the study of group membership ties modeled as a bipartite graph. For each pair of stocks, I tabulate the absolute number of overlapping and disjoint memberships and non-memberships in mutual fund portfolios, as follows:

		Stock j Member	Stock j Non-Member
Stock i	Member	n_{11}	n_{12}
	Non-Member	n_{21}	n_{22}

I then define the standardized stock network co-affiliation matrix, $\widetilde{\mathbf{X}}^S$, with entries:

$$\widetilde{x}_{ij}^S = \begin{cases} \ln(1+0.5) & \text{if } n_{11}n_{22} = n_{12}n_{21}, \\ \ln(1+(n_{11}n_{22} - \sqrt{n_{11}n_{22}n_{12}n_{21}})/(n_{11}n_{22} - n_{12}n_{21})) & \text{otherwise,} \end{cases} \tag{1.7}$$

where the transformation of matrix elements by taking natural logs serves to mitigate the impact of extreme observations in the holdings data. As a result of this normalization, we may roughly interpret a given entry, $\widetilde{x}_{ij}^S$, as expressing the extent to which the observed overlap between stocks i and j exceeds that expected by chance, given each stock's degree. Borgatti and Halgin (2011) suggest that such an adjustment is most appropriate when the reason for studying affiliation data is to infer hidden relationships from network ties, since the end result is a measure of each pair's "tendency or preference...to co-occur while controlling for nuisance variables"—in this case, the individual stocks' degrees.

1.3 Empirical Setup

1.3.1 Data

Data on mutual fund portfolio holdings, the principal building blocks used to construct the networks described in the previous section, come from Thomson Reuters. These holdings reports (often cited in the literature as the CDA/Spectrum data set) constitute quarterly disclosures of long equity positions by virtually all U.S.-based mutual funds, from March 1980 through December 2009. The primary source of Thomson's data is the SEC N-30D filing, which regulators require domestic mutual funds to publicly disclose on a semiannual basis (although, as Wermers (1999) demonstrates, a majority of mutual funds elect to file these reports at the end of every quarter). Thomson supplements these filings, as necessary, with data from fund prospectuses and, occasionally, with information obtained by direct communication with mutual fund management companies. For stocks held by managers in the Thomson Reuters holdings reports, I gather data related to trading activity and performance, including volume, prices, and returns, from the Center for Research in Security Prices (CRSP) Daily and Monthly Stock Files.

As a practical matter, in order to define a network of mutual funds and stocks at a single point in time, one must first synchronize holdings reports, which do not necessarily occur on the same day for all funds in the sample. Given that a majority of managers report holdings at a quarterly frequency, it seems most natural to update the network at the end of each calendar quarter. Accordingly, within a given quarter, I use the most recent filing provided by each fund, and take the approach (as is common in the mutual fund literature, beginning with Wermers, 1999) of assuming that funds pursue a buy-and-hold strategy, and carrying all holdings reported within the quarter forward to the end of the period. It is worth noting that, while the relative infrequency with which mutual funds report holdings raises concerns for studies using such disclosures as the basis for inferences about fund behavior and performance (for a thorough assessment, see Elton et al., 2010), because my hypotheses are principally concerned with the predictability of stock returns, any noise introduced by the low temporal resolution of portfolio holdings—missing a round-trip trade that occurred within the quarter, for example—will only make it more difficult for subsequent tests to produce statistically significant results and, importantly, should not interfere with the interpretation of my findings.

Because the hypotheses in Section 1.1 apply most directly to trading by active mutual fund managers, and since data on stock prices and characteristics only cover domestic securities, it is desirable to restrict attention to funds trading primarily in U.S. equities. Nevertheless, given the potential sensitivity of my network measures to missing data, I take a rather conservative approach to excluding funds from analysis. I begin by obtaining investment objective classification codes for the funds in my sample from the CRSP Survivor-Bias-Free U.S. Mutual Fund Database, which I match with Thomson holdings reports using the Wharton Research Data Services MFLINKS tables. The set of investment objective classes and the criteria for separating funds by strategy type have changed considerably over the thirty years covered by my data. As such, I evaluate a fund's objective by Lipper class code, when available, then by Lipper objective code, Strategic Insights objective code, and Wiesberger objective code; if I find none of these for a given fund, I rely on the CRSP investment policy code. I eliminate managers with investment objectives indicating a focus on anything but U.S. stocks (such as "Tax-Free Money Market" and "Japanese Funds"), but retain domestic equity funds that tilt toward a particular sector or investment style (for example, "Telecommunications Funds" and "Small Capitalization Growth").

Elton, Gruber, and Blake (2001) and Evans (2010) present evidence that mutual fund management companies occasionally field a number of new funds, and only report the past performance of those that experience early success. This practice gives rise to an "incubation bias," which results in overestimation of the returns of funds included in research databases and, implicitly, the performance of stocks held by such managers. To mitigate this bias, following Kacperczyk, Sialm, and Zheng (2008), I remove fund-quarter observations with performance reports that precede the fund's inception date, or for which the fund's name is missing from the database. Application of these filters, along with the investment objective-based exclusions described above, leaves 156,038 unique fund-quarter observations.

In the analysis that follows, it will also be important to make a distinction between active and passive mutual fund managers. If all funds operated passively, as index funds, then the structure of the holdings network would be uninteresting. Passive equity managers do not exercise discretion when choosing stocks, and do not adjust holdings quickly in response to new information or as a result of the actions of other fund managers. If such managers dominated the sample, all that one might infer from ownership ties, one could more easily learn by reading the index funds' guidelines for portfolio constitution. It follows that in constructing the holdings network, my primary interest is in the patterns of ownership traced out by the purchases and sales of active mutual fund managers. Moreover, to the extent that the presence of passive managers' holdings serves to obscure substantive ties among mutual funds—those connections that create opportunities for the sort of inter-manager influence and propagation of information or sentiment that motivate the hypotheses at the heart of this study—there is good reason to exclude index funds from the sample.

Unfortunately, neither the CRSP mutual fund files nor the Thomson holdings data consistently flag passively managed funds. In an attempt to systematically differentiate between active and passive funds, I follow the approach of Gil-Bazo and Ruiz-Verdú (2009), and Jiang, Verbeek, and Wang (2011), who use mutual fund names, along with a list of keywords likely associated with passive managers, to identify suspected indexers. I use a modified set of keywords, flagging as passive managers all CRSP funds with names containing one or more of the following strings: `Index`, `INDEX`, `index`, `Indx`, `Idx`, `Ix`, `NASDAQ`, `Nasdaq`, `Mkt`, `MKT`, `Dow`, `DJ`, `500`, `BARRA`, `Barra`, `iShare`, `MSCI`, `ProShare`. Using index fund identifiers provided by CRSP for the years 2008 and 2009, I find that my algorithm correctly classifies 70.5% of index funds in the test sample (716 of 1,015 index fund-quarter observations), with only 2.5% of flagged funds revealed as active managers (18 of 734 suspected index fund-quarter observations), confirming the usefulness of this procedure for minimizing the impact of index funds on network structure, while largely avoiding spurious deletion of active managers. In the full sample, application of this filter results in the exclusion of around 3% of all observations (4,877 of 156,038 fund-quarters), leaving a total of 151,161 fund-quarter observations.

Table 1.1 provides an overview of the universe of funds and stocks included in my analysis, constructed from Thomson Reuters holdings reports, as described above, before dropping suspected index funds. To highlight changes in the composition of the stocks and funds in the sample over the thirty years spanned by my data, I report average statistics for each five-year sub-period, from March 1980 to December 2009. Immediately apparent from the figures in Panel A is the striking expansion of the mutual fund industry over the course of the last three decades. A dramatic increase in the number of funds and equity assets under mutual fund management culminates during the stock market boom of the late 1990s, before leveling off in the volatile period surrounding the financial crisis in the last segment of the data. Perhaps unsurprisingly, one observes a corresponding—though somewhat less pronounced—increase in the size of the typical manager's portfolio, with the median fund holding around 70 equity positions, worth approximately $200 million, by the end of the sample. The growing popularity of index funds is also apparent in the data, with the number of suspected indexers rising steadily, beginning in the early 1990s.

Given the aforementioned importance of obtaining a complete representation of the connections among stocks for the validity of tests relating network structure to stock performance, it is natural to wonder whether the sample of mutual fund holdings picks up a substantial fraction of securities in the universe of U.S. equities. Panel B of Table 1.1 reports the sample's coverage of CRSP stocks, both in terms of the absolute percentage of CRSP stocks held by funds included in my analysis, and as measured by the fraction of CRSP market capitalization corresponding to stocks in the sample. The latter measure exceeds 90% in every sub-period, suggesting that the data indeed cover a vast majority of U.S. stocks, and that exceptions are quite small, and thus unlikely to present problems for the network measures employed in this study.

1.3.2 Network Construction

Based on the network model and methods described in Section 1.2, and using data on institutional investors' portfolio holdings outlined in the last subsection, it is finally possible to construct an empirical network representation of the links among stocks based on patterns in overlapping mutual fund equity positions. The process is straightforward: At the end of each quarter, from March 1980 to December 2009, I apply Equation (1.1) to portfolios defined by the most recent set of holdings reports, yielding the holdings network affiliation matrix encoding connections between funds and stocks; I then project these ties onto the space of stocks according to Equation (1.2), resulting in a stock network co-affiliation matrix, in which links among securities reflect relationships induced by common ownership.

Figure 1.1 presents a visualization of the network of US stocks based on holdings from December 2009. For the sake of clarity, I plot only the top five percent of connections based on the extent of raw overlap in ownership, and only display the largest connected component (removing, for example, stocks with no meaningful connections). The figure depicts a total of 1,508 stocks and 83,079 connections among them, plotted using Tulip, a software package for visualizing relational data. I accomplish the arrangement of nodes by the force-directed Fast Multipole Multilevel Method of Hachul and Jünger (2004), and bundle edges according to the algorithm of Lambert, Bourqui, and Auber (2010), making it easier to discern the global structure of links among stocks. Darker edges in the graph indicate nodes with higher network degree—that is, those that exhibit the most holdings overlap.

While the figure serves nicely as an illustration of the abstract network structures described in previous sections, one naturally wonders whether it also reveals something interesting about the relational configuration of stocks in the portfolios of mutual fund managers. The scale of the market and presumable complexity of the processes underlying the formation of stock-to-stock links make it difficult to draw deep conclusions from the simple depiction in Figure 1.1. Nevertheless, nodes in the concentrated mass at the top left quadrant of the network tend to be larger than those at the bottom right (median market capitalizations of around $750 million and $470 million, respectively) and are significantly more likely to belong to the S&P 500 index, suggesting that segmentation of the mutual fund industry (as reflected, for example, in a fund's choice of benchmark) might be responsible for the separation apparent in the graph. Figure 1.2, which presents similar visualizations at the end of each five-year sub-period in the sample, tends to support this interpretation, as the transition from one cluster of stocks to several dense subgroups seems to occur over a period in which diversity in mutual fund investment styles is known to have proliferated.

In addition to visual representations like those in Figures 1.1 and 1.2, many structural features of empirical networks are amenable to statistical analysis. Table 1.2 presents a variety of stock network statistics, averaged across quarters in each five-year sub-period of my sample. The number of nodes (stocks, in this case) increases steadily through the mid-2000s, reflecting the expansion of U.S. equity markets over the last thirty years. In the language of graph theory, a "connected component" is defined as a subgroup of nodes with the property that a path exists between any pair of nodes within the group, but no path exists from a node within the group to one without. An important limitation of eigenvector centrality, the principal measure used in tests throughout the remainder of this chapter, is that it only assigns scores to nodes in the largest connected component. As such, if the network of stocks frequently comprised several large, disconnected groups, my use of eigenvector centrality in Section 1.4 would present problems. Fortunately, based on the size of the largest connected component, which uniformly spans over 95% of stocks, this concern appears unfounded.[2] Indeed, the size of the next-largest component is almost always one, indicating that stocks falling outside of the largest component are almost always "isolates"—stocks with no connections to the rest of the network—and likely represent equities held only by index funds, the connections of which do not factor in the construction of the stock network.

[2]In fact, as Easley and Kleinberg (2010) observe, it is extremely difficult for two "giant components" to coexist, since adding just a single edge from one component to the other renders them connected; Newman (2010) provides a nice discussion of the theoretical conditions under which a giant component is likely to emerge.

Table 1.2 includes a summary of the distributions associated with standardized stock network degree, defined as the number of other stocks to which a given stock connects, divided by $n_S - 1$ to control for network size. Graph density represents the number of ties observed in the network as a proportion of those that could exist, if the graph were maximally connected, and is equivalent to the average standardized nodal degree. Together, these statistics reflect the fact that as portfolios have grown in size (see Table 1.1), the amount of overlap has mechanically increased. One implication of this trend is that, to the extent that it results in nodes becoming more similar with respect to connectivity, many network measures become less informative (one can imagine the extreme case in which all possible ties exist and nodes are, from the perspective of a network analyst, completely identical). Because eigenvector centrality operates on a weighted matrix of network ties, however, it is less susceptible to such effects, since similarity in the sheer number of outgoing ties need not imply that two nodes occupy similar positions within the network.

1.4 Asset Pricing Tests

1.4.1 Summary Statistics

Before delving into asset pricing tests of the hypotheses established in Section 1.1, it is worth taking a closer look at the basic statistical properties of holdings network centrality, along with a range of relevant stock characteristics, many of which show up as controls in the regressions featured throughout this section. To this end, Table 1.3 provides summary statistics, correlations, and cross-autocorrelations associated with all major variables employed in my analysis. As mentioned before, in the interest of obtaining a relatively complete representation of the relational structure of mutual fund holdings, my construction of the stock networks outlined in Subsection 1.3.2 utilizes all equities held by mutual funds in the sample. Consistent with past work on mutual fund holdings, however, I restrict asset pricing tests—including the summary statistics presented in this subsection—to common stocks (i.e., the set of securities with CRSP share codes 10 or 11, which excludes, for example, closed-end mutual funds, real estate investment trusts, and ADRs) and to stocks with prices of at least \$5 per share as of the beginning of the quarter (in an attempt to reduce the impact of outlier returns).

CENTRALITY represents a stock's eigenvector centrality score with respect to the normalized stock network defined by Equation (1.7), standardized to take a value between zero and one, and calculated for each stock-quarter observation in my sample, from March 1980 to December 2009. To the extent that a connection between two stocks suggests that members of the pair are meaningfully related in the eyes of mutual fund managers—in the sense that the observed degree of overlap in holdings is unlikely to have occurred by chance, given the number of ties associated with each stock—we may interpret centrality as a rough proxy for a stock's "visibility" or "popularity" within the network of equity holdings. ΔCENTRALITY is simply the change in a stock's centrality since the previous quarter, calculated for stocks included in the sample at the beginning and end of the last quarter, starting in June 1980. From Panel C of Table 1.3, it is clear that levels of centrality are highly autocorrelated, providing support for the decision to also consider changes in centrality in tests of the hypotheses outlined in the introduction. In the same panel, the negative correlation between CENTRALITY_t, the centrality score at time t, and $\Delta\text{CENTRALITY}_{t+1}$, the change in centrality over the next quarter, results from the fact that a stock beginning the quarter with a high centrality score has more to lose, so to speak, than a non-central stock. Likewise, the positive correlation between $\Delta\text{CENTRALITY}_t$ and

CENTRALITY_{t+1} simply reflects the fact that stocks gaining in centrality during the last quarter will have higher centrality scores, all else equal, in the current period.

As discussed in the introduction, two constructs from the previous literature on mutual fund holdings are conceptually related to my measure of holdings network centrality, and these will play an important role in the tests that follow. BREADTH and ΔBREADTH are, respectively, the levels of and changes in the fraction of mutual funds holding a given stock, as described by Chen, Hong, and Stein (2002). Likewise, MFO and ΔMFO represent the levels of and changes in the fraction of a firm's shares held by mutual funds, following Chen, Jegadeesh, and Wermers (2000). As with changes in centrality, observation of ΔBREADTH and ΔMFO begins in June 1980. Statistics on BREADTH and MFO presented in Table 1.3 correspond closely to those reported in Chen, Hong, and Stein (2002). Interestingly, despite the theoretical link between centrality scores, breadth, and mutual fund ownership, Panel B reveals only modest positive correlations among ΔCENTRALITY, ΔBREADTH, and ΔMFO.

In addition to measures of mutual fund holdings, Table 1.3 includes a summary of variables that one might expected to correlate with stock centrality or that have been demonstrated in the literature to explain cross-sectional variation in stock returns. I define ANALYSTS as the fraction of sell-side equity analysts in the I/B/E/S database providing full-year EPS forecasts for a stock as of the end of the previous year, beginning in 1984, when prior-year data are first available. LOGSIZE is a stock's log market capitalization at the end of the quarter, based on prices and shares outstanding from CRSP. The book-to-market ratio, BM, is computed as the book value of shareholders' equity at the end of the previous fiscal year (Compustat item SEQ), divided by the firm's most recent market value, taken from CRSP. The earnings-to-price ratio, EP, is equal to a firm's net income before extraordinary items from the previous fiscal year (Compustat item IB), over the firm's most recent market value.

Because negative valuation ratios are difficult to interpret, I only report statistics on BM and EP for firms with non-negative book value and earnings, respectively. As a measure of momentum, I calculate LAST1YR as a stock's cumulative return over the last 12 months, based on data from CRSP, for stocks with at least one year of past returns. Finally, to proxy for liquidity, I define a stock's TURNOVER as its average monthly trading volume, measured over the prior quarter, as a proportion of shares outstanding; I following the advice of Nagel (2005), and divide volume for NASDAQ stocks by two to correct for the double-counting of trades on that exchange.

Contemporaneous and lagged correlations between centrality scores (both in terms of levels and changes) and the characteristics described above are, for the most part, unremarkable, although I will provide commentary on a number of subtle connections between centrality and stock characteristics hinted at in Panels B and C—particularly the relationship between my measure of centrality and firm size—in the context of the portfolio sorts presented in the next subsection.

1.4.2 Portfolio Sorts

I hypothesized in the introduction that, consistent with the model of Chen, Hong, and Stein (2002) and the authors' corresponding "change in breadth of ownership" measure, stocks with significant recent declines in popularity—as reflected in the pattern of connections among mutual fund managers' portfolio holdings—should underperform stocks experiencing an increase in network status, as the rising dispersion of investors' beliefs signaled by the former, combined with binding short-sale constraints, should result in the overvaluation of low-centrality shares. As a first step in the evaluation of that claim, this subsection reports the composition and performance of portfolios formed by sorting stocks on the basis of holdings network centrality. I begin with an analysis of sorts conducted using levels of centrality, then proceed to the more interesting case of portfolios sorted on changes in centrality, which accounts for the general persistence of network structure revealed by the dynamics of stock position discussed in the last subsection, and accords with the intuition, emphasized by Chen, Jegadeesh, and Wermers (2000), that changes in holdings indicate stronger shifts in opinion on the part of mutual fund managers than passively maintained portfolio allocations.

At the beginning of each quarter, from March 1980 to December 2009, I construct a network of stocks based on the overlap in mutual fund holdings, then sort firms according to levels of centrality, calculated as described in Subsection 1.2.2 and scaled to take values between zero and one. I partition the universe of stocks into ten equally weighted portfolios, where Portfolio 1 consists of the least central stocks in the network, and Portfolio 10, the most central stocks, and assess the monthly performance of stocks in each partition over the quarter subsequent to portfolio formation. Table 1.4 presents statistics based on the monthly excess returns of stocks in each decile portfolio, calculated as the difference in monthly holding period returns, taken from CRSP, and risk-free bond returns, obtained from Professor Kenneth French's website, via WRDS. In addition to providing the annualized mean and standard deviation of returns over the thirty years in my sample, I report the annualized alpha captured by each portfolio, based on a series of nested factor models attempting to explain an asset's returns as a linear combination of exposures to commonly cited sources of aggregate risk. These include: the basic Capital Asset Pricing Model (CAPM); a 3-factor model, incorporating the size and value factors suggested by Fama and French (1993); a 4-factor model, adding the Carhart (1997) momentum factor; and a 5-factor model that includes the Pástor and Stambaugh (2003) liquidity factor. In each case, alpha is simply the intercept taken from a regression estimate of the model in question, using factor returns from French's website as inputs. In this and all subsequent asset pricing tests, I employ the method of Newey-West (1987) to correct standard errors for heteroskedasticity and autocorrelation in returns.

The pattern of decile portfolio returns in Table 1.4 is somewhat disappointing in light of the hypothesized relationship between holdings network centrality and future stock returns. We do not observe the expected pattern of increasing returns from the bottom portfolio to the top. After adjusting for factors including market, size, value, momentum, and liquidity, it is stocks with the *highest* centrality that yield the most negative alpha; the risk-adjusted performance of stocks in the first decile portfolio is negative, but statistically insignificant. Indeed, the biggest winners—with over 2% annualized alpha—are stocks exhibiting middle-of-the-road levels of centrality. Comparing CAPM alpha estimates for the bottom and top decile portfolios to those obtained from the 3-factor model, we find that adding the size and value factors results in a sharp decline in risk-adjusted performance of both low- and high-centrality stocks, suggesting that an examination of stock characteristics associated with each portfolio might lead to greater insight as to the source of the performance described above.

In Table 1.5, I provide a range of stock characteristics for each decile portfolio, calculated as the average value, over all quarters in the sample, of the characteristic in question for stocks in the corresponding portfolio. Most variables should be familiar from Table 1.3. Mutual fund holders is simply the number of mutual funds with a long position in a stock at the end of each quarter. Market beta is taken from the regression of a stock's monthly excess returns against the Fama-French (1993) market, size, and value factors, using up to 60 months of data for stocks with a minimum of 24 months' trailing excess returns. A given stock's residual volatility is the annualized error variance from this regression. To ease the interpretation of the size, book-to-market, earnings-to-price, and momentum characteristics, I express these measures as the average decile portfolio to which stocks in each column belong, based on a sort of all CRSP stocks with respect to the characteristic in question.

Reviewing the figures in Table 1.5, one quickly concludes that there is something quite different about stocks in the first decile portfolio, and those in the second. In the case of firm size, for example, firms with the lowest levels of centrality are quite small, with an average size decile of 5.7, while those having only slightly higher centrality are among the largest fifth of listed U.S. stocks. Excluding the first decile, the relationship between centrality and size is more regular, with market capitalization gradually declining as one moves up the ladder with respect to holdings network centrality. Trading volume, the number of mutual fund holders, and the degree of analyst coverage follow a similar, pronounced pattern.

As it turns out, the correspondence between centrality and firm size is not surprising: it is an artifact of the process by which I normalize the stock-to-stock co-affiliation matrix, which penalizes firms with many shareholders—including, naturally, most large companies—to account for the fact that such stocks will inevitably exhibit more overlap in the portfolios of institutional investors. This normalization makes it relatively more difficult for large firms to feature in the class of extraordinary connections occupying the center of the holdings network, so that in addition to median-size stocks with weak connectivity, the lower decile bins contain large-cap stocks with patterns of connections indistinguishable from those expected by chance, given the number of portfolios with allocations to such firms. Based on the correlations between firm size and other stock characteristics, presented in Table 1.3, it seems reasonable to conclude that it is in fact the relationship between levels of centrality and market capitalization that gives rise to the clear differences across decile portfolios in mutual fund ownership, stock breadth, the attention of sell-side analysts, and levels of turnover.

These observations, together with evidence of the marked persistence in levels of centrality, suggest that examining portfolios sorted on innovations in, rather than absolute measures of, a stock's network position might yield more interesting results. Tables 1.6 and 1.7 summarize the monthly performance of stocks sorted into ten portfolios on the basis of quarterly changes in centrality, calculated from June 1980 to December 2009, along with statistics related to a long/short trading strategy that calls for buying stocks experiencing the largest increase in holdings network centrality over the prior quarter, and selling stocks that declined most in centrality over the same period. Table 1.8 provides corresponding stock characteristics for each portfolio, averaged across quarterly observations, over the full sample.

In Table 1.6, we observe a pattern in returns that is much more consistent with the predictions of Section 1.1. Both raw and risk-adjusted returns are essentially flat across portfolios two through ten. By contrast, the bottom decile portfolio, consisting of stocks experiencing the greatest fall in importance over the preceding quarter, returns sharply lower performance, on average, with a highly significant 5-factor alpha of -3.3% per annum; this poor performance is relatively insensitive to the model one uses to adjust for risk. Indeed, from 1980 to 2009, an investor long the top decile and short the bottom decile would have earned annualized alpha in excess of 4%. That profits to this long/short strategy are asymmetrically concentrated on the short side of the portfolio supports the notion that underreaction to an increase in pessimism among mutual fund managers—a result, in the Chen, Hong, and Stein (2002) model, of risk aversion on the part of those rational arbitrageurs who *are* able to short—is actually driving my results.

Table 1.7 reports performance associated with the same portfolios, disaggregated by calendar month. The last column, which provides average monthly returns to the P10–P1 strategy, demonstrates that my findings do not depend on exceptional performance in a handful of months, ruling out explanations related to the well-documented January effect, reviewed by Thaler (1987), or due to window dressing on the part of fund managers, as described by Lakonishok et al. (1991). Nor are returns concentrated in months immediately following quarterly portfolio disclosures, allaying any concern that my results stem from a reaction by other investors to information contained in the holdings reports, themselves. Point estimates of long/short portfolio returns are fairly stable and positive in ten of twelve months, although the small number of observations in each month precludes serious statistical inference.

In Figure 1.3, I plot cumulative abnormal returns to the P10–P1 strategy in each year of the sample, calculating abnormal returns in each month by subtracting the portion of returns explained by variables in the 5-factor regression from Table 1.6, including market, size, value, momentum, and liquidity factors. Because only partial-year data on changes in centrality are available at the beginning and end of the sample, I measure performance in 1980 from July through December, and in 2010, from January through March. While the strategy performs extremely well at times, it is obvious from the plot that alpha is consistently positive over the three decades spanned by my data, and that the results do not stem from anomalous returns in just a few years of the sample.

The characteristics associated with portfolios sorted on changes in centrality, found in Table 1.8, further highlight the benefits of shifting the focus to innovations in stock position. First, note that firm size changes very little across decile portfolios; one no longer observes the mechanical relationship between market capitalization and levels of network centrality evident in Table 1.5. As expected, changes in local measures of stock importance, like BREADTH and MFO, exhibit moderate correlation with my global measure, ΔCENTRALITY. Finally, the U-shaped pattern in monthly turnover simply reflects the fact that trading activity necessarily underlies any meaningful change in network structure, whether an increase or decrease in centrality.

As a final step in my analysis of network-based portfolio sorts, Table 1.9 presents probabilities associated with quarter-to-quarter transitions of stocks across decile portfolios formed according to CENTRALITY (Panel A) and ΔCENTRALITY (Panel B). In each case, the category labeled "Out" corresponds to stocks that were not in the sample at the time of portfolio formation (for example, stocks removed as a result of falling below the $5-per-share threshold for inclusion). Referring to results in Panel A, the strong persistence in levels of network centrality is readily apparent in the high weight associated with transitions on or near the diagonal. I also observe that when a stock enters the sample, it is most likely to appear at low centrality. Likewise, firms on the fringe of the holdings network at time t are slightly more likely to drop from the sample at time $t+1$ than companies with high status. Turning to Panel B, which summarizes the intertemporal transit of stocks across portfolios formed on the basis of changes in centrality, two patterns emerge. First, the relatively high weights at the corners of the matrix suggest that stocks with big moves in centrality—positive or negative—are more likely to experience another large change in the next period, be it a significant continuation or a sharp reversal. Second, judging by the concentration of probability in transitions at the center of the matrix, stocks with no meaningful change in network position are somewhat less likely to feature in changes to network structure, going forward.

1.4.3 Factor Model Regressions

In the last subsection, I presented strong evidence that the returns on stocks showing the greatest increase in centrality significantly exceed the returns on stocks witnessing the greatest decline in network status, consistent with the prediction of Hypothesis

1.1. Of course, if changes in network structure merely capture a propensity for mutual fund managers to overweight stocks with characteristics known to generate high returns, then one would expect to observe this pattern even under the null hypothesis.[3] To rule out an explanation based solely on mutual fund preferences for stock characteristics, Table 1.6 included risk-adjusted estimates of long/short performance based on regressions controlling for factors known to explain cross-sectional variation in stock returns. Table 1.10 presents the full results of these regressions.

Specification 1 is simply the CAPM, where MKT is the return on a value-weighted portfolio of NYSE, AMEX, and NASDAQ stocks, less the one-month Treasury bill rate. Fama and French (1993) recommend two additional factors, leading to Specification 2: HML is the average return on high book-to-market (value) stocks, minus the average return on low book-to-market (growth) stocks, and SMB is the average return on small stocks, minus the average return on big stocks. Specification 3 presents the model proposed by Carhart (1997), which includes UMD, defined as the average return on past winners, minus the average return of past losers, where performance of winners and losers is measured over the prior 12 months, skipping the last month's return. Finally, Specification 4 adds the Pástor and Stambaugh (2003) liquidity factor, LIQ, defined as the average return of stocks with high beta to liquidity, minus the average return of stocks with low beta to liquidity. I obtain factor returns from WRDS, where all but the liquidity factor are identical to data available on Kenneth French's website.

As discussed with reference to Table 1.6, the long/short strategy's alpha is statistically and economically significant in each specification, confirming that the relationship between changes in network structure and future stock returns is robust to

[3] Falkenstein (1996) finds that mutual funds appear to show a preference for small stocks and firms with high book-to-market ratios, while Grinblatt, Titman, and Wermers (1995) present evidence that funds target stocks with high past returns, consistent with a "momentum" strategy.

controlling for previously documented stock characteristics. Market beta is highly significant in each regression, but quite small. Given the care taken to insulate portfolio sorts from mechanical size effects, it is perhaps reassuring that SMB is statistically indistinguishable from zero across specifications. Interestingly, coefficients on the book-to-market and momentum factors are both negative, with the latter coming in significant at the 5% level. The negative UMD coefficient indicates that, holding other factors constant, the P10–P1 strategy mimics a portfolio long past losers and short past winners—a strategy known, historically, to produce *low* returns. These findings directly contradict the alternative explanation that changes in centrality proxy for mutual fund bias toward favorable stock characteristics. Finally, the fact that LIQ comes out as zero in Regression 4 suggests that the long/short strategy's exceptional performance is not a result of tilting toward illiquid stocks (consistent with the observation, in Table 1.8, that stocks in both the top- and bottom-decile portfolios exhibit relatively high turnover). Along with the low adjusted R^2 statistics corresponding to each regression, these results clearly strengthen the case for large decreases in holdings network centrality as an indicator of overvaluation.

1.4.4 Predictive Regressions

So far, I've only provided explicit evidence that the relationship between changes in centrality and future stock returns is consistent with the prediction of Hypothesis 1.1, according to which binding short-sale constraints and increased disagreement among a stock's investors serve as an indication of lower subsequent performance. This claim derives from the model of Chen, Hong, and Stein (2002), who perform a test of the same hypothesis using quarterly differences in the percentage of mutual funds holding a stock as a measure of investor disagreement, finding empirical support for the model's predictions. As explained in the introduction, my primary purpose in developing a network-based proxy for the evolving attitudes of investors was to demonstrate that the global structure of equity ownership has something important to tell us, beyond the information about a stock's direct links revealed by a local measure like ΔBREADTH. In this subsection, using methodology similar to that introduced by Fama and MacBeth (1973), I add changes in breadth of ownership as a control in order to test whether, as proposed in Hypothesis 1.2, ΔBREADTH reduces but fails to eliminate the predictive power of ΔCENTRALITY.

Before running predictive regressions with returns as the dependent variable, I perform a set of tests to identify the determinants of changes in centrality from among a list of stock characteristics: some expected to correlate with my measure, others known from the literature to predict future stock returns. First, for each quarter with available data, I perform a cross-sectional regression of changes in centrality, represented by ΔCENTRALITY, against contemporaneous values for a range of explanatory variables introduced and defined in Subsection 1.4.1. Next, in an approach akin to that of Fama and MacBeth (1973), I calculate a time-series average of quarterly estimates for each coefficient in the cross-sectional regressions, which I report in Table 1.11. The average adjusted R^2 listed under a given specification is simply the time-series average of quarterly adjusted R^2 statistics from the corresponding set of cross-sectional regressions, and the number of observations at the bottom of each column indicates the number of quarterly coefficients used to compute time-series means associated with each specification.

Regressions 1 and 2 in Table 1.11 provide yet another view of the highly significant positive correlation between my measure of shifts in stock prominence and other authors' proxies for changes in a stock's popularity with mutual fund managers, ΔBREADTH and ΔMFO. Regression 3, which includes both alternative measures, suggests that my centrality score is more closely related to breadth of ownership than to MFO, although the low R^2 coefficient in each of the first three specifications makes it clear that neither alternative measure explains much of the change in a stock's holdings network centrality. Specifications 4 and 5 include additional controls, although the limited availability of data on sell-side analyst coverage—which turns out not to correlate with changes in centrality, in any case—restricts estimation of Regression 4 to a shorter time series, from March 1984 to December 2009. Turning to the last column of Table 1.11, and consistent with the pattern in stock characteristic across decile portfolios in Table 1.8, I find that log market capitalization has no explanatory power, while turnover correlates positively with ΔCENTRALITY.

The negative coefficient on BM_t and significant positive coefficient on EP_t mirror those found in similar regressions of ΔBREADTH on stock characteristics reported by Chen, Hong, and Stein (2002). Interestingly, while Table 1.11 reports a strong inverse relationship between a stock's past performance and its current change in centrality, Chen, Hong, and Stein (2002) find exactly the opposite: changes in breadth of ownership exhibit a heavy *positive* loading on trailing stock performance. The authors interpret this as evidence that momentum profits must derive, at least in part, from binding short-sale constraints, as suggested by Hong, Lim, and Stein (2000). Alternatively, it is possible that this relationship between changes in breadth of ownership and momentum arises as a result of the previously cited tendency for mutual funds to hold past winners in preference to past losers. Under this interpretation, because changes in a stock's eigenvector centrality do not depend on its being dumped by a herd of funds (unlike changes in breadth of ownership, which can *only* come about as a result of sales by many holders), ΔCENTRALITY need not exibit the same correlation as ΔBREADTH to a stock's past performance, although the significant negative coefficient on $\mathrm{LAST1YR}_t$ in Table 1.11 is still somewhat puzzling.

As a final step in the evaluation of Hypothesis 1.2, Tables 1.12–1.15 present an evaluation of the ability of changes in holdings network centrality to predict the future performance of individual stocks. Unlike the tests in previous subsections, the regressions that follow control for changes in the fraction of mutual funds holding each stock (ΔBREADTH), as well as changes in the fraction of each stock's shares held by mutual funds (ΔMFO), along with many of the stock characteristics featured in previous tables. Again, I proceed by running quarterly cross-sectional regressions, then average coefficients in the time series to obtain the numbers in each table. Now, however, the dependent variable in the cross-sectional regressions is the future return of each stock, and changes in centrality feature on the right-hand side. Finally, in keeping with the portfolio sorts presented throughout the rest of the chapter, and because extreme changes in centrality appear to be most informative (I estimate excess kurtosis of 19 for ΔCENTRALITY across stock-quarter observations in my sample), all subsequent tests express changes in stock prominence as a pair of binary indicator variables: dummy variables equal to one when a stock belongs to the portfolios of equities facing the biggest decrease and increase in centrality, respectively.

In Table 1.12, the dependent variable in quarterly cross-sectional regressions, R_{t+1}, represents a stock's cumulative return from the end of quarter t to the end of quarter $t+1$. $\mathrm{P1}_t$ is a dummy variable equal to one if a stock is in the bottom decile of firms sorted on changes in centrality, and $\mathrm{P10}_t$ is a dummy variable equal to one if a stock is in the top decile of stocks sorted on changes in centrality. The first column simply reiterates that stocks with the greatest drop in network status significantly underperform those with the greatest increase in importance, and that the effect is primarily a result of poor future returns to stocks in the former class. The second regression suggests that, by itself, $\Delta\mathrm{BREADTH}_t$ fails to predict future returns, contradicting the results in Chen, Hong, and Stein (2002).[4] Perhaps the most important specification is Regression 3, which includes the breadth of ownership measure alongside my holdings network centrality dummy variables. Note that both measures are significant at the 5% level, and bear the appropriate signs. While the value of holdings network centrality as a predictor is slightly attenuated, it remains both economically and statistically significant. Based on the average adjusted R^2, including both local and global measures of stock importance explains roughly twice as much of the variation in stock returns than using either forecasting variable in isolation. Finally, adding changes in mutual fund ownership, as defined by Chen, Jegadeesh, and Wermers (2000), fails to drive out either $\Delta\text{BREADTH}_t$ or P1_t.

The last two specifications in Table 1.12 incorporate additional controls to account for the size effect (Banz, 1981), the value effect (Basu, 1977), price momentum (Jegadeesh and Titman, 1993), the neglected firm effect (Arbel and Strebel, 1982), and the tendency for stocks with high turnover to experience lower future returns (Brennen, Chordia, and Subrahmanyam, 1998). With the exception of ANALYSTS_t, estimates for controls match results from the past literature on return predictability. The value of both holdings-based measures, ΔBREADTH and ΔCENTRALITY, is diminished, although each remains statistically and economically significant. To place my findings in perspective, based on the estimates in Regression 6 and holding everything else constant, stocks facing the sharpest drop in network centrality during the previous quarter will return 1.2% less, on average, over the next four quarters; likewise, a two-standard-deviation decrease in a stock's breadth of ownership corresponds

[4]There are a number of possible explanations for the discrepancy. My sample covers eleven additional years, from 1999 to 2009, and I make several exclusions that the previous authors did not (e.g., dropping stocks priced below $5/share, eliminating funds with unsuitable investment objectives). Moreover, Chen, Hong, and Stein (2002) only include stocks in size quintiles two through five in their baseline results—ΔBREADTH is unable to forecast returns for the smallest stocks—acknowledging that doing so gives rise to a selection bias. Because I place no such restrictions on the size of firms in my sample, my estimates of each measure's success are relatively conservative.

to returns that are 1.5% lower, on average, over the next year. Overall, these results confirm that holdings structure has meaningful implications for stock performance, and support the claim in Hypothesis 1.2 that the importance of network connections among stocks and managers goes beyond local measures of ownership.

Noting that the distribution of changes in centrality is rather fat-tailed, Table 1.13 reports the performance of those stocks facing only the most extreme changes in network status, sorting stocks into twenty portfolios on the basis of changes in centrality, and replacing the dummy variables from Table 1.12 with P1_t and P20_t, which record membership in the bottom and top five percent of stocks, respectively. Consistent with the view that changes in centrality at the tails of the distribution are most informative, the effects under this specification are considerably more pronounced: I find that stocks in the bottom portfolio underperform by roughly 2.2% in the subsequent year, a result that is significant at the 5% level.

I conclude by evaluating the predictive performance of ΔCENTRALITY at alternative horizons, employing the same procedure and controls used in the last column of Table 1.12. Each column of Table 1.14 presents the time-series average of coefficients from quarterly cross-sectional regressions, in which the dependent variable is a stock's cumulative return over the τ months following portfolio formation, for horizons ranging from one month to one year. Table 1.15 has the same format, but applies to stocks sorted into twenty portfolios on the basis of changes in centrality. Both tables paint a similar picture: the value of information contained in ΔCENTRALITY is highest in the month immediately following portfolio formation, decaying significantly over the next eleven months. For example, the 10% of stocks with the biggest declines in centrality over a given quarter underperform by annualized rates of 2.4%, 1.5%, and 0.4%, over the next one month, six months, and twelve months, respectively. The effect is similar for more extreme changes in centrality: the 5% of stocks with the greatest drop in network status underperform by annualized rates of 4.1%, 2.3%, and 1.0%, over the next one month, six months, and twelve months. In the context of the model described in Section 1.1, these results are consistent with limits to arbitrage slackening over time, as the trading of rational investors not subject to short-sale constraints gradually incorporates the disagreement suggested by low ΔCENTRALITY and ΔBREADTH into stock prices. Incidentally, changes in breadth of ownership do not appear useful for predicting short-run stock returns, only attaining significance after three months subsequent to portfolio formation, and seem to display a similar—if less extreme—reduction in performance over the next nine months.

1.5 Conclusion

I have outlined a framework for viewing the equity portfolios of active mutual fund managers as a network in which stocks connect to one another through overlap in mutual fund holdings. Modeled as a graph, this network gives rise to a novel measure of stock importance, holdings network centrality, which exploits the global structure of stock ownership to gauge the prominence of each stock in the minds of investors. This relational perspective represents a significant break from the previous literature, which focuses on the attributes of individual firms viewed in isolation to come up with necessarily local measures of an equity's status. I apply my network methodology to a model of disagreement and constrained short sales proposed by Chen, Hong, and Stein (2002), and find that holdings network centrality improves upon the authors' breadth of ownership measure in the identification of overvalued stocks. Specifically, controlling for breadth of ownership and a range of stock characteristics known to predict future returns, a portfolio short those stocks experiencing the greatest decrease in centrality over the preceding quarter outperforms by roughly 2.2%, annually.

While my measure of stock importance based on mutual fund ownership ties is an interesting application of relational thinking to a specific problem in finance, it is important to note that my network model of investors' portfolio holdings is highly flexible and amenable to a range of applications beyond the calculation of stock centrality. For example, if one assumes that ties promote the transmission of information, this framework is ideal for studying lead-lag relationships in returns based on the flow of stock- and sector-specific news across the "community structure" clearly visible in the holdings networks depicted in Figure 1.2. Likewise, there is no reason to restrict oneself to the analysis of mutual fund portfolios; my methods are equally suited to the analysis of interlocking hedge fund portfolios, or to the holdings of retail investors like those studied by Grinblatt and Keloharju (2000). Given the vast social networks literature, and new data sets combining individual investors' stock holdings with personal information ranging from home address to place of employment, the holdings network model offers a new way to analyze financial decision making in a social context. Finally, the clear connection established in this chapter between the overall structure of institutional investor's portfolios and the cross-sectional performance of stocks embedded in that network suggests that theoretical models based on network processes are likely important for understanding the behavior of financial markets in an increasingly connected world.

1.6 References

Antón, M., & Polk, C. (in press). Connected stocks. *The Journal of Finance.*

Arbel, A., & Strebel, P. (2005). The neglected and small firm effects. *Financial Review, 17*(4), 201-218.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3-18.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance, 32*(3), 663-682.

Blocher, J. (2013, March). The Externalities of Crowded Trades. *Working Paper*, Vanderbilt University.

Bonacich, P. (1972a). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology, 2*(1), 113-120.

Bonacich, P. (1972b). Technique for analyzing overlapping memberships. *Sociological Methodology, 4*, 176-185.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55-71.

Borgatti, S. P., & Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks, 19*(3), 243-269.

Borgatti, S. P., & Halgin, D. S. (2011). Analyzing affiliation networks. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 417-433). London, England: SAGE Publications.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics, 49*(3), 345-373.

Carhart, M. M. (2012). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57-82.

Chen, J., Hong, H., & Stein, J. C. (2002). Breadth of ownership and stock returns. *Journal of Financial Economics, 66*(2), 171-205.

Chen, H. L., Jegadeesh, N., & Wermers, R. (2000). The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and Quantitative Analysis, 35*(3), 343-368.

Cohen, R. B., Coval, J. D., & Pástor, Ľ. (2005). Judging fund managers by the company they keep. *The Journal of Finance, 60*(3), 1057-1096.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance, 63*(4), 1977-2011.

Cohen, L., Frazzini, A., & Malloy, C. (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy, 116*(5), 951-979.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge, England: Cambridge University Press.

Elton, E. J., Gruber, M. J., & Blake, C. R. (2001). A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *The Journal of Finance, 56*(6), 2415-2430.

Elton, E. J., Gruber, M. J., Blake, C. R., Krasny, Y., & Ozelge, S. O. (2010). The effect of holdings data frequency on conclusions about mutual fund behavior. *Journal of Banking & Finance, 34*(5), 912-922.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance, 65*(4), 1581-1611.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance, 51*(1), 111-35.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607-636.

Gil-Bazo, J., & Ruiz-Verdú, P. (2009). The relation between price and performance in the mutual fund industry. *The Journal of Finance, 64*(5), 2153-2183.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor types: A study of Finland's unique data set. *Journal of Financial Economics, 55*(1), 43-67.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review, 85*(5), 1088-1105.

Hachul, S., & Jünger, M. (2005). Drawing large graphs with a potential-field-based multilevel algorithm. In Graph drawing (pp. 285-295). Berlin/Heidelberg: Springer.

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *The Journal of Finance, 62*(1), 251-301.

Hong, H., Kubik, J. D., & Stein, J. C. (2004). Social interaction and stock-market participation. *The Journal of Finance, 59*(1), 137-163.

Hong, H., Kubik, J. D., & Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *The Journal of Finance, 60*(6), 2801-2824.

Hong, H., Lim, T., & Stein, J. C. (2002). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance, 55*(1), 265-295.

Ivković, Z., & Weisbenner, S. (2007). Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices. *Review of Financial Studies, 20*(4), 1327-1357.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Jiang, H., Verbeek, M., & Wang, Y. (2011, December). Information content when mutual funds deviate from benchmarks. *Working Paper*, Erasmus University.

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. *Review of Financial Studies, 21*(6), 2379-2416.

Lakonishok, J., Shleifer, A., Thaler, R., & Vishny, R. (1991). Window dressing by pension fund managers. *The American Economic Review, 81*(2), 227-231.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance, 49*(5), 1541-1578.

Lambert, A., Bourqui, R., & Auber, D. (2010). Winding roads: Routing edges into bundles. In Computer graphics forum (Vol. 29, No. 3, pp. 853-862). Hoboken, NJ: Blackwell Publishing.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of Finance, 32*(4), 1151-68.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics, 78*(2), 277-309.

Newey, W., & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703-708.

Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks, 27*(1), 39-54.

Newman, M. E. J. (2010). *Networks: An introduction.* Oxford, England: Oxford University Press.

Ozsoylev, H., Walden, J., Yavuz, M. D., & Bildik, R. (2011, March). Investor networks in the stock market. *Working Paper*, University of Oxford.

Pástor, Ľ, & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*(3), 642-685.

Solis, R. (2009). Visualizing stock-mutual fund relationships through social network analysis. *Global Journal of Finance and Banking Issues, 3*(3), 8-22.

Stephenson, K., & Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks, 11*(1), 1-37.

Thaler, R. H. (1987). Anomalies: The January effect. *The Journal of Economic Perspectives, 1*(1), 197-201.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge, England: Cambridge University Press.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *The Journal of Finance, 54*(2), 581-622.

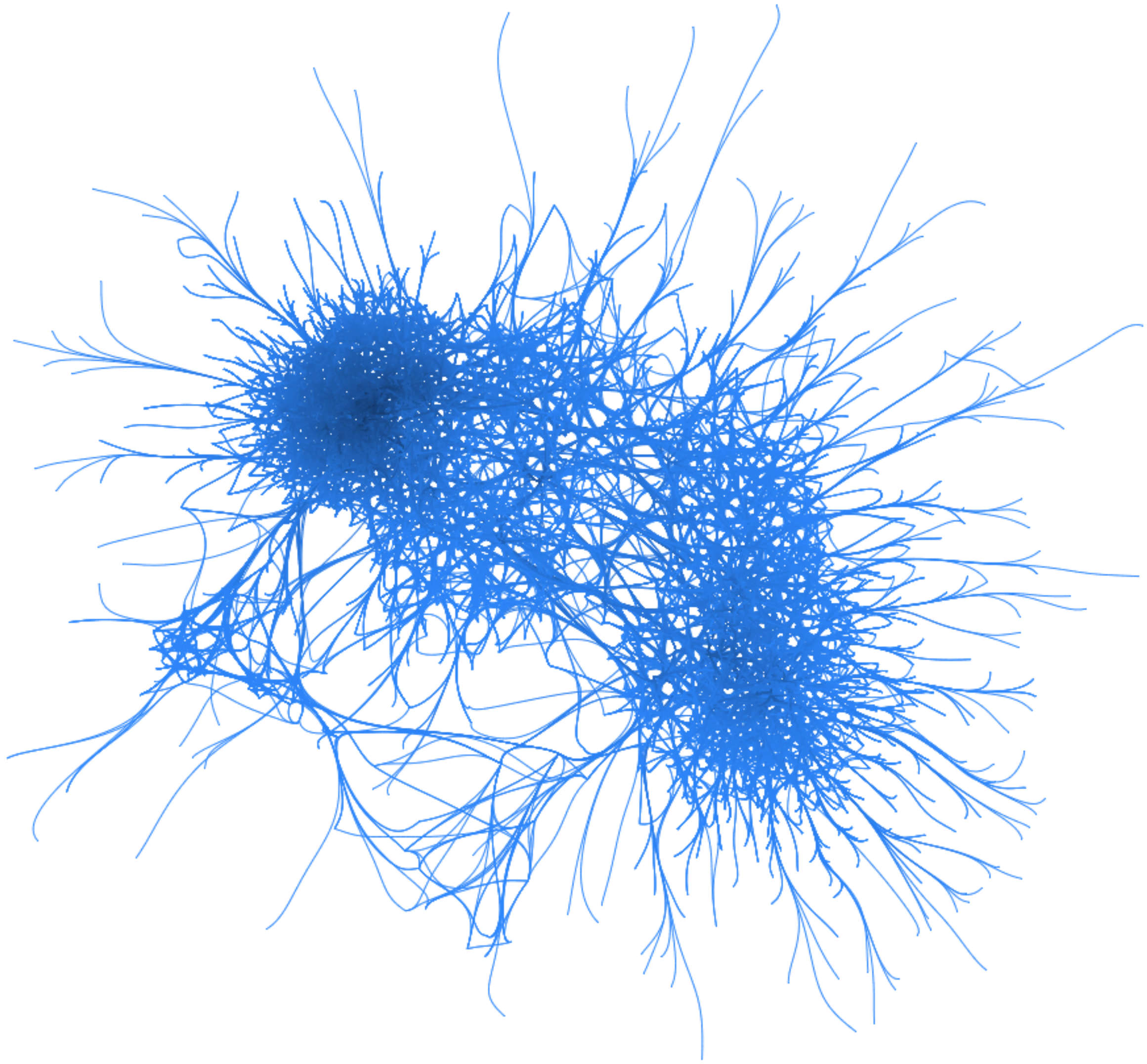

Figure 1.1. Network of Mutual Fund Stock Holdings, December 31, 2009. A visualization of the network of US stocks based on holdings from December 2009. For the sake of clarity, only the top five percent of connections based on the extent of raw overlap in ownership have been plotted, and only the largest connected component is shown. A total of 1,508 stocks and 83,079 connections are depicted. The arrangement of nodes is accomplished by the force-directed Fast Multipole Multilevel Method of Hachul and Jünger (2004), and edges have been bundled according to the algorithm of Lambert, Bourqui, and Auber (2010); both methods are implemented through Tulip, a software package for visualizing relational data. Darker edges indicate nodes with higher network degree. Stocks in the concentrated mass at the top left quadrant of the network tend to be larger than those at the bottom right—median market capitalizations of around \$750 million and \$470 million, respectively—and are significantly more likely to belong to the S&P 500 index.

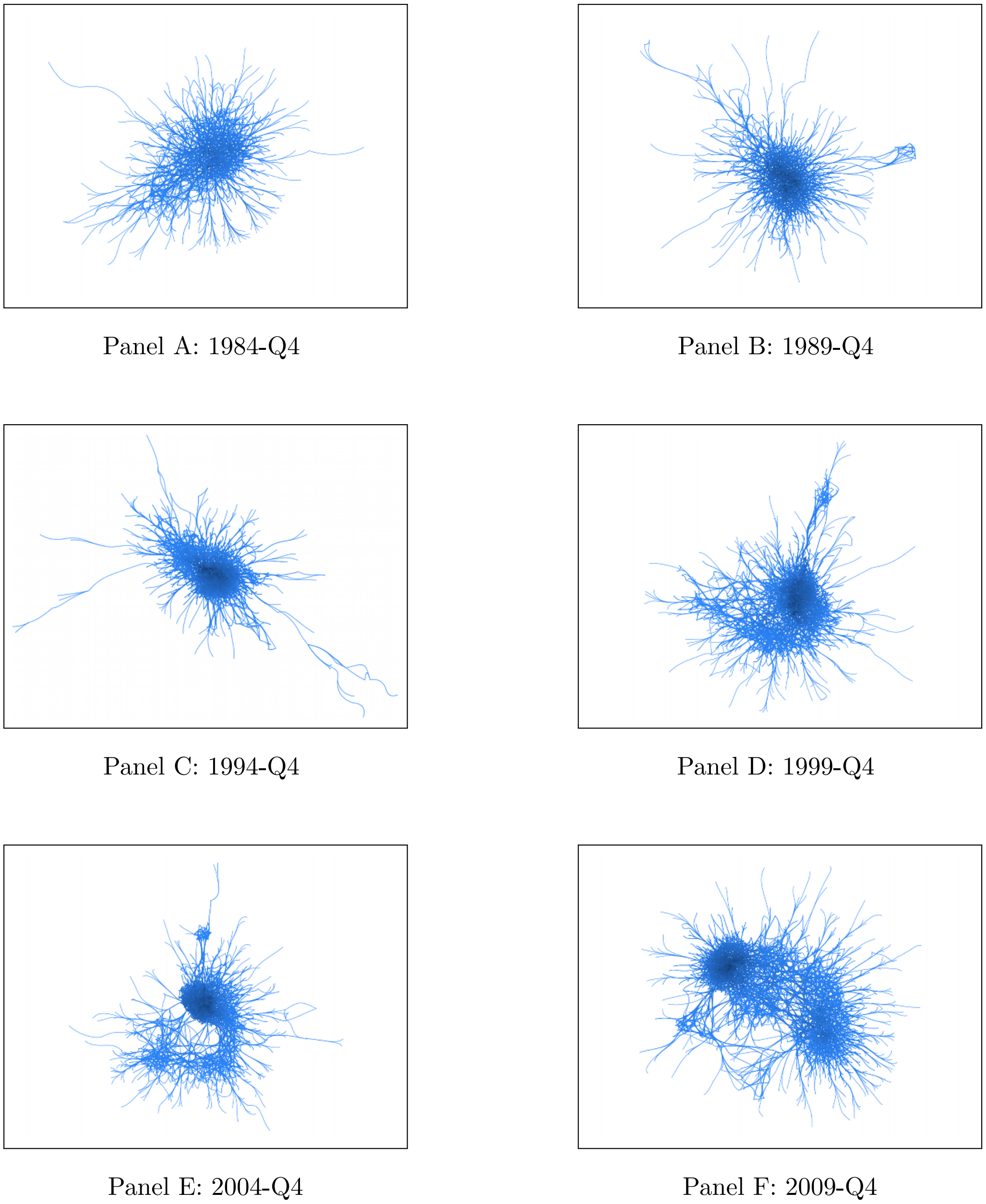

Figure 1.2. Network of Mutual Fund Stock Holdings, 1984–2009. A visualization of the network of US stocks based on holdings at the end of each 5-year sub-period in the sample, from 1984 to 2009. For the sake of clarity, only the top five percent of connections based on the extent of raw overlap in ownership have been plotted, and only the largest connected component is shown. Darker edges indicate nodes with higher network degree. The apparent transformation in network structure over the second half of the sample is likely a result of changes in the mutual fund industry.

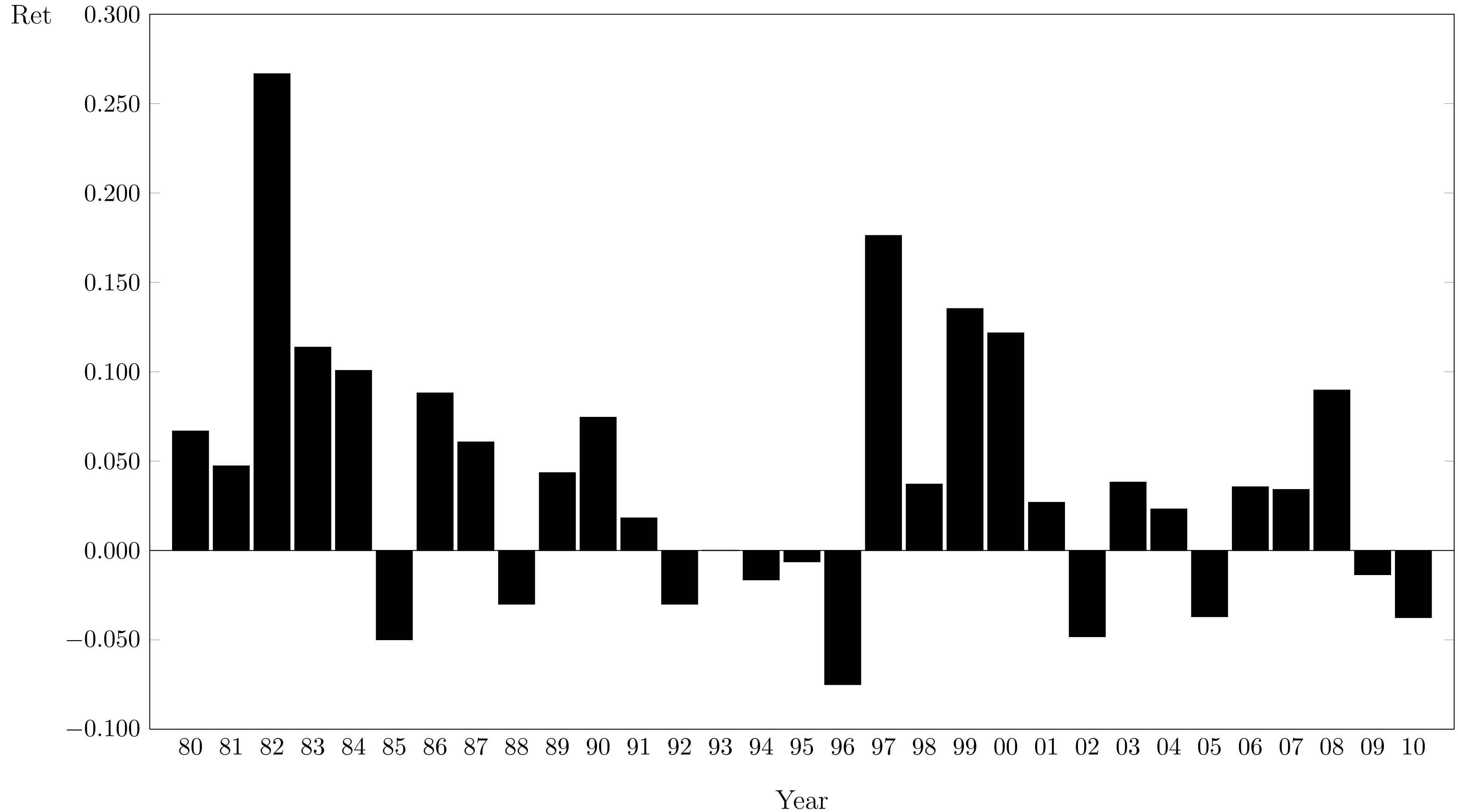

Figure 1.3. Annual Alpha, P10–P1 Strategy. Bars measure the annual cumulative abnormal returns to a strategy that is long the top decile of stocks sorted on changes in network centrality and short the bottom decile, as described in the text. Abnormal returns are calculated in each month by subtracting the portion of returns explained by factors from Regression 4 in Table 1.10, including MKT, HML, SMB, UMD, and LIQ. Performance in 1980 is measured from July through December, and in 2010, from January through March.

Table 1.1
Summary Statistics: Mutual Fund Holdings Data

At the end of each quarter, from March 1980 to December 2009, a network of stocks is constructed on the basis of overlapping mutual fund holdings, as described in the text. For each summary measure, figures represent the average of that measure across quarters in the corresponding sub-period. As this study focuses on the common stock holdings of mutual funds, statistics in the table reflect only the fraction of mutual fund portfolios invested in equities. The items in Panel A are self-explanatory. In Panel B, coverage of CRSP stocks is expressed in two forms: the raw percentage of CRSP stocks held by mutual funds in the sample, and the percentage of CRSP market capitalization corresponding to stocks in the sample.

Panel A: Funds	Sub-period					
	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09
Count of funds in the database	309.9	483.4	789.1	1,583.3	2,306.9	2,174.7
Suspected index funds	1.0	3.6	20.9	44.5	77.2	89.7
Total equity assets ($billion)	43.6	121.5	287.8	1,307.2	2,079.0	2,507.1
Equity portfolio size						
Average ($million)	138.8	248.5	352.5	794.9	909.5	1,147.5
Median ($million)	48.8	74.8	87.5	133.4	140.1	198.1
Min ($thousand)	345.0	328.0	381.5	63.0	9.5	23.0
Max ($billion)	1.7	6.9	18.6	59.1	72.3	108.6
Number of positions						
Average	54.1	67.6	91.2	111.1	127.4	146.6
Median	42.8	48.7	56.9	66.9	69.8	72.9
Min	3.2	2.9	3.2	2.1	1.0	1.1
Max	547.8	1,388.0	1,740.0	3,348.6	3,676.0	3,547.3

Panel B: Stocks	Sub-period					
	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09
Count of stocks in the database	2,685.8	3,773.5	4,545.6	6,883.8	6,101.1	5,473.0
Coverage of CRSP stocks (%)						
By number	48.1	55.9	62.9	79.0	83.4	80.5
By dollar value	91.2	95.3	96.8	98.1	98.4	97.5
Number of mutual fund holders						
Average	6.3	8.7	15.8	25.4	48.9	57.9
Median	3.0	3.9	7.2	10.2	22.6	33.9
Min	1.0	1.0	1.0	1.0	1.0	1.0
Max	147.6	191.6	272.1	498.2	847.1	708.2

Table 1.2
Summary Statistics: Holdings Network, Graph-Level Measures

At the end of each quarter, from March 1980 to December 2009, a network of stocks is constructed on the basis of overlapping mutual fund holdings, as described in the text. For each summary measure, figures represent the average of that measure across quarters in the corresponding sub-period. Number of nodes is the size of the group of stocks used to construct the network. A component is a maximally connected subgroup of stocks (i.e., a path exists between each pair of stocks within the subgroup, but no path exists to a stock outside of the subgroup); this total includes trivial components of size one, enumerated separately as the number of isolates in the graph. Nodal degree is the number of stocks to which a given stock has a direct tie, divided by $n_S - 1$ to control for network size. Graph density—the number of ties observed as a proportion of those that could exist, if the graph were maximally connected—is equivalent to the average standardized nodal degree. Finally, clustering is calculated as the number of complete triples in a graph (sets of three stocks with links between all three pairs), divided by the number of triangles in the graph (distinct sets of three stocks), and provides a coarse measure of the extent to which stocks cluster in the portfolios of mutual fund managers.

Stock Network	Sub-period					
	1980-84	1985-89	1990-94	1995-99	2000-04	2005-09
Number of nodes	2,685.8	3,773.5	4,545.6	6,883.8	6,101.1	5,473.0
Number of components	9.0	37.0	58.5	112.4	253.3	156.9
Size of largest component						
Number of nodes	2,677.8	3,737.5	4,488.1	6,772.5	5,848.8	5,317.0
Percentage of nodes	99.7	99.1	98.7	98.4	95.9	97.2
Size of next-largest component	1.0	1.0	1.0	1.0	1.0	1.2
Number of isolates	8.0	36.0	57.5	111.4	252.3	155.8
Nodal degree, standardized						
Quintile 1	4.0	7.7	17.0	16.2	23.2	25.3
Quintile 2	8.6	19.7	32.6	39.5	51.2	62.2
Quintile 3	15.0	32.6	41.3	52.6	60.0	69.6
Quintile 4	22.6	41.5	49.5	61.5	71.4	76.3
Quintile 5	61.4	69.6	77.4	82.8	84.8	84.3
St. Dev.	10.4	17.0	17.6	22.6	24.7	26.8
Density	13.8	25.9	34.4	41.6	49.3	55.0
Clustering	52.1	73.2	68.8	78.5	81.0	85.3

Table 1.3
Summary Statistics and Correlations for Assorted Stock Characteristics

Panel A reports summary statistics for a number of stock characteristics appearing throughout the chapter, calculated across all stock-quarter observations used in subsequent asset pricing tests, covering the period from March 1980 to December 2009. CENTRALITY is a stock's eigenvector centrality score with respect to the network defined by overlapping mutual fund holdings, as described in the text. ΔCENTRALITY records the change in a stock's centrality since the previous quarter, beginning in June 1980. BREADTH and ΔBREADTH are, respectively, the levels of and changes in the fraction of mutual funds holding a stock, as described by Chen, Hong, and Stein (2002). Likewise, MFO and ΔMFO represent the levels of and changes in the fraction of a stock's shares held by mutual funds, following Chen, Jegadeesh, and Wermers (2000). As with centrality scores, observation of ΔBREADTH and ΔMFO begins in June 1980. ANALYSTS is the fraction of I/B/E/S analysts providing full-year EPS forecasts for a stock as of the end of the previous year, beginning in March 1984, when prior-year data are first available, expressed as a percentage. LOGSIZE is a stock's log market capitalization, based on price and shares outstanding from CRSP. The book-to-market ratio, BM, is computed as book value (Compustat item SEQ) at the end of the previous fiscal year, divided by the firm's most recent market value. The earnings-to-price ratio, EP, is equal to net income (Compustat item IB) from the previous fiscal year, over the firm's most recent market value. LAST1YR is a stock's cumulative return, over the last 12 months, expressed as a percentage, based on data from CRSP. Finally, TURNOVER is a stock's average monthly volume as a proportion of shares outstanding, measured over the prior quarter, also expressed as a percentage. Panel B provides contemporaneous correlations associated with each pair of characteristics, while Panel C presents cross-autocorrelations for each pair, based on a lag of one quarter. The last row of Panel A reports the number of stock-quarter observations used to calculate summary statistics and correlations associated with each characteristic.

Panel A: Summary statistics

	CENT	ΔCENT	BDTH	ΔBDTH	MFO	ΔMFO	ANLYS	LOGSZ	BM	EP	LST1YR	TURN
Mean	0.5573	0.0123	2.5127	0.0603	9.7725	0.3244	0.6820	5.6936	0.6578	0.0753	25.2981	8.0288
Median	0.6191	0.0007	1.3328	0.0010	6.7900	0.0500	0.2916	5.5074	0.5196	0.0611	12.2641	4.8262
St. Dev.	0.3428	0.1282	3.2847	0.5401	9.2796	2.6730	1.0356	1.7454	0.7245	0.0870	81.3395	11.5157
Obs.	426,077	397,908	426,077	424,211	426,077	424,211	388,956	422,300	385,997	327,342	392,963	415,832

Table 1.3—*Continued*

Panel B: Contemporaneous correlations

	$CENT_t$	$\Delta CENT_t$	$BDTH_t$	$\Delta BDTH_t$	MFO_t	ΔMFO_t	$ANLYS_t$	$LOGSZ_t$	BM_t	EP_t	$LST1YR_t$	$TURN_t$
$CENT_t$		0.1709	−0.0507	0.0131	0.2819	−0.0273	−0.1658	−0.0719	−0.0008	−0.0496	−0.0093	0.0883
$\Delta CENT_t$			−0.0370	0.1403	−0.0137	0.0702	−0.0390	−0.0385	−0.0084	−0.0061	−0.0032	0.0003
$BDTH_t$				0.0861	0.3801	−0.0098	0.6573	0.7871	−0.1370	−0.0788	−0.0228	0.2217
$\Delta BDTH_t$					0.0527	0.3462	−0.0371	0.0538	−0.0673	−0.0609	0.1837	0.0480
MFO_t						0.1383	0.1843	0.3937	−0.0985	−0.1029	−0.0254	0.3366
ΔMFO_t							−0.0559	−0.0131	−0.0361	−0.0233	0.1060	0.0200
$ANLYS_t$								0.5898	−0.0609	−0.0314	−0.0668	0.1772
$LOGSZ_t$									−0.2177	−0.1195	0.0268	0.2625
BM_t										0.5861	−0.1829	−0.0973
EP_t											−0.1572	−0.0459
$LST1YR_t$												0.1620
$TURN_t$												

Panel C: Autocorrelations and cross-autocorrelations

	$CENT_{t+1}$	$\Delta CENT_{t+1}$	$BDTH_{t+1}$	$\Delta BDTH_{t+1}$	MFO_{t+1}	ΔMFO_{t+1}	$ANLYS_{t+1}$	$LOGSZ_{t+1}$	BM_{t+1}	EP_{t+1}	$LST1YR_{t+1}$	$TURN_{t+1}$
$CENT_t$	0.9284	−0.2078	−0.0547	−0.0054	0.2675	−0.0307	−0.1788	−0.0768	0.0146	−0.0341	−0.0084	0.1152
$\Delta CENT_t$	0.1456	−0.0642	−0.0359	0.0042	−0.0095	0.0110	−0.0391	−0.0388	−0.0066	−0.0045	0.0034	0.0002
$BDTH_t$	−0.0850	−0.0580	0.9847	−0.0608	0.3659	−0.0296	0.6586	0.7808	−0.1220	−0.0682	−0.0276	0.2585
$\Delta BDTH_t$	0.0089	−0.0085	0.1083	0.1291	0.0754	0.0747	−0.0366	0.0610	−0.0705	−0.0555	0.1805	0.0549
MFO_t	0.2684	−0.0330	0.3736	−0.0181	0.9504	−0.1299	0.1826	0.3879	−0.0759	−0.0852	−0.0345	0.3963
ΔMFO_t	−0.0204	0.0094	0.0044	0.0711	0.1453	0.0180	−0.0555	−0.0061	−0.0423	−0.0274	0.1087	0.0266
$ANLYS_t$	−0.1973	−0.0417	0.6488	−0.0298	0.1722	−0.0280	0.9738	0.5858	−0.0575	−0.0329	−0.0477	0.2038
$LOGSZ_t$	−0.1020	−0.0389	0.7830	0.0011	0.3901	0.0014	0.5914	0.9887	−0.1915	−0.0996	0.0095	0.2997
BM_t	−0.0085	−0.0070	−0.1405	−0.0238	−0.1052	−0.0239	−0.0639	−0.2130	0.8747	0.4974	−0.1352	−0.1075
EP_t	−0.0555	−0.0051	−0.0820	−0.0212	−0.1063	−0.0151	−0.0295	−0.1169	0.5081	0.8425	−0.1199	−0.0515
$LST1YR_t$	−0.0077	−0.0022	−0.0086	0.0973	−0.0102	0.0602	−0.0724	0.0290	−0.1768	−0.1261	0.6765	0.1504
$TURN_t$	0.0938	0.0038	0.2147	−0.0186	0.3255	−0.0253	0.1868	0.2481	−0.0658	−0.0258	0.1064	0.7832

Table 1.4
Portfolios Sorted on Levels of Centrality: Stock Returns

The table decomposes the performance of stocks sorted on levels of centrality, as described in the text, using monthly returns from April 1980 to March 2010. Excess returns for each portfolio are calculated by subtracting the return on risk-free bonds. The CAPM alpha is simply the intercept from a regression of monthly excess returns on the CRSP value-weighted market portfolio. The 3-factor alpha is obtained by regressing monthly excess returns on the Fama-French (1993) market, size, and value factors; the 4-factor model includes, as an explanatory variable, the Carhart (1997) momentum factor; the 5-factor model adds the Pástor-Stambaugh (2003) liquidity factor. Returns, standard deviations, and regression coefficients have been annualized and are expressed as percentages. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation in returns. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Portfolio									
	1	2	3	4	5	6	7	8	9	10
Excess Ret.	6.33*	6.21*	7.95**	8.47**	9.18**	10.73***	10.44***	9.08**	7.77**	6.45*
	(1.72)	(1.65)	(2.13)	(2.20)	(2.41)	(2.79)	(2.75)	(2.42)	(2.00)	(1.66)
St. Dev.	17.69	18.46	18.62	19.37	19.09	19.36	18.91	18.70	19.12	19.43
CAPM α	−0.18	−1.13	0.29	0.52	1.54	3.24*	3.13*	1.96	0.49	−0.85
	(0.09)	(0.72)	(0.25)	(0.43)	(1.03)	(1.90)	(1.79)	(1.08)	(0.26)	(0.44)
3-Factor α	−1.63	−1.73	−0.57	−0.29	−0.27	1.68**	0.87	−0.37	−2.23**	−3.21***
	(1.29)	(1.65)	(0.59)	(0.37)	(0.27)	(1.98)	(0.92)	(0.41)	(2.17)	(2.90)
4-Factor α	−1.94	−1.78	0.91	0.74	0.98	2.09**	0.93	−0.27	−1.68	−2.59**
	(1.54)	(1.58)	(1.02)	(1.02)	(1.04)	(2.41)	(1.01)	(0.30)	(1.61)	(2.29)
5-Factor α	−1.81	−1.95*	0.91	0.70	1.03	2.37***	1.48	0.05	−1.41	−2.33**
	(1.37)	(1.66)	(0.98)	(0.93)	(1.10)	(2.66)	(1.62)	(0.05)	(1.34)	(1.99)
Centrality	0.078	0.214	0.287	0.377	0.508	0.613	0.727	0.800	0.856	0.913

Table 1.5
Portfolios Sorted on Levels of Centrality: Stock Characteristics

The table presents characteristics of stocks sorted on levels of centrality, as described in the text, over the period from March 1980 to December 2009. Specifically, each figure represents the average value, over all quarters in the sample, of the characteristic in question for stocks in the corresponding portfolio. Mutual fund holders is the number of funds in the sample with a long position in the stock. Chen, Hong, and Stein (2002) define Breadth as the fraction of mutual funds holding a stock at the end of the quarter. Chen, Jegadeesh, and Wermers (2000) define MFO as the fraction of a stock's shares owned by mutual funds at the end of the quarter. The fraction of analysts covering a stock is measured by the percentage of active I/B/E/S analysts providing full-year EPS forecasts for a firm as of the end of the last calendar year, expressed as a percentage (beginning in March 1984, when prior-year data are first available). Turnover is calculated as the average of the last three months' trading volume as a proportion of shares outstanding. Size is simply market price multiplied by shares outstanding, using data from CRSP. The book-to-market ratio is computed as book value (Compustat item SEQ) at the end of the previous fiscal year, divided by the firm's most recent market value. The earnings-to-price ratio is equal to net income (Compustat item IB) from the previous fiscal year, over the firm's most recent market value. Momentum is measured by the stock's average monthly return over the last 12 months, using CRSP data. Market beta is taken from regressions of the stocks' monthly excess returns on the Fama-French (1993) market, size, and value factors, using up to 60 months of data; residual volatility is the error variance from these regressions, annualized. Only stocks with at least 24 months of trailing excess returns are included in the regression analysis. To ease interpretation of the size, book-to-market, earnings-to-price, and momentum characteristics, these measures are expressed as the average decile portfolio to which stocks in each column belong.

Characteristic	Portfolio									
	1	2	3	4	5	6	7	8	9	10
Centrality score	0.078	0.214	0.287	0.377	0.508	0.613	0.727	0.800	0.856	0.913
Mutual fund holders	13.2	68.4	67.4	53.0	36.3	27.3	23.7	21.8	22.6	23.7
Breadth, %	0.8	3.7	4.1	3.6	2.9	2.5	2.1	1.8	1.8	2.2
MFO, %	3.1	8.5	11.0	11.9	11.2	10.3	9.7	9.1	8.9	9.1
Analysts, %	0.24	1.02	1.26	1.11	0.91	0.80	0.64	0.45	0.36	0.34
Turnover, % monthly	5.9	8.6	9.8	10.3	8.8	7.7	7.1	6.6	6.5	6.3
Size decile	5.7	8.0	8.5	8.3	7.6	7.1	6.7	6.4	6.4	6.6
Book-to-market decile	5.3	4.6	4.6	4.7	4.7	5.1	5.4	5.6	5.6	5.7
Earnings-to-price decile	5.6	5.5	5.3	5.3	5.3	5.4	5.5	5.6	5.6	5.7
Momentum decile	6.2	6.0	6.1	6.3	6.2	6.3	6.3	6.1	6.1	6.0
Market beta	0.8	1.0	1.1	1.1	1.0	1.0	1.0	1.0	1.0	1.1
Residual vol., % annual	38.9	29.6	30.7	33.1	36.3	37.2	38.4	39.0	39.3	38.9

Table 1.6
Portfolios Sorted on Changes in Centrality: Stock Returns

The table decomposes the performance of stocks sorted on changes in centrality, as described in the text, using monthly returns from July 1980 to March 2010. Excess returns for each portfolio are calculated by subtracting the return on risk-free bonds. The CAPM alpha is simply the intercept from a regression of monthly excess returns on the CRSP value-weighted market portfolio. The 3-factor alpha is obtained by regressing monthly excess returns on the Fama-French (1993) market, size, and value factors; the 4-factor model includes, as an explanatory variable, the Carhart (1997) momentum factor; the 5-factor model adds the Pástor-Stambaugh (2003) liquidity factor. Returns, standard deviations, and regression coefficients have been annualized and are expressed as percentages. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation in returns. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Portfolio										
	1	2	3	4	5	6	7	8	9	10	10–1
Excess Ret.	4.83	7.82**	8.00**	7.73**	8.53**	8.21**	7.76**	7.56**	8.04**	7.91*	3.08**
	(1.27)	(2.15)	(2.22)	(2.16)	(2.47)	(2.32)	(2.21)	(2.00)	(2.07)	(1.86)	(2.25)
St. Dev.	19.02	18.19	17.87	17.59	16.96	17.47	17.33	18.74	19.53	21.06	7.38
CAPM α	−2.35	0.99	1.37	1.20	2.28	1.72	1.25	0.51	0.62	−0.06	2.29*
	(1.49)	(0.65)	(0.85)	(0.78)	(1.47)	(1.10)	(0.90)	(0.33)	(0.41)	(0.04)	(1.73)
3-Factor α	−3.38***	−0.62	−0.82	−0.88	0.07	−0.48	−0.58	−1.19	−0.73	−0.81	2.56**
	(3.52)	(0.69)	(0.96)	(1.06)	(0.08)	(0.57)	(0.69)	(1.45)	(0.75)	(0.72)	(2.11)
4-Factor α	−3.30***	−0.40	−0.36	−0.45	0.68	−0.06	−0.06	−0.25	0.38	0.81	4.11***
	(3.33)	(0.43)	(0.44)	(0.52)	(0.80)	(0.07)	(0.07)	(0.31)	(0.41)	(0.75)	(3.12)
5-Factor α	−3.30***	−0.31	−0.19	−0.33	0.99	0.14	0.25	0.09	0.63	0.82	4.12***
	(3.18)	(0.33)	(0.23)	(0.37)	(1.13)	(0.17)	(0.29)	(0.11)	(0.67)	(0.78)	(3.06)
ΔCentrality	−0.134	−0.042	−0.023	−0.011	−0.002	0.005	0.013	0.029	0.058	0.219	-

Table 1.7

Portfolios Sorted on Changes in Centrality: Returns by Calendar Month

The table presents the average excess returns of stocks sorted on changes in centrality, as described in the text, reported by calendar month. Returns data cover the period from July 1980 to March 2010, with 29 observations for months April through June, and 30 observations for months July through March. Excess returns for each portfolio are calculated by subtracting the return on risk-free bonds. Returns are expressed as monthly percentages. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation in returns. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Portfolio										
	1	2	3	4	5	6	7	8	9	10	10–1
Jan.	1.21	1.54	1.65	1.63*	1.75*	1.57*	1.22	1.41	1.57	1.59	0.38
	(1.30)	(1.59)	(1.68)	(1.86)	(1.91)	(1.75)	(1.37)	(1.50)	(1.57)	(1.56)	(0.94)
Feb.	0.72	0.74	0.87	0.81	0.72	1.00	0.84	0.58	0.44	−0.04	−0.76
	(0.88)	(0.89)	(1.01)	(1.02)	(0.88)	(1.21)	(0.96)	(0.62)	(0.44)	(0.03)	(1.13)
Mar.	1.29*	1.60**	1.54**	1.36**	1.15*	1.31**	1.38**	1.16*	1.46**	1.45**	0.16
	(1.95)	(2.61)	(2.55)	(2.15)	(2.01)	(2.15)	(2.30)	(1.87)	(2.38)	(2.16)	(0.50)
Apr.	1.05	0.92	1.05	0.87	1.34	1.37	1.53*	1.38	1.54	1.56	0.51
	(1.44)	(1.23)	(1.31)	(1.13)	(1.57)	(1.52)	(1.80)	(1.45)	(1.54)	(1.52)	(1.15)
May	1.18*	1.39**	1.63***	1.92***	1.66**	1.46**	1.71***	1.79***	1.80***	1.82**	0.64*
	(1.85)	(2.39)	(2.84)	(3.14)	(2.58)	(2.49)	(3.01)	(2.99)	(2.82)	(2.63)	(2.04)
Jun.	0.05	0.56	0.28	0.35	0.20	−0.01	0.26	0.27	0.02	0.33	0.28
	(0.06)	(0.68)	(0.41)	(0.51)	(0.32)	(0.02)	(0.39)	(0.35)	(0.02)	(0.41)	(1.28)
Jul.	−0.97	−0.54	−0.67	−0.73	−0.37	−0.16	−0.44	−0.52	−0.51	−0.79	0.19
	(1.15)	(0.67)	(0.86)	(0.93)	(0.50)	(0.19)	(0.58)	(0.60)	(0.66)	(0.96)	(0.65)
Aug.	0.27	0.51	0.61	0.25	0.43	0.36	0.24	0.21	0.42	0.15	−0.11
	(0.25)	(0.50)	(0.66)	(0.27)	(0.46)	(0.40)	(0.27)	(0.22)	(0.40)	(0.14)	(0.38)
Sep.	−1.45	−1.31	−1.01	−0.92	−0.69	−0.82	−0.80	−0.97	−1.13	−1.11	0.34
	(1.30)	(1.26)	(1.03)	(0.92)	(0.80)	(0.90)	(0.85)	(0.95)	(1.10)	(1.01)	(1.00)
Oct.	−1.13	−0.99	−0.99	−0.83	−0.69	−0.75	−0.93	−1.01	−1.04	−0.73	0.40
	(0.78)	(0.69)	(0.69)	(0.57)	(0.49)	(0.53)	(0.66)	(0.67)	(0.70)	(0.44)	(1.03)
Nov.	0.97	1.20	1.17	1.23	1.35	1.15	1.22	1.47	1.69	1.54	0.56
	(0.90)	(1.19)	(1.19)	(1.23)	(1.42)	(1.18)	(1.25)	(1.38)	(1.43)	(1.28)	(1.25)
Dec.	1.68**	2.23***	1.91***	1.85***	1.71***	1.75***	1.59***	1.85**	1.84**	2.19***	0.51
	(2.61)	(3.43)	(3.09)	(3.44)	(3.09)	(3.02)	(2.95)	(2.71)	(2.61)	(2.93)	(1.37)
All	0.40	0.65**	0.67**	0.64**	0.71**	0.68**	0.65**	0.63**	0.67**	0.66*	0.26**
	(1.27)	(2.15)	(2.22)	(2.16)	(2.47)	(2.32)	(2.21)	(2.00)	(2.07)	(1.86)	(2.26)

Table 1.8
Portfolios Sorted on Changes in Centrality: Stock Characteristics

The table presents characteristics of stocks sorted on changes in centrality, as described in the text, over the period from June 1980 to December 2009. Specifically, each figure represents the average value, over all quarters in the sample, of the characteristic in question for stocks in the corresponding portfolio. Mutual fund holders is the number of funds in the sample with a long position in the stock. Chen, Hong, and Stein (2002) define ΔBreadth as the change in the fraction of mutual funds holding a stock from the beginning to the end of the quarter. Chen, Jegadeesh, and Wermers (2000) define ΔMFO as the change in the fraction of a stock's shares owned by mutual funds from the beginning to the end of the quarter. The fraction of analysts covering a stock is measured by the percentage of active I/B/E/S analysts providing full-year EPS forecasts for a firm as of the end of the last calendar year, expressed as a percentage (beginning in March 1984, when prior-year data are first available). Turnover is calculated as the average of the last three months' trading volume as a proportion of shares outstanding. Size is simply market price multiplied by shares outstanding, using data from CRSP. The book-to-market ratio is computed as book value (Compustat item SEQ) at the end of the previous fiscal year, divided by the firm's most recent market value. The earnings-to-price ratio is equal to net income (Compustat item IB) from the previous fiscal year, over the firm's most recent market value. Momentum is measured by the stock's average monthly return over the last 12 months, using CRSP data. Market beta is taken from regressions of the stocks' monthly excess returns on the Fama-French (1993) market, size, and value factors, using up to 60 months of data; residual volatility is the error variance from these regressions, annualized. Only stocks with at least 24 months of trailing excess returns are included in the regression analysis. To ease interpretation of the size, book-to-market, earnings-to-price, and momentum characteristics, these measures are expressed as the average decile portfolio to which stocks in each column belong.

Characteristic	Portfolio									
	1	2	3	4	5	6	7	8	9	10
ΔCentrality	−0.134	−0.042	−0.023	−0.011	−0.002	0.005	0.013	0.029	0.058	0.219
Mutual fund holders	41.7	41.6	37.2	33.4	32.5	34.1	36.5	41.9	41.8	34.8
ΔBreadth, %	−0.14	−0.02	0.00	0.01	0.02	0.03	0.05	0.08	0.12	0.27
ΔMFO, %	−0.11	0.13	0.11	0.15	0.16	0.16	0.18	0.22	0.34	0.77
Analysts, %	0.84	0.80	0.78	0.67	0.62	0.68	0.75	0.84	0.81	0.71
Turnover, % monthly	9.4	8.0	7.1	6.6	6.3	6.7	7.0	7.9	8.6	8.8
Size decile	7.7	7.4	7.2	6.9	6.8	6.9	7.1	7.4	7.5	7.4
Book-to-market decile	4.9	5.1	5.3	5.3	5.3	5.3	5.2	5.1	5.1	4.9
Earnings-to-price decile	5.4	5.5	5.6	5.6	5.6	5.6	5.5	5.5	5.4	5.3
Momentum decile	6.2	6.1	6.1	6.1	6.1	6.1	6.2	6.2	6.2	6.1
Market beta	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1
Residual vol., % annual	36.3	35.5	34.9	35.2	35.1	35.0	35.4	35.2	35.7	36.7

Table 1.9
Centrality-Based Portfolio Sorts: Transition Matrices

For portfolios sorted on levels of centrality (Panel A, using data from March 1980 to December 2009) and changes in centrality (Panel B, using data from June 1980 to December 2009), transition matrices are constructed by computing the probability, in each quarter, of a stock moving between each pair of portfolios, then averaging these probabilities across time. Each entry in the table represents the average probability of a stock moving from the row portfolio to the column portfolio, expressed as a percentage. "Out" corresponds to stocks that were not in the sample at the time of portfolio formation (e.g., stocks removed for falling below the $5/share threshold for inclusion in the sample).

Panel A: Levels		Portfolio$_{t+1}$										
		1	2	3	4	5	6	7	8	9	10	Out
	1	65.8	10.9	3.0	2.1	1.6	0.9	1.0	0.8	0.7	0.3	12.9
	2	9.4	61.0	15.0	4.1	1.9	1.0	0.9	0.7	0.5	0.3	5.2
	3	1.8	15.1	56.5	14.6	3.8	1.8	0.9	0.7	0.6	0.3	3.9
	4	1.0	2.6	16.2	56.0	13.6	3.3	1.5	1.0	0.7	0.6	3.4
	5	0.5	1.3	2.8	15.7	53.1	14.3	4.0	1.9	1.3	1.2	3.9
Portfolio$_t$	6	0.3	0.5	1.4	2.7	16.3	49.4	15.4	5.5	2.6	1.5	4.3
	7	0.2	0.3	0.5	1.2	3.3	17.6	46.6	16.3	6.1	2.4	5.5
	8	0.4	0.3	0.4	0.5	1.7	4.9	18.3	44.9	17.7	4.6	6.2
	9	0.3	0.4	0.4	0.4	1.1	2.3	5.1	19.0	48.4	15.7	7.0
	10	0.2	0.2	0.2	0.3	0.7	1.3	2.1	4.4	15.2	69.1	6.3
	Out	6.1	2.8	1.7	1.1	1.0	0.9	1.3	1.3	1.8	1.2	80.9

Panel B: Changes		Portfolio$_{t+1}$										
		1	2	3	4	5	6	7	8	9	10	Out
	1	11.0	9.7	8.0	8.1	7.2	7.5	7.8	8.9	10.6	15.2	5.9
	2	11.6	12.1	9.7	8.2	8.0	8.6	8.8	9.2	10.7	8.4	4.8
	3	9.7	10.1	10.6	9.9	9.8	9.2	10.2	9.8	9.0	7.2	4.7
	4	7.7	8.3	10.3	10.7	11.5	11.0	10.3	10.0	7.7	7.2	5.3
	5	6.7	8.2	9.3	10.8	11.9	11.8	11.5	9.4	7.7	7.2	5.5
Portfolio$_t$	6	7.2	8.3	9.8	10.6	11.7	11.5	10.8	9.0	7.9	7.6	5.4
	7	7.8	9.1	10.8	11.6	10.3	10.6	10.2	9.3	7.8	7.5	4.9
	8	9.3	10.6	10.8	10.0	9.2	9.4	9.2	10.2	9.4	7.9	4.1
	9	11.1	11.0	8.9	8.4	8.6	8.3	8.7	10.4	11.0	9.8	4.0
	10	14.5	9.5	8.4	7.0	6.5	7.0	7.1	9.6	12.5	13.7	4.2
	Out	0.8	0.8	0.9	1.2	1.4	1.4	1.4	1.1	1.4	1.8	87.8

Table 1.10
Portfolios Sorted on Changes in Centrality: Factor Model Regressions

The table presents the results of time-series regressions of monthly returns to a strategy long firms in the top decile of stocks sorted on changes in centrality, and short firms in the bottom decile of stocks sorted on changes in centrality, along with a number of factors known to explain cross-sectional variation in equity returns, using data from July 1980 to March 2010. Fama and French (1993) recommend three factors: MKT is the return on a value-weighted portfolio of NYSE, AMEX, and NASDAQ stocks, less the one-month Treasury bill rate. HML is the average return on high book-to-market (value) stocks, minus the average return on low book-to-market (growth) stocks. SMB is the average return on small stocks, minus the average return on big stocks. Carhart (1997) adds a fourth factor, UMD, defined as the average return on past winners, minus the average return on past losers, where performance of winners and losers is measured over the prior 12 months, skipping the last month's return. Finally, Pástor and Stambaugh (2003) suggest a liquidity factor, LIQ, defined as the average return of stocks with high beta to liquidity, minus the average return of stocks with low beta to liquidity. Alphas are annualized and expressed as percentages. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: Monthly Returns, P10–P1			
	1	2	3	4
Intercept (α)	2.29*	2.56**	4.11***	4.12***
	(1.73)	(2.11)	(3.12)	(3.06)
MKT	0.12***	0.11***	0.08**	0.08**
	(3.11)	(2.92)	(2.54)	(2.54)
HML		−0.04	−0.10	−0.10
		(0.51)	(1.30)	(1.30)
SMB		−0.01	−0.01	−0.02
		(0.19)	(0.22)	(0.22)
UMD			−0.13**	−0.13**
			(2.36)	(2.36)
LIQ				0.00
				(0.04)
Adj. R^2	0.065	0.062	0.145	0.143
Obs.	357	357	357	357

Table 1.11
Determinants of Changes in Centrality

The table presents time-series averages of coefficients from cross-sectional regressions of quarterly changes in centrality on a number of potential explanatory variables, in the spirit of Fama-Macbeth (1973), using data from March 1980 to December 2009. In each specification, the dependent variable, $\Delta\text{CENTRALITY}_t$, is the change in a stock's network centrality during quarter t. $\Delta\text{BREADTH}_t$ is the change in the fraction of mutual funds holding a stock, following Chen, Hong, and Stein (2002), and ΔMFO_t is the change in the fraction of a stock's shares held by mutual funds, as in Chen, Jegadeesh, and Wermers (2000). ANALYSTS_t is the fraction of I/B/E/S analysts providing full-year EPS forecasts for a stock as of the end of the previous year, beginning in March 1984, when prior-year data are first available. LOGSIZE_t is a stock's log market capitalization at the end of quarter t. BM_t and EP_t are a stock's book-to-market ratio and earnings-to-price ratio at the end of quarter t. LAST1YR_t is a stock's cumulative return over the last 12 months, as of the end of quarter t. TURNOVER_t is a stock's average monthly volume as a proportion of shares outstanding during quarter t. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: $\Delta\text{CENTRALITY}_t$				
	1	2	3	4	5
Intercept	−0.0007	−0.0008	−0.0010	0.0268	0.0227
	(0.38)	(0.42)	(0.51)	(0.57)	(0.56)
$\Delta\text{BREADTH}_t$	3.8523***		3.4952***	3.3082***	3.6530***
	(10.62)		(10.98)	(10.24)	(11.77)
ΔMFO_t		0.7039***	0.2979***	0.1489**	0.2347***
		(5.61)	(3.31)	(2.05)	(3.31)
ANALYSTS_t				0.0601	
				(0.79)	
LOGSIZE_t				−0.0015	−0.0013
				(0.59)	(0.59)
BM_t				−0.0006	−0.0004
				(0.54)	(0.33)
EP_t				0.0081	0.0121*
				(1.31)	(1.87)
LAST1YR_t				−0.0081***	−0.0090***
				(3.79)	(4.74)
TURNOVER_t				0.0195**	0.0213**
				(2.09)	(2.37)
Avg. Adj. R^2	0.028	0.012	0.031	0.053	0.066
Obs.	119	119	119	104	119

Table 1.12
Using Changes in Centrality to Predict Future Stock Returns (10 Bins)

The table presents time-series averages, in the spirit of Fama-Macbeth (1973), of coefficients from cross-sectional regressions of quarterly stock returns on changes in centrality and a number of other variables known to predict future stock returns, using data from July 1980 to March 2010. In each specification, the dependent variable, R_{t+1}, is a stock's cumulative return from the end of quarter t to the end of quarter $t+1$. Because extreme changes in centrality are expected to be most informative, and in keeping with the portfolio sorts presented in Tables 1.4 through 1.8, changes in centrality are expressed as two binary indicators: $P1_t$ is a dummy variable equal to one if a stock is in the bottom decile of stocks sorted on changes in centrality, and $P10_t$ is a dummy variable equal to one if a stock is in the top decile of stocks sorted on changes in centrality. All other explanatory variables are as previously described. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: R_{t+1}					
	1	2	3	4	5	6
Intercept	0.0339***	0.0372***	0.0338***	0.0337***	0.0504*	0.0382
	(3.69)	(3.38)	(3.66)	(3.67)	(1.68)	(1.32)
$P1_t$	−0.0057**		−0.0047**	−0.0046**	−0.0032*	−0.0031*
	(2.52)		(2.03)	(2.00)	(1.86)	(1.87)
$P10_t$	−0.0009		−0.0027	−0.0029	−0.0006	−0.0013
	(0.31)		(0.99)	(1.05)	(0.30)	(0.72)
$\Delta BREADTH_t$		0.3958	0.6784**	0.5980**	0.3064**	0.3387**
		(1.04)	(2.55)	(2.38)	(2.07)	(2.38)
ΔMFO_t				0.0675*	−0.0142	0.0149
				(1.74)	(0.54)	(0.53)
$ANALYSTS_t$					0.5839***	
					(3.88)	
$LOGSIZE_t$					−0.0014	−0.0005
					(1.08)	(0.41)
BM_t					0.0021	0.0052**
					(0.94)	(2.17)
EP_t					0.0673***	0.0596***
					(3.21)	(2.90)
$LAST1YR_t$					0.0173***	0.0191***
					(3.06)	(3.51)
$TURNOVER_t$					−0.0985***	−0.1185***
					(2.67)	(3.18)
Avg. Adj. R^2	0.002	0.002	0.004	0.005	0.049	0.053
Obs.	119	119	119	119	104	119

Table 1.13
Using Changes in Centrality to Predict Future Stock Returns (20 Bins)

The table presents time-series averages, in the spirit of Fama-Macbeth (1973), of coefficients from cross-sectional regressions of quarterly stock returns on changes in centrality and a number of other variables known to predict future stock returns, using data from July 1980 to March 2010. As in Table 1.12, the dependent variable in each specification, R_{t+1}, is a stock's cumulative return from the end of quarter t to the end of quarter $t+1$. In contrast to the results presented in that table, the tests below reflect the predictive value of only the most extreme changes in centrality, based on sorts of stocks into twenty portfolios. Specifically, in each regression, changes in centrality are expressed as $P1_t$, a dummy variable equal to one if a stock is in the bottom 5% of stocks sorted on changes in centrality, and $P20_t$, a dummy variable equal to one if a stock is in the top 5% of stocks sorted on changes in centrality. All other explanatory variables are as previously described. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: R_{t+1}					
	1	2	3	4	5	6
Intercept	0.0337***	0.0372***	0.0336***	0.0335***	0.0512*	0.0391
	(3.66)	(3.38)	(3.63)	(3.63)	(1.71)	(1.35)
$P1_t$	−0.0080***		−0.0070**	−0.0070**	−0.0055**	−0.0054**
	(2.87)		(2.46)	(2.48)	(2.35)	(2.49)
$P20_t$	−0.0018		−0.0038	−0.0039	−0.0028	−0.0031
	(0.51)		(1.12)	(1.18)	(1.01)	(1.20)
$\Delta BREADTH_t$		0.3958	0.6723**	0.5939**	0.3124**	0.3364**
		(1.04)	(2.52)	(2.37)	(2.12)	(2.39)
ΔMFO_t				0.0647*	−0.0149	0.0136
				(1.69)	(0.62)	(0.49)
$ANALYSTS_t$					0.5828***	
					(3.86)	
$LOGSIZE_t$					−0.0014	−0.0005
					(1.11)	(0.45)
BM_t					0.0021	0.0051**
					(0.92)	(2.15)
EP_t					0.0677***	0.0601***
					(3.24)	(2.92)
$LAST1YR_t$					0.0172***	0.0191***
					(3.04)	(3.50)
$TURNOVER_t$					−0.0984***	−0.1185***
					(2.67)	(3.18)
Avg. Adj. R^2	0.001	0.002	0.004	0.005	0.049	0.053
Obs.	119	119	119	119	104	119

Table 1.14
Changes in Centrality: Return Predictability at Alternative Horizons (10 Bins)

The table presents time-series averages, in the spirit of Fama-Macbeth (1973), of coefficients from cross-sectional regressions of stock returns at various horizons on changes in centrality and a number of other variables known to predict future stock returns, using data from July 1980 to December 2010. In each specification, the dependent variable, $R_{t+\tau}$, is a stock's cumulative τ-month return, beginning at the end of quarter t. Because extreme changes in centrality are expected to be most informative, and in keeping with the portfolio sorts presented in Tables 1.4 through 1.8, changes in centrality are expressed as two binary indicators: $P1_t$ is a dummy variable equal to one if a stock is in the bottom decile of stocks sorted on changes in centrality, and $P10_t$ is a dummy variable equal to one if a stock is in the top decile of stocks sorted on changes in centrality. All other explanatory variables are as previously described. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: $R_{t+\tau}$					
	τ=1	τ=2	τ=3	τ=6	τ=9	τ=12
Intercept	0.0199	0.0276	0.0382	0.0722	0.1113	0.1518
	(1.30)	(1.11)	(1.32)	(1.28)	(1.24)	(1.25)
$P1_t$	−0.0020**	−0.0022	−0.0031*	−0.0076***	−0.0059**	−0.0041
	(2.13)	(1.58)	(1.87)	(3.50)	(2.35)	(1.33)
$P10_t$	0.0003	−0.0015	−0.0031	−0.0030	−0.0069**	−0.0071**
	(0.25)	(0.89)	(0.72)	(1.23)	(2.26)	(2.12)
$\Delta BREADTH_t$	0.1068	0.0215	0.3387**	0.4623**	0.7373**	0.6566*
	(0.25)	(0.17)	(2.38)	(2.14)	(2.35)	(1.72)
ΔMFO_t	−0.0138	−0.0192	0.0149	0.0593	0.0225	0.0054
	(0.84)	(0.84)	(0.53)	(1.15)	(0.31)	(0.07)
$LOGSIZE_t$	−0.0008	−0.0005	−0.0005	−0.0007	−0.0011	−0.0015
	(1.21)	(0.50)	(0.41)	(0.32)	(0.29)	(0.29)
BM_t	0.0023**	0.0041**	0.0052**	0.0093**	0.0137**	0.0197***
	(2.04)	(2.16)	(2.17)	(2.29)	(2.39)	(2.67)
EP_t	0.0295***	0.0354**	0.0596***	0.1061**	0.1343**	0.1654*
	(2.97)	(2.18)	(2.90)	(2.42)	(2.03)	(1.90)
$LAST1YR_t$	0.0019	0.0056	0.0191***	0.0339***	0.0386**	0.0416**
	(0.57)	(1.26)	(3.51)	(2.97)	(2.41)	(2.16)
$TURNOVER_t$	−0.0202	−0.0190	−0.1185***	−0.2555***	−0.3887***	−0.4877***
	(0.91)	(0.63)	(3.18)	(3.86)	(4.09)	(3.80)
Avg. Adj. R^2	0.046	0.053	0.053	0.056	0.055	0.055
Obs.	119	119	119	119	119	119

Table 1.15

Changes in Centrality: Return Predictability at Alternative Horizons (20 Bins)

The table presents time-series averages, in the spirit of Fama-Macbeth (1973), of coefficients from cross-sectional regressions of stock returns at various horizons on changes in centrality and a number of other variables known to predict future stock returns, using data from July 1980 to December 2010. As in Table 1.14, the dependent variable in each specification, $R_{t+\tau}$, is a stock's cumulative τ-month return, beginning at the end of quarter t. In contrast to the results presented in that table, the tests below reflect the predictive value of only the most extreme changes in centrality, based on sorts of stocks into twenty portfolios. Specifically, in each regression, changes in centrality are expressed as $P1_t$, a dummy variable equal to one if a stock is in the bottom 5% of stocks sorted on changes in centrality, and $P20_t$, a dummy variable equal to one if a stock is in the top 5% of stocks sorted on changes in centrality. All other explanatory variables are as previously described. The t-statistics, given in parentheses, have been adjusted according to the method of Newey-West (1987) to correct for heteroskedasticity and autocorrelation. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Dependent Variable: $R_{t+\tau}$					
	τ=1	τ=2	τ=3	τ=6	τ=9	τ=12
Intercept	0.0200	0.0281	0.0391	0.0733	0.1123	0.1530
	(1.31)	(1.13)	(1.35)	(1.30)	(1.25)	(1.26)
$P1_t$	−0.0034***	−0.0033*	−0.0054**	−0.0113***	−0.0118***	−0.0099**
	(2.78)	(1.69)	(2.49)	(3.30)	(3.29)	(2.37)
$P20_t$	0.0004	−0.0026	−0.0031	−0.0057	−0.0089*	−0.0114**
	(0.30)	(1.29)	(1.20)	(1.64)	(1.82)	(2.08)
$\Delta BREADTH_t$	0.1038	0.0205	0.3364**	0.4566**	0.7128**	0.6446*
	(1.05)	(0.16)	(2.39)	(2.15)	(2.31)	(1.68)
ΔMFO_t	−0.0143	−0.0197	0.0136	0.0589	0.0236	0.0101
	(0.86)	(0.86)	(0.49)	(1.14)	(0.33)	(0.13)
$LOGSIZE_t$	−0.0008	−0.0005	−0.0005	−0.0008	−0.0012	−0.0016
	(1.22)	(0.52)	(0.45)	(0.34)	(0.30)	(0.30)
BM_t	0.0023**	0.0040**	0.0051**	0.0092**	0.0137**	0.0196***
	(2.05)	(2.16)	(2.15)	(2.28)	(2.39)	(2.66)
EP_t	0.0294***	0.0356**	0.0601***	0.1063**	0.1342**	0.1657*
	(2.98)	(2.19)	(2.92)	(2.43)	(2.03)	(1.90)
$LAST1YR_t$	0.0019	0.0056	0.0191***	0.0339***	0.0386**	0.0415**
	(0.57)	(1.26)	(3.50)	(2.96)	(2.42)	(2.16)
$TURNOVER_t$	−0.0200	−0.0191	−0.1185***	−0.2559***	−0.3886***	−0.4885***
	(0.89)	(0.64)	(3.18)	(3.87)	(4.09)	(3.81)
Avg. Adj. R^2	0.046	0.053	0.053	0.056	0.055	0.055
Obs.	119	119	119	119	119	119

Chapter 2

Thank Goodness it's Friday? Opportunistic Filing of Section 13(d) Disclosures by Institutional Investors

2.1 Introduction

Retail investors pay close attention to the investing activities of professional money managers. For decades, the investing magazine *Barron's* has published a weekly column detailing the purchases and sales of activist money managers, while *The Wall Street Journal*'s "Heard on the Street" feature provides a daily roundup of rumored actions by high-profile investors. More recently, a number of websites allow users to track quarterly disclosures of prominent hedge fund investors' holdings. On trading desks across Wall Street, Bloomberg terminals deliver real-time notification of regulatory disclosures reflecting portfolio actions by institutional investors, indicating that even professional money managers, themselves, monitor the investing activity of their peers.

Knowledge of the trades made by large investors might be valuable for a number of reasons. If, as some have suggested (see, for example, Chakravarty, 2001), institutional investors possess private information or exhibit superior ability in processing public information, an institution's decision to buy or sell provides a signal as to the true value of a security, which observers may exploit. Moreover, Brunnermeier and Pedersen (2005) show that even the *uninformed* trading of a large, distressed investor creates an opportunity for strategic "predatory" traders to profit from the price impact of anticipated sales. In any case, if an observer can identify the nature of trading activity—whether orders are the result of informed trading or forced liquidations—the observer may profit by front-running future trades. Of course, such profits come at the expense of the original trader, who must compete with the observer in the market and will realize less favorable prices as a result.

Accordingly, beginning with Kyle (1985), theoretical models of trading suggest that an investor with private information will spread out purchases and sales over time, such that his private signal cannot be easily inferred from the resulting order flow. In fact, as Chan and Lakonishok (1995) document for a sample of prominent investment managers, any institution wishing to transact in large numbers of shares will likely split orders into smaller trades to minimize price impact and execution costs; thus, even investors trading for non-informational reasons will not necessarily execute an entire order in a single trade. To the extent that observers correctly attribute early trades in a sequence to a particular type of investor, later trades will be subject to the kind of exploitation outlined above.

There are, however, circumstances under which an investor, rather than disguising orders, must explicitly *disclose* his trading activity. Investors subject to Section 13(d) filing requirements, for example, must promptly report material changes in their holdings of certain securities—even if the transactions in question only reflect the first in a series of planned purchases or sales—exposing themselves to front-running on future trading in those securities. Although filers may not choose *what* information goes into the disclosure, they do have limited discretion as to *when* the filing is released to the public. It seems natural, then, to ask whether 13D filers take advantage of this flexibility in an effort to minimize the impact of the disclosure on future transactions.

A growing strand of the behavioral finance literature studies the implications of investor inattention—the tendency for investors with limited information-processing resources to make poorer decisions in the face of distractions—for asset prices and market efficiency. Hirshleifer, Lim, and Teoh (2009), for instance, find that when companies release earnings news during days on which many other companies report results, prices exhibit a delayed reaction to earnings announcements, and a more pronounced post-earnings announcement drift. The authors attribute the substantial increase in underreaction for "high-news" days to the great burden such a flood of information places on investors with finite time and cognitive resources. Another empirical test of the investor inattention hypothesis comes from Cohen and Frazzini (2008), who find that the stocks of suppliers only gradually adjust to news about customers' fundamentals, suggesting that a meaningful proportion of investors do not immediately incorporate all available information into security prices.

DellaVigna and Pollet (2009) introduce a theoretical model in which all investors receive a signal about future dividends, but only a fraction pays attention to this news. The authors allow the fraction of attentive investors to vary by day of the week, and prove that prices exhibit a delayed response to news released on high-distraction days; they also show that the long-term price response does not depend on the fraction of inattentive investors. One immediate implication is that managers concerned with the short-run stock price reaction to news will schedule the release of negative information to coincide with high-distraction days. DellaVigna and Pollet (2009) then test for inattention to news released on Fridays, based on the intuition that the weekend, itself, serves to distract many investors. The authors find that

Friday earnings releases meet with a delayed price reaction and a greater degree of post-earnings announcement drift than announcements made on other weekdays. These results corroborate earlier studies documenting a tendency for managers to release poor earnings announcements on Fridays (Penman, 1987).

In this chapter, I extend the literature on investor inattention, investigating the disclosure strategies of institutional investors using a novel data set consisting of Schedule 13D filings and periodic amendments to those filings for a comprehensive sample of large money managers. If a filer believes that public disclosure of the first in a sequence of trades will increase the cost of transacting orders in the near future, then, under the model of DellaVigna and Pollet (2009), that filer has a clear incentive to make the disclosure at a time of high investor distraction. I find evidence that portfolio managers engage in an "opportunistic" filing strategy, timing disclosures in an effort to exploit periods of presumed investor inattention and conceal the liquidation of large positions. Specifically, filers disclosing a material decrease in ownership are considerably more likely to do so on a Friday. Analysis of price formation around disclosures of large sales, however, suggests that investors aren't fooled by opportunistic filers. Instead, the market appears to react quickly to Schedule 13D disclosures, regardless of the filing day, with no detectable post-filing drift in the wake of Friday or Monday through Thursday filings. These results support the view that market participants are rational and well-apprised of the information contained in regulatory disclosures by large investors, and contradict earlier evidence of investor inattention as a source of securities mispricings.

The remainder of the chapter is organized as follows. Subsection 2.2.1 provides a thorough description of the regulatory environment, which imposes constraints on investors that create favorable circumstances for the tests that follow. In Subsection 2.2.2, I describe my data collection procedure and explain the series of filters that I employ in the interest of isolating active portfolio managers. Subsection 2.2.3 formalizes my hypotheses regarding strategic disclosure by Schedule 13D filers and the asset pricing implications of investor inattention to Friday filings. In Subsection 2.3.1, I begin by addressing the investment manager's decision with respect to the timing of Section 13(d) disclosures, showing evidence in support of the hypothesis that institutional investors employ a strategy of opportunistic disclosure. Subsection 2.3.2 presents the results of an event-study analysis which fails to detect underreaction to 13D amendments filed on Fridays. In Subsection 2.3.3, I construct a sample of filers with a particularly strong incentive to hide sales—namely, filers disclosing a sale who go on to sell additional shares in the near future—and show that tests using this subset of filers strengthen support for the conclusions obtained in previous subsections. In Subsection 2.4.1, I consider an alternative explanation for the results documented in Subsection 2.3.1, but find no evidence in support of this hypothesis. Finally, Subsection 2.4.2 summarizes the results of the investigation, and discusses the broader asset pricing implications of my work.

2.2 Preliminaries

2.2.1 The Regulatory Environment

In 1968, Congress enacted the Williams Act, adding Section 13(d) to the Securities & Exchange Act of 1934, requiring an investor or group of investors acquiring greater than five percent beneficial ownership of a company's stock to disclose certain information to the public, including the size of the ownership stake and the purpose of the individual or group's acquisition. Additionally, when information disclosed in the initial filing changes "materially"—for example, when the size of an ownership stake changes significantly—the filer is required to post an amendment updating the original filing. Passage of the Williams Act was intended to provide market participants with an early warning of ownership accumulations that might constitute the beginnings of a change in corporate control (e.g., through a subsequent tender offer or proxy contest). Thus, by design, Schedule 13D filings contain information that is likely not public at the time of disclosure, and that has the potential to significantly affect valuation of the subject firm's shares.

Technically, an individual is considered a beneficial owner of shares if that individual possesses—alone or as part of a group—voting power or investment power (e.g., the power to dispose of or direct the disposition of a security) with respect to the shares in question, or both. An individual with the right to acquire investment power and/or voting power with respect to a company's shares within 60 days also qualifies as a beneficial owner under the regulation. Moreover, when two or more investors have an agreement (not necessarily formal) to act in concert for the purpose of acquiring, holding, or disposing of shares, and those investors' aggregate holdings exceed five percent of the company, each member of the group is deemed a beneficial owner of the group's total holdings for the purposes of determining reporting requirements.[1] In this case, the SEC only requires that one group member submit a filing, to which the filer must append a list of all other members of the group.

For the hypotheses tested in this chapter, timing is an important issue. An initial 13D filing must be submitted within ten days of a breach of the reporting threshold. The requirement for filing amendments to the initial disclosure is more ambiguous, as the Exchange Act only stipulates that amendments be filed "promptly" after a

[1] As an example, suppose that three investors, acting in coordination, each acquire a two percent equity stake in a firm. Because the group's aggregate ownership of six percent exceeds the Section 13(d) reporting threshold, each member of the group is required to file, and each will report beneficial ownership of six percent.

material change. Over time, courts have come to differing conclusions as to what constitutes a "prompt" disclosure. In 1974, the courts determined that a filing within ten days was prompt "by any standard"; later, in a 1989 motion, the SEC ruled that an amendment filed 12 days subsequent to an investor acquiring an additional six percent of a company's shares did not meet the requirement of a timely disclosure (Levy, 2010). Based on these precedents, I operate under the assumption that amendments must be filed within a ten-day window following a change, but that the filer has discretion as to the exact date of submission within this interval. In the case of the initial 13D filings and all associated amendments, the filer has some choice as to when information becomes public.

In 1977, in part to ease the compliance burden associated with Section 13(d) requirements, the SEC created Schedule 13G, a "short-form" beneficial ownership report, which may be filed by passive investors or qualified institutional investors, in lieu of Schedule 13D. A qualified institutional investor is one whose acquisition takes place in the ordinary course of business, and who has no intention of influencing control of the issuer (examples include broker dealers, banks, and insurance companies); a passive investor is any other investor with no intention of changing or influencing control, who is the beneficial owner of no more than twenty percent of the stock in question. Although there are some circumstances under which a passive investor might still elect to file the "long-form" 13D (when, for instance, the investor wishes to retain an option to influence control of the firm in the future) the availability of Schedule 13G ensures that a vast majority of 13D filings stem from the holdings of "active" investors, the subset of investors with whom this study is primarily concerned.

I focus on active, Schedule 13D investors because the motives to exploit investor inattention seem particularly strong for this group. Consider an active investor who holds shares for the purposes of changing or influencing control. It is reasonable to expect that a substantial reduction in this investor's ownership—which significantly diminishes the investor's voting power—signals, with high probability, that the investor has abandoned his original purpose and will eventually liquidate any remaining holdings. Intuitively, such sales can be assumed to occur sooner rather than later, as many active investors (activist hedge funds, for example) have less mandate for maintaining passive positions, which they must hold at the expense of other active investment opportunities. Market observers following this logic will react to Schedule 13D amendments disclosing large decreases in an active investor's position by revising upward their beliefs about the probability of sizable future sales; trading based on these beliefs will exacerbate the investor's price impact on subsequent orders. Thus, the active investor has a clear incentive to minimize the visibility of a 13D filing, possibly by strategically timing its release.

2.2.2 Data

Beginning in 1996, the SEC made electronic submission mandatory for many filings, including all Section 13(d) disclosures and amendments. Filers transmit forms in plain text or HTML format directly to the SEC through an automated system. The filings receive an electronic timestamp, and become immediately available to the public in the SEC's EDGAR database. The SEC accepts submissions on weekdays, from 6 a.m. to 10 p.m., Eastern time (ET), except on federal holidays. I collect full text submissions for all 13D filings and amendments (henceforth, 13D/A filings) from EDGAR, along with associated date and timestamps (109,625 filings, in all). Prior to 2003, electronic filings bear a datestamp, but not a timestamp; this prevents me from verifying the accuracy of the filing date. Because correctly identifying the date of each filing is of critical importance in performing the event studies of Subsection

2.3.2, my data begin in January 2003, at the inception of the electronic timestamp, and continue through January 2010. I design a text-mining algorithm to extract, from each filing, a CUSIP identifying the subject company, and information about the individual or group's beneficial ownership. Using the filers' and subjects' unique SEC identifiers, I am able to track 13D filings and amendments over a seven-year period. Owing to idiosyncrasies in the formatting of some filings, it was not possible to automatically process all disclosures. For the purposes of this study, rather than hand-collect these data, I have simply excluded those filings for which my algorithm failed to extract the relevant items.

Any institutional investment manager (defined by the SEC as an entity that invests or transacts in securities for its own account, or has investment discretion over the account of any other person or entity) with in excess of $100 million in assets under management is required to submit Form 13F, a quarterly report containing information on all of its long positions in Section 13(f) securities, including exchange-traded and NASDAQ stocks, equity options and warrants, and shares of closed-end funds. In this chapter, one objective is to better understand the filing behavior of institutional money managers. To this end, following Greenwood and Shor (2009), I cross-reference the collection of 13D filings and amendments with a database of 13F filings collected from EDGAR, with the goal of separating portfolio investments from other types of 13D positions (corporate cross-holdings, for example, established in advance of a merger). Specifically, for each 13D filing in my original sample, I determine whether the filer has also submitted form 13F at some point prior to submission of the 13D, and remove from my sample any 13D filing (and all subsequent amendments) associated with a filer who has not also submitted a 13F, effectively restricting my investigation to large, institutional investors (e.g., mutual funds, hedge funds, pension funds). This filter leaves me with a sample of 6,962 13D and 13D/A filings, representing 2,129 individual filing "chains" (a 13D filing, along with all associated amendments).

For each of the 1,311 subject companies in this sample, I attempt to collect prices from the CRSP database; I exclude from my sample 250 companies for which CRSP data are not available. For the remaining 1,061 firms, I collect data on prices, returns, volume, closing bid-ask spreads, and shares outstanding. This leaves 2,785 filings for which I have data on ownership and stock characteristics. When available, I also determine the number of analysts making full-year EPS forecasts (from I/B/E/S), along with the number of institutional holders and number of shares held by institutions (both from the Thomson Reuters 13F holders database); I interpret these measures as rough proxies for firm visibility.

I present summary statistics for the filings in my sample in Table 2.1, along with identical statistics calculated for S&P 500 firms over the same period, for comparison.

2.2.3 Hypotheses

An important implication of the model presented in DellaVigna and Pollet (2009) is that when investors face distractions on a particular day of the week, a manager wishing to minimize the effect of a news release on the stock price in the near-term will choose to disclose the news when investors are most distracted. In the present context, the manager is a decision maker at a large, "active" investment firm, and the news release relates to a material change in the firm's investment with respect to a particular stock. The amendment in question might reflect a change in the size of the manager's position, or simply some modification to the manager's stated reasons for holding the security. Given the regulatory constraints outlined in Subsection 2.2.1, the manager has no discretion over what information the filing contains, but does have a choice as to when other investors first observe the news. Specifically, the manager may file on any weekday within a window of roughly ten days, effectively allowing the manager to submit on the weekday of his choice. The manager filing to disclose a substantial decrease in the size of his holdings, but who hasn't yet disposed of the entire position, must be concerned that publicizing the liquidation will result in higher costs to selling his remaining shares. I hypothesize that such a manager will be significantly more likely than other managers to opt for a Friday disclosure.

Hypothesis 2.1. *Filers disclosing a substantial decrease in holdings, but who have not disposed of the entire position, will be more likely to submit the filing on a Friday, when investor inattention is believed to be greatest.*

An investor attempting to quietly build an activist position, but required under Section 13(d) to disclose the initial purchase or a material increase in existing holdings, might likewise wish to avoid detection—and hence minimize the adverse impact of front-runners—by filing on days of high investor inattention. There is also the possibility that, having already established a substantial stake, the filer will welcome purchases by other investors, which only serve to increase the value of the filer's existing position. Because the motive to opportunistically disclose is less clear for buyers than it is for sellers, I restrict my analysis to Section 13(d) investors reporting a decrease in holdings.

Of course, if a manager wishing to sell all holdings in a particular stock is able to liquidate the entire position in the ten calendar days before submitting an amendment, there is little motivation to strategically time the disclosure. Positions subject to Section 13(d) requirements are necessarily large, however, making it difficult for a seller to dispose of all holdings over a short horizon without incurring steep transactions costs. For each filing in my sample, I compute the average number of days that a manager would require to completely liquidate the shares remaining *after* disclosure of the first decrease in holdings (simply the number of remaining shares, divided by the average trading volume of the subject company's stock). This measure is conservative, as it assumes that the filer is the only one selling shares, and doesn't account for any shares that the filer actually sold in the days preceding the 13D/A disclosure. I find that filers reporting a decrease in holdings require an average of 19.4 days (and a median of 6.7 days) to dispose of any remaining shares. Given that the subject companies in my sample tend to be quite small, with relatively large bid-ask spreads, it is therefore unlikely that most managers have the option to liquidate all holdings during the pre-filing window, increasing the incentive to strategically time submission of 13D/A filings.

As in DellaVigna and Pollet (2009), I take Friday to be the weekday of highest distraction to market participants. While other stories have been proposed—some have cited Monday as a particularly busy day, since investors must process and trade on a weekend's worth of information—most of the evidence from the papers cited supports my decision to focus on Friday filing activity. Alternatively, it is highly likely that a holiday (e.g., Good Friday) during which no trading takes place, but on which the SEC accepts submissions, is a day of high investor inattention; retail investors might be less inclined to follow financial news during market closures, and not all professional investors will work on these days. Unfortunately, it is unlikely that an active institutional investor will have the flexibility to time large sales to coincide with such holidays, and the filers that do happen to sell near these holidays constitute too small a proportion of my sample to base tests on these days, alone. Similarly, summer months (during which investors might be less attentive, in general) have too little filing activity in my sample to provide useful tests of the above hypothesis.

There are, of course, two players in the story presented above: managers and investors. Hypothesis 2.1 addresses the decision by managers to release bad news during periods of investor inattention. Another critical question concerns the reaction of investors to such disclosures: Do investors exhibit greater underreaction to Friday filings than to disclosures on other weekdays? Simply put, it is important to know whether markets quickly process and incorporate all relevant public information, regardless of the day on which the information becomes available; this is a straightforward test of investor rationality and market efficiency.

Hypothesis 2.2. *Investors respond more slowly to information obtained during times of high distraction, inducing greater underreaction of prices to the information content of Friday filings.*

2.3 Tests

2.3.1 The Filing Decision

It is a simple matter to test whether or not filers engage in a strategy of "opportunistic" disclosure, more frequently releasing news perceived as unfavorable on Fridays. I consider two groups of filers: those filing amendments to disclose a decrease in holdings of twenty percent or more (measured as a proportion of the position claimed in the initial 13D), but not a complete closure of the position; and the group made up of all other 13D/A filers. I use twenty percent as the cutoff because, for a hypothetical five-percent holder, a smaller reduction in ownership does not meet the threshold of "materiality," and does not require an amendment under Section 13(d); filings associated with such changes likely reflect some other modification to the original 13D, which might be good or bad news from the perspective of the manager. Discrepancies in submission frequencies across filer types are readily apparent in Figure 2.1. Notice that the "control" group—filers *not* disclosing a large decrease in holdings—appears uniformly distributed over weekdays. However, a filer revealing a substantial sale is significantly more likely to do so on a Friday, and less inclined to submit the filing on a Monday.

In Table 2.2, I present corresponding two-tailed tests for differences in the proportion of filings observed for each day of the week, based on a normal approximation to the binomial distribution, for filers disclosing ten-, twenty-, and thirty-percent reductions in holdings. The pattern of Friday filing is strongest for filers disclosing the largest declines in holdings (Panel C). When I include 13D/A filings that report a decrease in holdings of greater than 10%, but less than 20% (Panel A), the effect vanishes, supporting the notion that this represents a logical value for the threshold in Figure 2.1. These results provide strong support for Hypothesis 2.1, in the sense that a manager's choice of submission day is clearly related to the content of the disclosure, with managers reporting large sales more likely to do so on a Friday.

Based on the timestamps associated with each filing in my sample, it is possible to consider whether filers also time disclosures *within the day*, in an attempt to exploit hours of perceived investor inattention. As mentioned in Subsection 2.2.2, filings must be submitted electronically, between 6 a.m. and 10 p.m. ET. Additionally, for a filing to be considered submitted on a given day for the purposes of compliance with Section 13(d) deadlines, the disclosure must have been filed before 5:30 p.m. ET on that day. U.S. exchanges close for trading at 4 p.m. ET, leaving 90 minutes after the close (from 4 p.m. to 5:30 p.m. ET), in which all submissions are still considered official as of that day, but during which investors may not trade based on the information content of new disclosures. Filers seeking to delay the disclosure of large sales should favor submission in this 90 minute window over submission during trading hours.

Table 2.3 provides the hourly distribution of electronic submissions on each weekday for a group of large sellers (specifically, those 13D/A filers reporting a decrease in holdings in excess of 20%, but not a complete liquidation), and for all other filers. In both subsamples and across weekdays, submissions are concentrated in the late afternoon, consistent with the notion that—at least from the perspective of a portfolio manager—regulatory compliance is secondary to the execution of trades and can be more easily addressed after markets close. Table 2.4 reports the proportion of disclosures submitted in the 4 p.m. to 5:30 p.m. window (previously identified as a logical period of high intraday investor inattention) for large sellers and for all other filers. Consistent with the hypothesis that filers reporting a large decrease in holdings will favor disclosure during periods of perceived investor inattention, I find that large sellers are much more likely than other filers to submit disclosures in the 90-minute period immediately following the market close. These findings suggest that managers strategically time not only the day of the filing, but also the exact time of the disclosure, in an effort to exploit investor inattention.

2.3.2 Investors' Reaction

In this subsection, I investigate the asset pricing implications of opportunistic filing behavior detected in my previous tests. Although a manager might attempt to exploit periods of presumed investor distraction, it is not necessarily the case that investors fail to incorporate information from a Friday filing into the stock price of the subject company. In order to test Hypothesis 2.2—that is, in order to determine whether investors exhibit more pronounced underreaction to Friday filings than to filings submitted on other weekdays—I perform a series of event studies using data from my sample of Schedule 13D filers, calculating abnormal returns in the window surrounding each filing, and searching for significant differences in investors' reactions to filings submitted on different days of the week.

For each filing in my sample, I define the event date (designated by τ) as the date on which the disclosure becomes public, and construct an event window spanning 38 trading days, from date T_1 (seven days before the filing date), to T_3 (30 days after the disclosure occurs). Because the filer must submit an amendment within approximately ten calendar days of a material change, including the day of the filing, this leaves seven trading days and two weekend days prior to the filing date on which additional sales might have taken place. Since the event study only records activity on trading days, I select seven days as the length of the pre-event measurement period. I also define an estimation window, consisting of the 200 trading days immediately prior to the event window, beginning at date T_0. It is not possible to construct the event and estimation windows described for all filings in my sample due to limitations in the availability of CRSP data for some stocks. As a result, I must exclude an additional 353 filings from the analysis in this subsection, leaving a sample of 2,342 filings.

For each event in the sample, I estimate parameters of the normal return model:

$$\begin{aligned} R_{it} - R_{ft} &= \alpha_i + \beta_{MON,i} I_{MON,t} + \beta_{MKT,i}(R_{MKT,t} - R_{ft}) \\ &\quad + \beta_{SMB,i} R_{SMB,t} + \beta_{HML,i} R_{HML,t} + \beta_{UMD,i} R_{UMD,t} + \varepsilon_{it}, \end{aligned} \tag{2.1}$$

where $R_{it} - R_{ft}$ represents the date-t excess return on the subject company stock, and $R_{MKT,t} - R_{ft}$ is the date-t excess return on all NYSE, AMEX, and NASDAQ stocks, from the CRSP database. To account for the "weekend effect" documented in French (1980)—which is potentially important, given my interest in comparing abnormal returns by day of the week—I include the indicator variable, $I_{MON,t}$, equal to unity when date t falls on a Monday, and zero otherwise. $R_{SMB,t}$ and $R_{HML,t}$ represent the returns on the Fama and French (1993) "small-minus-big" and "high-minus-low" factors, while $R_{UMD,t}$ is the return on the Carhart (1997) momentum factor (all from French's website). Finally, ε_{it} is the disturbance term, assumed to have zero mean. I estimate abnormal returns for the subject company stock on date t, denoted $\widehat{AR}_{it}$ as:

$$\widehat{AR}_{it} = R_{it} - \widehat{NR}_{it} \tag{2.2}$$

where R_{it} is the observed date-t return on the subject company stock, and $\widehat{NR}_{it}$ is the stock's normal return at date t, estimated from the model expressed in Equation (2.1). Given the large estimation window, the abnormal return has variance $\sigma^2(\widehat{AR}_{it}) = \sigma^2_{\varepsilon_i}$.

For each subject company stock, I aggregate abnormal returns over time, defining the estimated cumulative abnormal return from date τ_1 to τ_2 as:

$$\widehat{CAR}_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} \widehat{AR}_{i\tau}, \tag{2.3}$$

which has asymptotic variance equal to $\sigma_i^2(\tau_1, \tau_2) = (\tau_2 - \tau_1 + 1)\sigma^2_{\varepsilon_i}$. I then average across securities, obtaining the estimated aggregated cumulative abnormal return:

$$\overline{CAR}(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^{N} \widehat{CAR}_i(\tau_1, \tau_2), \tag{2.4}$$

which has estimated variance given by:

$$\widehat{var}[\overline{CAR}(\tau_1, \tau_2)] = \frac{1}{N^2} \sum_{i=1}^{N} \hat{\sigma}_i^2(\tau_1, \tau_2). \tag{2.5}$$

Table 2.5 provides sample average abnormal returns computed for each day in the event window, estimates of aggregated cumulative abnormal returns (each beginning with day -7), and corresponding p-values for a test of the null hypothesis that $\overline{CAR}(-7, \tau) = 0$, for two classes of events: (1) filings submitted on Monday through Thursday, which disclose a decrease of ownership in excess of 20%, but not a complete liquidation of the position, and (2) filings with the same characteristics, but submitted on Friday. I present this information graphically in Figures 2.2 and 2.3. The plots reveal a similar pattern: subject companies experience an initial accumulation of negative abnormal returns, which appears to cease on the filing date.

Negative abnormal returns in advance of the filing date are consistent with the price impact caused by a large sale of shares—the event that triggered a disclosure in the first place. As such, this component of the results is not altogether surprising. Due to the small number of events in each subset of filings and the magnitude of measurement errors in the market model regressions, it is somewhat difficult to draw inferences from estimates of abnormal returns observed subsequent to the filing date. One thing is relatively clear: there appears to be little support for underreaction in *either* subsample, based on the results in Table 2.5. We do not observe a significant negative drift after the disclosure of "bad" news for either the Monday through Thursday group, or the group of Friday filers. My results suggest that, if anything, the shares of subject companies recover a portion of the losses due to the initial sales in the six trading weeks following a disclosure.

Next, I compare the price response to 13D/A disclosures of large sales (but not complete liquidations) by Monday through Thursday filers with the price reaction to the same disclosures made by Friday filers. Table 2.6 presents the results of this comparison; I provide an estimate of the difference in $\overline{CAR}$, along with p-values corresponding to two-tailed tests of the hypothesis that there is no difference in responses. I break each event window into a pre-filing period and a post-filing period, looking for differences in either the price impact of the liquidation itself, or the market response in the wake of a disclosure. The evidence in Table 2.6 uniformly refutes Hypothesis 2.2. It appears that the market reaction to information conveyed in Schedule 13D/A disclosures is independent of the weekday on which a filing becomes public.

2.3.3 Analysis of Sequential Sales

The tests presented in the preceding sections rest on the assumption that managers reporting a large decrease in holdings are likely to pursue additional sales in the near future, and should thus wish to minimize the price impact on subsequent liquidations by hiding the initial sale. Those filers with explicit plans to conduct further sales in the near future should have a particularly strong incentive to time disclosures such that they coincide with periods of high investor inattention. Of course, not all managers go on to report additional decreases in holdings in the months following the initial 13D/A filing. Examining each series of filings in my sample, however, I am able to identify "sequential sellers": filers making additional sales shortly after the initial report of a decrease in holdings. In this subsection, I present additional tests of Hypotheses 2.1 and 2.2 based on this refined segmentation of my sample.

First, I assign each Section 13(d) amendment to one of three groups: (1) 13D/A filings disclosing a decrease in holdings, but not a complete liquidation of the position, followed by at least one additional sale within a fixed number of days, denoted as "sequential sales"; (2) 13D/A filings disclosing a decrease in holdings, but not a complete liquidation of the position, but with *no* additional sales within the same fixed number of days, denoted as "isolated sales"; and (3) all 13D/A disclosures not classified as "sequential sales", denoted as "other filings". Table 2.7 presents the proportion of filings from each group occurring on a Friday for various specifications of the cutoff for counting near-future sales. There is no significant difference in the frequency of Friday filings associated with sequential sales and isolated sales. Moreover, consistent with the results in Panel A of Table 2.2, filers reporting a decrease in holdings—whether sequential or isolated—show no greater propensity to file on the last day of the week than the control group of all other filers.

Table 2.8 presents a similar test, where sequential and isolated sales are redefined to include only those initial sales representing a decrease in holdings in excess of 20%, measured as a proportion of initial ownership, but not a complete liquidation of the position. Accordingly, the category of "other" filings expands to include all filings reporting a decrease in holdings of 20% or less. Sorting on the size of the initial sale, I find that large sequential sellers file much more frequently on Fridays than large isolated sellers for most specifications of the cutoff used to determine near-future sales, with significance at the 5% level in the case of the three-week and one-month cutoffs (although it seems likely that with more observations, the results would become significant for additional values of this threshold). The increased propensity for filing on Fridays is much more pronounced for the group of large sequential sellers than for the sample of all large sellers presented in Panel B of Table 2.2, suggesting that the incentive to hide an initial sale increases substantially with the expectation of subsequent liquidations, strengthening support for Hypothesis 2.1.

Using the sample of sequential sales described above, and based on the methodology outlined in Subsection 2.3.2, I perform event study tests of Hypothesis 2.2, with and without a sort on size, and for all values of the cutoff used to count near-future sales. Unfortunately, the small sample sizes obtained from the identification of sequential sellers render most specifications of the event study unusable. For the two cases in Table 2.8, Panel B for which the differences in the frequency of Friday filing between both the sequential/other groups and the sequential/isolated groups are significant—i.e., the three-week and one-month cutoffs for large sequential sellers—I present event study results in Tables 2.9 and 2.10. In each case, there is no sign that investors underreact; neither is there any strong indication that the market reacts differently to the filings of sequential sellers submitted on different days of the week. I conclude that, while the sample of sequential sales provides greater support for the opportunistic filing behavior hypothesized in Subsection 2.2.3, there is no evidence of investor inattention to Friday filings by sequential sellers.

2.4 Discussion

2.4.1 Alternative Explanations

Having presented evidence that institutional money managers engage in an opportunistic filing strategy, scheduling the release of potentially negative information to occur on days of perceived investor inattention—and summarily rejected the notion that this strategy confers any benefits on managers in the way of diminished price response to the information content of such disclosures—it seems natural to consider whether alternative explanations exist to reconcile the filing behavior of Section 13(d) investors with the lack of underreaction by the market to Friday filings. In this subsection, I briefly consider a number of competing stories which might rationalize the findings presented in Subsection 2.3.1.

It is quite possible that microstructure concerns are not the only factors influencing a filer's timing decision. Aside from telegraphing future order flow, the disclosure of large sales might also signal to observers that a fund is in distress, or that a portfolio manager has lower ability than previously assumed. While my results in Subsection 2.3.2 seem to suggest that some subset of market participants is incorporating the information from 13D/A filings into prices, it is possible that Friday filings do escape notice by another subset of investors, whose beliefs—while not important for price formation—matter to the filer in some other context. In this case, the decision to file on Friday might be optimal, even if it affords the manager no benefits with respect to minimizing the price impact on future transactions.

In the preceding tests I have intentionally excluded filings disclosing a complete liquidation of holdings from consideration as "large sales"; if the reason for exploiting perceived investor inattention is to create more favorable conditions for the disposal of one's remaining shares, once a manager has closed the position, there is little motive for a Friday filing. However, if the manager is concerned with the reputational consequences of revealing large sales, then Hypothesis 2.1 should apply to completed liquidations, as well. Unfortunately, my sample of filings reflecting closed positions is particularly small (this is likely because once an investor reports sales bringing ownership below the five percent threshold requiring disclosures under Section 13(d), amendments to the original filing are no longer necessary). Table 2.11 summarizes my data on completed liquidations. Preliminary indications provide no support for the alternative view that Friday filers have motives other than reducing price impact.

2.4.2 Conclusions

Does the investor inattention hypothesis—documented in studies of earnings announcements and customer-supplier relationships, among others—apply to the release of Section 13(d) regulatory disclosures? Using a novel data set consisting of 13D and 13D/A filings over seven years, and building on the implications of a theoretical model of investor inattention, I search for evidence that managers of large investment portfolios exploit perceived periods of high investor distraction to minimize the adverse impact of disclosing large sales. I confirm that portfolio managers tend to submit

Schedule 13D/A opportunistically: managers reporting substantial decreases in holdings clearly favor Friday disclosure to disclosure on other weekdays, relative to filers reporting small sales or other changes in status. An analysis of timestamps reveals that large sellers also favor disclosure in the 90 minutes immediately following the market close, a time of day during which investors are less likely to notice new filings.

Using event study methodology, I test the hypothesis that opportunistic filings by sellers induce greater underreaction to Friday filings than filings made on other weekdays. Both Monday through Thursday disclosures and Friday filings exhibit a pattern of significant, negative abnormal returns in the days leading up to the filing, however, I find that the subject company stock price reaction in both the pre- and post-filing periods is independent of the day on which managers disclose sales. Simply put, I find no support for the underreaction hypothesis. Analysis of "sequential sellers"—filers disclosing the first in a series of sales—strengthens the case for strategic filing on the part of managers, but fails to uncover abnormal return drift in the wake of Friday filings by sequential sellers. Finally, I entertain an alternative explanation that might tie together the results in Subsections 2.3.1 and 2.3.2—namely, that something like reputational concerns might motivate Friday filings, even in the absence of any benefits resulting from delayed price reaction. I can find no conclusive support for this explanation in a small sample of filings disclosing completed liquidations.

My findings suggest that, despite evidence of investor inattention to other categories of news about stocks, a large enough subset of market participants diligently processes Section 13(d) disclosures to ensure that market prices fully reflect the information content of Schedule 13D filings and amendments. The fact that managers persist in attempts to exploit investor irrationality, despite the efficiency of stock price reactions to Friday filings, might be viewed as supporting those researchers of corporate finance who—observing the interactions between managers and investors—

perceive “a rational market dealing with irrational managers” (Subrahmanyam, 2007; examples of papers in this spirit include Roll, 1986, and Malmendier and Tate, 2005). On the other hand, because filing on a Friday as opposed to a Thursday or Monday is a relatively low-cost decision, managers lose little by attempting to “fool” investors, even if such efforts are ultimately in vain.

2.5 References

Brunnermeier, M. K., & Pedersen, L. H. (2005). Predatory trading. *The Journal of Finance, 60*(4), 1825-1863.

Chakravarty, S. (2001). Stealth-trading: Which traders' trades move stock prices?. *Journal of Financial Economics, 61*(2), 289-307.

Chan, L. K., & Lakonishok, J. (1995). The behavior of stock prices around institutional trades. *The Journal of Finance, 50*(4), 1147-74.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance, 63*(4), 1977-2011.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance, 64*(2), 709-749.

Doyle, J. T., & Magilke, M. J. (2009). The timing of earnings announcements: An examination of the strategic disclosure hypothesis. *The Accounting Review, 84*(1), 157-182.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of Financial Economics, 8*(1), 55-69.

Greenwood, R., & Schor, M. (2009). Investor activism and takeovers. *Journal of Financial Economics, 92*(3), 362-375.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance, 64*(5), 2289-2325.

Kothari, S., & Warner, J. (2007). Econometrics of event studies. In B. E. Eckbo (Ed.), *Handbook of corporate finance: Empirical corporate finance* (pp. 3-36). Amsterdam: North-Holland.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 53*(6), 1315-1336.

Levy, S. (2010). *Regulation of securities: SEC compliance and practice.* New York: Aspen Law & Business.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature, 35*(1), 13-39.

Malmendier, U., & Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics, 124*(4), 1593-1638.

Penman, S. H. (1987). The distribution of earnings news over time and seasonalities in aggregate stock returns. *Journal of Financial Economics, 18*(2), 199-228.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *The Journal of Business, 59*(2 Part 1), 197-216.

Subrahmanyam, A. (2007). Behavioural finance: A review and synthesis. *European Financial Management, 14*(1), 12-29.

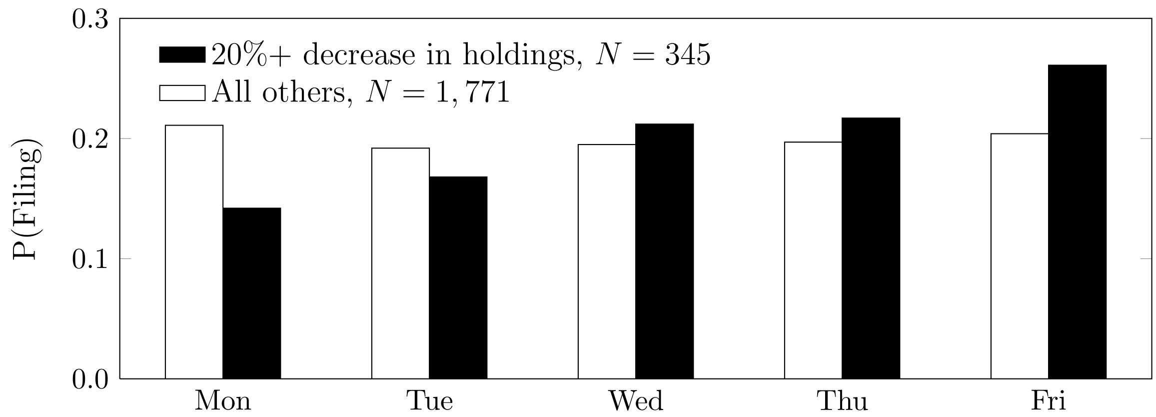

Figure 2.1. Frequency of Filing by Day of Week. A plot of weekday distributions corresponding to: (1) 13D/A filings disclosing a decrease in holdings in excess of 20%, as a proportion of the original 13D position, but not a complete liquidation of the position (black bars); and (2) all other 13D/A filings (white bars).

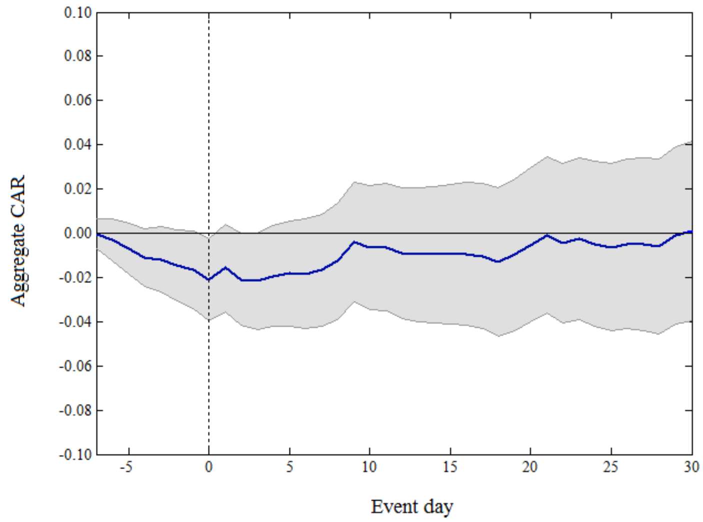

Figure 2.2. Price Impact, Monday through Thursday Filers. Price reaction to the disclosure of large sales by institutional investors on Mondays through Thursdays (with 95% confidence band shaded). The data include all filers in the sample disclosing a reduction in ownership in excess of 20%, but not closure of the position ($N = 193$).

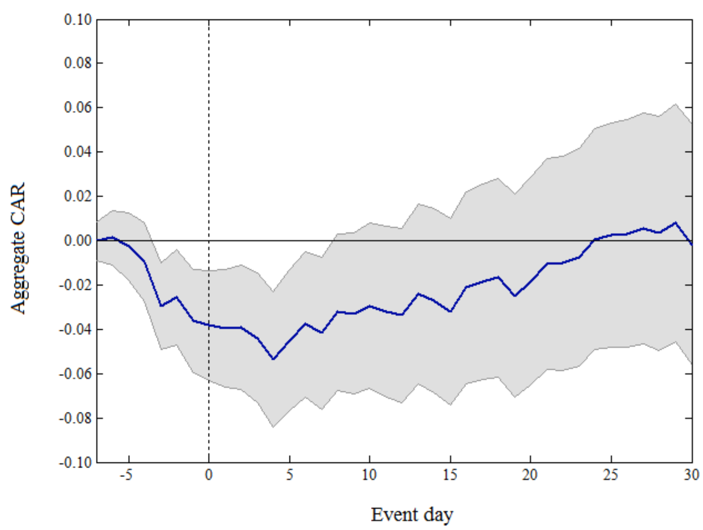

Figure 2.3. Price Impact, Friday Filers. Price reaction to the disclosure of large sales by institutional investors on Fridays (with 95% confidence band shaded). The data include all filers in the sample disclosing a reduction in ownership in excess of 20%, but not a complete liquidation of the position ($N = 72$).

Table 2.1
Subject Firms, Schedule 13D Filing Database: Summary Statistics

The table provides summary statistics for subject companies associated with filings in the final sample, covering 01/2003 through 01/2010, as well as statistics on S&P 500 firms, from 01/2003 through 12/2007. Firm characteristics include: market capitalization (in $ millions), as well as price-to-earnings ratios and book-to-market ratios (both calculated from Compustat data). Proxies for firm visibility are: the number of analysts making full fiscal-year EPS forecasts (from I/B/E/S), along with the number of institutional holders and percentage of outstanding shares held by institutions (based on holdings reported in the Thomson Reuters 13F holders database). Measures of liquidity consist of: a stock's average turnover, computed as the mean daily trading volume divided by shares outstanding; and average bid-ask spread, based on closing bid and ask quotes. Both liquidity measures use CRSP data, with averages for the sample of filings calculated for the six months prior to each filing date.

Panel A: Firms from sample of 13D and 13D/A filings								$N = 2,159$
	Firm characteristics			Firm visibility			Liquidity	
	Size, $M	P/E	B/M	#Analysts	#Holders	%Held	Turnover	Bid-ask
Mean	$1,191.9$	31.8	0.989	7	107	0.717	0.010	0.054
St. dev.	$2,908.8$	53.5	1.583	7	91	0.267	0.009	0.070
25%	129.3	12.7	0.349	3	45	0.529	0.004	0.022
50%	367.2	19.6	0.624	6	88	0.765	0.008	0.032
75%	$1,230.3$	31.6	1.050	10	138	0.899	0.013	0.056

Panel B: S&P 500 firms								$N = 2,450$
	Firm characteristics			Firm visibility			Liquidity	
	Size, $M	P/E	B/M	#Analysts	#Holders	%Held	Turnover	Bid-ask
Mean	$23,595.9$	28.2	0.459	18	428	0.691	0.009	0.040
St. dev.	$41,464.2$	75.4	0.302	11	261	0.232	0.007	0.025
25%	$5,602.7$	14.8	0.248	12	278	0.604	0.005	0.023
50%	$11,070.1$	19.2	0.388	18	366	0.731	0.007	0.032
75%	$22,114.5$	25.5	0.600	25	510	0.832	0.010	0.050

Table 2.2
Disclosure of Large Sales Compared to All Other 13D/A Filings

The table presents the weekday distributions of: (1) 13D/A filings disclosing a large decrease in holdings, denoted as "Large Sellers"; and (2) all other 13D/A filings, denoted as "All others". The last column provides an estimate of the difference in proportions for each day of the week, with the p-value for a two-tailed test of the null hypothesis that the proportions are equal, based on a normal approximation to the binomial distribution, in parentheses. In Panel A, the "Seller" has disclosed a decrease of holdings in excess of 10%, measured as a proportion of initial ownership, but has not closed the position; in Panels B and C, the "Seller" has disclosed a decrease in excess of 20% and 30%, respectively. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

Panel A: "Large Sellers" report a decrease in holdings of greater than 10%					
	Large Sellers		All others		Difference, %
	N	%	*N*	%	
Monday	73	15.9	350	21.1	−5.2 (0.014) **
Tuesday	79	17.2	319	19.2	−2.0 (0.334)
Wednesday	103	22.5	316	19.1	3.4 (0.103)
Thursday	98	21.4	326	19.7	1.7 (0.412)
Friday	105	22.9	347	20.9	2.0 (0.356)
All	*458*		*1,658*		

Panel B: "Large Sellers" report a decrease in holdings of greater than 20%					
	Large Sellers		All others		Difference, %
	N	%	*N*	%	
Monday	49	14.2	374	21.1	−6.9 (0.003) ***
Tuesday	58	16.8	340	19.2	−2.4 (0.299)
Wednesday	73	21.2	346	19.5	1.6 (0.489)
Thursday	75	21.7	349	19.7	2.0 (0.388)
Friday	90	26.1	362	20.4	5.6 (0.019) **
All	*345*		*1,771*		

Panel C: "Large Sellers" report a decrease in holdings of greater than 30%					
	Large Sellers		All others		Difference, %
	N	%	*N*	%	
Monday	35	14.2	388	20.8	−6.6 (0.015) **
Tuesday	40	16.2	358	19.2	−3.0 (0.263)
Wednesday	54	21.9	365	19.5	2.3 (0.387)
Thursday	52	21.1	372	19.9	1.1 (0.672)
Friday	66	26.7	386	20.7	6.1 (0.029) **
All	*247*		*1,869*		

Table 2.3
13D/A Filings: Timing of Electronic Submissions

The table presents the distribution of disclosures, by hour and weekday, for: (1) 13D/A filings disclosing a decrease in holdings in excess of 20%, as a proportion of the original 13D position, but not a complete liquidation of the position, denoted as "Sale"; and (2) all other 13D/A filings, denoted as "Other". Figures represent the percentage of each filing type submitted in a given hour on a given weekday, as a proportion of all filings of that type submitted on that day of the week.

Time (ET)	Mon		Tue		Wed		Thu		Fri	
	Sale	Other	Sale	Other	Sale	Other	Sale	Other	Sale	Other
06:00–06:59	2.0	0.8	0.0	0.9	0.0	0.9	0.0	1.1	1.1	0.0
07:00–07:59	0.0	2.1	0.0	0.3	0.0	0.9	0.0	0.6	0.0	0.6
08:00–08:59	0.0	2.9	0.0	1.2	0.0	1.4	1.3	2.9	1.1	1.1
09:00–09:59	2.0	2.7	5.2	4.4	1.4	2.9	2.7	1.1	1.1	3.3
10:00–10:59	0.0	4.3	3.4	3.5	1.4	4.6	4.0	5.2	0.0	1.7
11:00–11:59	4.1	3.2	3.4	4.7	4.1	3.5	1.3	3.7	1.1	5.2
12:00–12:59	4.1	4.0	3.4	4.4	0.0	5.2	1.3	3.2	3.3	7.2
13:00–13:59	8.2	3.7	3.4	3.5	5.5	4.0	4.0	6.3	6.7	5.8
14:00–14:59	2.0	6.1	6.9	9.1	6.8	7.8	5.3	8.3	7.8	8.0
15:00–15:59	6.1	15.0	12.1	13.2	9.6	11.6	2.7	16.3	8.9	10.5
16:00–16:59	26.5	20.1	20.7	18.5	17.8	21.1	42.7	19.8	32.2	23.2
17:00–17:59	32.7	24.6	34.5	24.4	42.5	23.4	28.0	22.3	34.4	23.8
18:00–18:59	6.1	2.7	1.7	4.7	2.7	3.5	1.3	2.9	1.1	3.9
19:00–19:59	2.0	2.9	1.7	2.4	6.8	4.9	4.0	2.6	1.1	1.9
20:00–20:59	4.1	2.4	0.0	2.6	1.4	2.0	0.0	2.3	0.0	1.9
21:00–21:59	2.0	0.8	0.0	0.9	0.0	0.9	0.0	1.1	1.1	0.0

Table 2.4
13D/A Filings: Submissions After Market Close

The table presents the proportion of disclosures on each weekday submitted between 16:00 and 17:30 ET, for: (1) 13D/A filings disclosing a decrease in holdings in excess of 20%, as a proportion of the original 13D position, but not a complete liquidation of the position, denoted as "Large Sellers"; and (2) all other 13D/A filings, denoted as "All others". The last column provides an estimate of the difference in proportions for each day of the week, with the p-value for a two-tailed test of the null hypothesis that the proportions are equal, based on a normal approximation to the binomial distribution, in parentheses. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Large Sellers		All others		Difference, %
	N	%	*N*	%	
Monday	29	59.2	158	42.2	16.9 (0.025) **
Tuesday	32	55.2	135	39.7	15.5 (0.027) **
Wednesday	42	57.5	148	42.8	14.8 (0.021) **
Thursday	53	70.7	140	40.1	30.6 (0.000) ***
Friday	59	65.6	163	45.0	20.5 (0.000) ***
All weekdays	215	62.3	744	42.0	20.3 (0.000) ***

Table 2.5
Price Response Around Filing Date: Event Study Results

The table presents abnormal returns over various event windows for two subsets of filings, where: $\overline{AR}$ is the aggregated abnormal return on the day in question, $\overline{CAR}$ is the aggregated cumulative abnormal return from day −7 through the day in question, and the p-value is calculated with respect to the null hypothesis that $\overline{CAR}$ is zero. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively. There are $N = 193$ events in the "Monday–Thursday" sample, and $N = 72$ events in the "Friday" sample.

Event Day	Monday–Thursday			Friday		
	$\overline{AR}$, %	$\overline{CAR}$, %	p-value	$\overline{AR}$, %	$\overline{CAR}$, %	p-value
−7	−0.03	−0.03	0.398	−0.02	−0.02	0.399
−6	−0.28	−0.31	0.322	0.14	0.13	0.391
−5	−0.39	−0.70	0.190	−0.40	−0.27	0.375
−4	−0.40	−1.09	0.100	−0.70	−0.97	0.218
−3	−0.09	−1.19	0.109	−1.99	−2.96	0.004 ***
−2	−0.25	−1.44	0.081 *	0.40	−2.56	0.024 **
−1	−0.22	−1.66	0.065 *	−1.07	−3.62	0.003 ***
0	−0.43	−2.10	0.032 **	−0.21	−3.83	0.004 ***
1	0.52	−1.58	0.112	−0.13	−3.96	0.005 ***
2	−0.51	−2.09	0.054 *	0.06	−3.91	0.008 ***
3	−0.09	−2.17	0.055 *	−0.50	−4.40	0.004 ***
4	0.23	−1.95	0.093 *	−0.97	−5.37	0.001 ***
5	0.14	−1.81	0.125	0.88	−4.49	0.007 ***
6	−0.02	−1.83	0.132	0.71	−3.78	0.029 **
7	0.16	−1.68	0.168	−0.40	−4.18	0.020 **
8	0.40	−1.28	0.250	0.95	−3.23	0.075 *
9	0.87	−0.41	0.381	−0.08	−3.30	0.077 *
10	−0.25	−0.66	0.357	0.36	−2.94	0.116
11	0.02	−0.64	0.362	−0.27	−3.22	0.098 *
12	−0.28	−0.91	0.330	−0.15	−3.37	0.093 *
13	−0.06	−0.98	0.324	0.97	−2.40	0.197
14	0.01	−0.97	0.328	−0.30	−2.71	0.170
15	0.00	−0.97	0.330	−0.50	−3.21	0.127
16	0.04	−0.93	0.338	1.08	−2.13	0.246
17	−0.11	−1.04	0.328	0.28	−1.84	0.281
18	−0.26	−1.30	0.296	0.17	−1.67	0.303
19	0.33	−0.97	0.340	−0.83	−2.50	0.220
20	0.44	−0.53	0.381	0.67	−1.84	0.293
21	0.46	−0.08	0.399	0.78	−1.06	0.361
22	−0.36	−0.44	0.387	0.04	−1.02	0.365
23	0.17	−0.26	0.395	0.26	−0.76	0.380
24	−0.23	−0.49	0.385	0.83	0.07	0.399
25	−0.14	−0.63	0.378	0.16	0.23	0.397
26	0.15	−0.48	0.387	0.08	0.31	0.396
27	−0.02	−0.50	0.386	0.24	0.55	0.390
28	−0.10	−0.60	0.381	−0.22	0.33	0.396
29	0.52	−0.08	0.399	0.47	0.80	0.382
30	0.16	0.09	0.399	−0.99	−0.20	0.398

Table 2.6

Price Response Around Filing Date: Test for a Difference in Reactions

The table presents two-tailed tests for differences between large sellers disclosing on Monday through Thursday ("MonThu"), and those reporting on Friday ("Fri"), with respect to price response in the pre-filing window (Panel A) and the post-filing window (Panel B). In the second and third columns, numbers in parentheses are p-values calculated with respect to the null hypothesis that the $\overline{CAR}$ is zero; in the last column, numbers in parentheses are p-values calculated with respect to the null hypothesis that the difference in $\overline{CAR}$ is zero. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively. There are $N = 193$ events in the "MonThu" sample, and $N = 72$ events in the "Fri" sample.

Panel A: Pre-filing window $\overline{CAR}$			
Event Day	MonThu, %	Fri, %	MonThu–Fri, %
−7	−0.03 (0.398)	−0.02 (0.399)	−0.01 (0.987)
−6	−0.31 (0.322)	0.13 (0.391)	−0.43 (0.580)
−5	−0.70 (0.190)	−0.27 (0.375)	−0.43 (0.655)
−4	−1.09 (0.100)	−0.97 (0.218)	−0.13 (0.908)
−3	−1.19 (0.109)	−2.96 (0.004) ***	1.77 (0.152)
−2	−1.44 (0.081) *	−2.56 (0.024) **	1.12 (0.408)
−1	−1.66 (0.065) *	−3.62 (0.003) ***	1.96 (0.180)
0	−2.10 (0.032) **	−3.83 (0.004) ***	1.73 (0.267)

Panel B: Post-filing window $\overline{CAR}$			
Event Day	MonThu, %	Fri, %	MonThu–Fri, %
1	0.52 (0.114)	−0.13 (0.382)	0.65 (0.238)
2	0.01 (0.399)	−0.07 (0.396)	0.09 (0.913)
3	−0.08 (0.395)	−0.57 (0.302)	0.49 (0.605)
4	0.15 (0.389)	−1.54 (0.086) *	1.69 (0.127)
5	0.29 (0.370)	−0.66 (0.319)	0.95 (0.443)
6	0.26 (0.378)	0.05 (0.399)	0.22 (0.873)
7	0.42 (0.355)	−0.35 (0.381)	0.77 (0.598)
8	0.82 (0.271)	0.60 (0.355)	0.22 (0.889)
9	1.69 (0.093) *	0.53 (0.368)	1.16 (0.484)
10	1.44 (0.154)	0.89 (0.326)	0.55 (0.752)
11	1.46 (0.163)	0.61 (0.365)	0.85 (0.643)
12	1.19 (0.233)	0.46 (0.381)	0.73 (0.704)
13	1.12 (0.256)	1.43 (0.267)	−0.31 (0.878)
14	1.13 (0.263)	1.12 (0.316)	0.00 (0.999)
15	1.12 (0.271)	0.63 (0.373)	0.50 (0.815)
16	1.17 (0.270)	1.71 (0.250)	−0.54 (0.807)
17	1.06 (0.294)	1.99 (0.219)	−0.93 (0.684)
18	0.80 (0.339)	2.16 (0.205)	−1.36 (0.562)
19	1.13 (0.293)	1.33 (0.314)	−0.20 (0.933)
20	1.57 (0.227)	2.00 (0.239)	−0.43 (0.861)
21	2.02 (0.163)	2.77 (0.155)	−0.75 (0.765)
22	1.66 (0.224)	2.81 (0.158)	−1.15 (0.656)
23	1.83 (0.204)	3.07 (0.139)	−1.24 (0.639)
24	1.61 (0.243)	3.90 (0.078) *	−2.30 (0.396)
25	1.47 (0.268)	4.06 (0.073) *	−2.59 (0.348)
26	1.62 (0.251)	4.14 (0.073) *	−2.52 (0.371)
27	1.60 (0.258)	4.38 (0.064) *	−2.78 (0.333)
28	1.50 (0.276)	4.16 (0.081) *	−2.66 (0.362)
29	2.02 (0.209)	4.63 (0.060) *	−2.61 (0.380)
30	2.18 (0.192)	3.64 (0.129)	−1.45 (0.631)

Table 2.7
Disclosure of Sequential Sales: All Sellers

Panel A lists the proportion of Friday disclosures for three groups of filings: (1) 13D/A filings disclosing a decrease in holdings, but not a complete liquidation of the position, followed by at least one additional sale within a fixed number of days (specified in the first column), denoted as "Sequential Sales"; (2) 13D/A filings disclosing a decrease in holdings, but not a complete liquidation of the position, with no additional sales within a fixed number of days (also in the first column), denoted as "Isolated Sales"; and (3) all 13D/A disclosures not classified as "Sequential Sales", denoted "Other Filings". Panel B provides an estimate of the difference in proportions of Friday filings across each pair of filing types, with the p-value for a two-tailed test of the null hypothesis that the proportions are equal, based on a normal approximation to the binomial distribution, in parentheses. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

Panel A: Frequency of Friday disclosures for each group of filings

Additional sale(s) occurs within:	Sequential Sales		Isolated Sales		Other Filings	
	N	%	*N*	%	*N*	%
7 days	8	16.3	139	23.2	444	21.5
14 days	15	18.1	132	23.4	437	21.5
21 days	23	20.5	124	23.2	429	21.4
30 days	27	21.3	120	23.1	425	21.4
45 days	33	21.6	114	23.1	419	21.3
60 days	37	22.2	110	22.9	415	21.3

Panel B: Differences in proportions of Friday disclosures across groups

Additional sale(s) occurs within:	Seq − Oth		Iso − Oth		Seq − Iso	
	Difference, %		Difference, %		Difference, %	
7 days	−5.2	(0.384)	1.8	(0.358)	−6.9	(0.267)
14 days	−3.4	(0.456)	1.9	(0.332)	−5.3	(0.279)
21 days	−0.9	(0.827)	1.8	(0.378)	−2.6	(0.544)
30 days	−0.1	(0.977)	1.7	(0.400)	−1.8	(0.661)
45 days	0.2	(0.948)	1.7	(0.404)	−1.5	(0.697)
60 days	0.9	(0.794)	1.6	(0.439)	−0.8	(0.840)

Table 2.8
Disclosure of Sequential Sales: Large Sellers

Panel A lists the proportion of Friday disclosures for three groups of filings: (1) 13D/A filings disclosing a decrease in holdings in excess of 20%, measured as a proportion of initial ownership, but not a complete liquidation of the position, followed by at least one additional sale within a fixed number of days (specified in the first column), denoted as “Sequential Sales”; (2) 13D/A filings disclosing a decrease in holdings in excess of 20%, measured as a proportion of initial ownership, but not a complete liquidation of the position, with no additional sales within a fixed number of days (also in the first column), denoted as “Isolated Sales”; and (3) all 13D/A disclosures not classified as “Sequential Sales”, denoted “Other Filings”. Panel B provides an estimate of the difference in proportions of Friday filings across each pair of filing types, with the p-value for a two-tailed test of the null hypothesis that the proportions are equal, based on a normal approximation to the binomial distribution, in parentheses. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

Panel A: Frequency of Friday disclosures for each group of filings

Additional sale(s) occurs within:	Sequential Sales		Isolated Sales		Other Filings	
	N	%	N	%	N	%
7 days	7	29.2	83	25.9	445	21.3
14 days	13	39.4	77	24.7	439	21.1
21 days	18	40.9	72	23.9	434	20.9
30 days	19	38.8	71	24.0	433	20.9
45 days	20	35.1	70	24.3	432	21.0
60 days	22	34.4	68	24.2	430	21.0

Panel B: Differences in proportions of Friday disclosures across groups

Additional sale(s) occurs within:	Seq − Oth	Iso − Oth	Seq − Iso
	Difference, %	Difference, %	Difference, %
7 days	7.9 (0.348)	4.6 (0.064) *	3.3 (0.722)
14 days	18.3 (0.011) **	3.6 (0.149)	14.7 (0.067) *
21 days	20.0 (0.001) ***	3.0 (0.239)	17.0 (0.017) **
30 days	17.8 (0.003) ***	3.0 (0.233)	14.8 (0.029) **
45 days	14.1 (0.010) **	3.3 (0.197)	10.8 (0.090) *
60 days	13.4 (0.010) ***	3.2 (0.213)	10.2 (0.094) *

Table 2.9

Price Response Around Filing Date: Sequential Sales, Three-Week Cutoff

The table presents two-tailed tests for differences between large sequential sellers—defined using a 21-day cutoff—disclosing on Monday through Thursday ("MonThu"), and those reporting on Friday ("Fri"), with respect to price response in the pre-filing window (Panel A) and the post-filing window (Panel B). In the second and third columns, numbers in parentheses are p-values calculated with respect to the null hypothesis that the $\overline{CAR}$ is zero; in the last column, p-values in parentheses are calculated with respect to the null hypothesis that the difference in $\overline{CAR}$ is zero. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively. There are $N = 22$ events in the "MonThu" sample, and $N = 15$ events in the "Fri" sample.

Panel A: Pre-filing window $\overline{CAR}$ (sequential sales, 21-day cutoff)

Event Day	MonThu, %	Fri, %	MonThu–Fri, %
−7	−0.77 (0.283)	−0.37 (0.352)	−0.40 (0.741)
−6	−1.14 (0.273)	1.51 (0.142)	−2.65 (0.123)
−5	−2.74 (0.092) *	1.13 (0.272)	−3.87 (0.069) *
−4	−4.26 (0.028) **	0.61 (0.367)	−4.87 (0.048) **
−3	−4.02 (0.061) *	1.02 (0.331)	−5.03 (0.067) *
−2	−4.60 (0.051) *	1.44 (0.292)	−6.04 (0.045) **
−1	−4.08 (0.099) *	−0.31 (0.394)	−3.77 (0.238)
0	−5.41 (0.047) **	−0.70 (0.377)	−4.71 (0.170)

Panel B: Post-filing window $\overline{CAR}$ (sequential sales, 21-day cutoff)

Event Day	MonThu, %	Fri, %	MonThu–Fri, %
1	−0.94 (0.237)	−0.72 (0.249)	−0.22 (0.855)
2	−1.22 (0.259)	−0.02 (0.399)	−1.20 (0.479)
3	−0.32 (0.391)	0.44 (0.376)	−0.77 (0.711)
4	−0.83 (0.361)	−0.51 (0.376)	−0.32 (0.895)
5	−0.29 (0.395)	0.98 (0.336)	−1.27 (0.635)
6	−0.35 (0.394)	0.50 (0.384)	−0.84 (0.773)
7	−2.12 (0.274)	−0.32 (0.394)	−1.80 (0.571)
8	2.75 (0.229)	−0.25 (0.396)	3.01 (0.377)
9	6.29 (0.031) **	−0.65 (0.382)	6.94 (0.060) *
10	7.30 (0.018) **	0.60 (0.386)	6.70 (0.083) *
11	6.46 (0.043) **	1.99 (0.288)	4.47 (0.265)
12	6.93 (0.038) **	2.65 (0.236)	4.28 (0.305)
13	5.97 (0.081) *	2.45 (0.263)	3.51 (0.417)
14	6.54 (0.067) *	2.94 (0.228)	3.59 (0.424)
15	5.25 (0.136)	2.73 (0.255)	2.51 (0.588)
16	4.55 (0.187)	2.62 (0.271)	1.92 (0.688)
17	4.28 (0.213)	3.67 (0.195)	0.60 (0.902)
18	3.87 (0.246)	2.42 (0.298)	1.45 (0.776)
19	3.98 (0.245)	1.46 (0.361)	2.52 (0.629)
20	3.61 (0.273)	0.48 (0.395)	3.13 (0.560)
21	3.93 (0.260)	0.85 (0.387)	3.08 (0.575)
22	2.60 (0.333)	0.33 (0.397)	2.27 (0.687)
23	2.19 (0.353)	0.44 (0.396)	1.75 (0.760)
24	2.31 (0.350)	1.03 (0.383)	1.29 (0.826)
25	1.10 (0.388)	0.43 (0.396)	0.67 (0.911)
26	0.94 (0.391)	0.98 (0.386)	−0.03 (0.995)
27	1.32 (0.384)	0.60 (0.394)	0.72 (0.908)
28	2.43 (0.353)	0.83 (0.390)	1.59 (0.801)
29	2.78 (0.341)	1.81 (0.360)	0.97 (0.881)
30	3.14 (0.329)	0.17 (0.399)	2.97 (0.651)

Table 2.10

Price Response Around Filing Date: Sequential Sales, One-Month Cutoff

The table presents two-tailed tests for differences between large sequential sellers—defined using a 30-day cutoff—disclosing on Monday through Thursday ("MonThu"), and those reporting on Friday ("Fri"), with respect to price response in the pre-filing window (Panel A) and the post-filing window (Panel B). In the second and third columns, numbers in parentheses are p-values calculated with respect to the null hypothesis that the $\overline{CAR}$ is zero; in the last column, p-values in parentheses are calculated with respect to the null hypothesis that the difference in $\overline{CAR}$ is zero. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively. There are $N = 26$ events in the "MonThu" sample, and $N = 15$ events in the "Fri" sample.

Panel A: Pre-filing window $\overline{CAR}$ (sequential sales, 30-day cutoff)

Event Day	MonThu, %	Fri, %	MonThu–Fri, %
−7	−0.05 (0.398)	−0.37 (0.352)	0.32 (0.787)
−6	−0.84 (0.324)	1.51 (0.142)	−2.35 (0.168)
−5	−2.05 (0.174)	1.13 (0.272)	−3.18 (0.129)
−4	−3.76 (0.050) **	0.61 (0.367)	−4.37 (0.073) *
−3	−3.60 (0.086) *	1.02 (0.331)	−4.62 (0.089) *
−2	−4.42 (0.059) *	1.44 (0.292)	−5.86 (0.051) *
−1	−3.84 (0.115)	−0.31 (0.394)	−3.53 (0.267)
0	−4.99 (0.064) *	−0.70 (0.377)	−4.29 (0.208)

Panel B: Post-filing window $\overline{CAR}$ (sequential sales, 30-day cutoff)

Event Day	MonThu, %	Fri, %	MonThu–Fri, %
1	−1.05 (0.208)	−0.72 (0.249)	−0.33 (0.783)
2	−1.12 (0.276)	−0.02 (0.399)	−1.10 (0.515)
3	−0.32 (0.391)	0.44 (0.376)	−0.76 (0.712)
4	−0.48 (0.386)	−0.51 (0.376)	0.03 (0.990)
5	−0.11 (0.398)	0.98 (0.336)	−1.08 (0.685)
6	−0.32 (0.395)	0.50 (0.384)	−0.82 (0.779)
7	−1.84 (0.300)	−0.32 (0.394)	−1.52 (0.631)
8	1.98 (0.299)	−0.25 (0.396)	2.23 (0.509)
9	4.92 (0.081) *	−0.65 (0.382)	5.58 (0.125)
10	5.89 (0.051) *	0.60 (0.386)	5.29 (0.166)
11	4.92 (0.109)	1.99 (0.288)	2.93 (0.461)
12	5.34 (0.098) *	2.65 (0.236)	2.69 (0.515)
13	5.29 (0.112)	2.45 (0.263)	2.84 (0.510)
14	6.53 (0.066) *	2.94 (0.228)	3.59 (0.423)
15	4.75 (0.164)	2.73 (0.255)	2.02 (0.662)
16	4.30 (0.202)	2.62 (0.271)	1.68 (0.725)
17	4.48 (0.199)	3.67 (0.195)	0.81 (0.869)
18	4.06 (0.232)	2.42 (0.298)	1.64 (0.745)
19	4.26 (0.227)	1.46 (0.361)	2.81 (0.590)
20	4.04 (0.246)	0.48 (0.395)	3.56 (0.505)
21	4.15 (0.246)	0.85 (0.387)	3.29 (0.547)
22	2.87 (0.320)	0.33 (0.397)	2.53 (0.651)
23	2.19 (0.353)	0.44 (0.396)	1.75 (0.760)
24	2.13 (0.357)	1.03 (0.383)	1.10 (0.850)
25	1.11 (0.388)	0.43 (0.396)	0.67 (0.910)
26	1.13 (0.388)	0.98 (0.386)	0.16 (0.979)
27	1.53 (0.379)	0.60 (0.394)	0.93 (0.881)
28	2.36 (0.355)	0.83 (0.390)	1.53 (0.809)
29	2.70 (0.344)	1.81 (0.360)	0.89 (0.890)
30	2.91 (0.338)	0.17 (0.399)	2.73 (0.676)

Table 2.11
Disclosure of Closed Positions Compared to All Other 13D/A Filings

The table presents weekday distributions of: (1) 13D/A filings disclosing a complete liquidation of shares, denoted as "Closed"; and (2) all other 13D/A filings, denoted as "All others". The last column provides an estimate of the difference in proportions for each day of the week, with the p-value for a two-tailed test of the null hypothesis that the proportions are equal, based on a normal approximation to the binomial distribution, in parentheses. Estimates with significance at the 1%, 5%, and 10% levels are indicated by ***, **, and *, respectively.

	Closed		All others		Difference, %
	N	%	*N*	%	
Monday	19	26.8	404	19.8	7.0 (0.147)
Tuesday	4	5.6	394	19.3	−13.6 (0.004) ***
Wednesday	20	28.2	399	19.5	8.7 (0.072) *
Thursday	13	18.3	411	20.1	−1.8 (0.711)
Friday	15	21.1	437	21.4	−0.2 (0.961)
All	*71*		*2,045*		