NORTHWESTERN UNIVERSITY

Market Imperfections and Asset Prices

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Finance

By

Pengjie Gao

EVANSTON, ILLINOIS

June 2007

ABSTRACT

Market Imperfections and Asset Prices

Pengjie Gao

This dissertation studies two long-standing asset pricing anomalies: "Value and Growth Effects" and "Momentum Effects" via the channel of market imperfections. These market imperfections stem from basic information asymmetry problem, and take forms of contracting problems and less than perfectly competitive information intermediaries.

The first essay (joint with Zhi Da) provides empirical evidence supporting the view that a sharp rise in a firm's default likelihood causes a change in its shareholder clientele. The market imperfection in this essay is institutional investors' investment policy (or investment mandate) constraints, by which they cannot hold stocks falling below certain market capitalization, price, analyst coverage trigger. As institutions decrease their holdings of the firm's share, trading volume and cost increase; the order imbalance measure indicates large selling pressure. The resulting liquidity shock leads to a further concession in the stock price, recovering though, in the subsequent month. Such price recovery explains the first-month abnormal high return earned by stocks with high default likelihood documented in Vassalou and Xing (2004). The abnormal high return is therefore

mostly reward for providing liquidity when it is most needed rather than compensation for bearing a systematic default risk.

The second essay studies the biased information intermediary (sell-side financial analysts) and the momentum effects. Sell-side equity analysts at times have a tendency to herd toward the consensus estimate when making their quarterly earnings forecasts. I argue that such tendency to herd leads to inefficient aggregation of private information and consequently price momentum in stocks. I demonstrate that the Jegadeesh and Titman (1993) price momentum phenomenon is present among stocks only during those time periods when analysts who follow those stocks herd together. I find that the herding tendency is stronger among smaller stocks, growth stocks, and stocks with higher share turnover ratio and more news media coverage. I provide diagnostics suggesting that my findings are distinct from the earnings momentum effects, information uncertainty effects and liquidity risk already documented in the literature.

Acknowledgements

I would like to thank the members of my dissertation committee: Kent Daniel, Peter Klibanoff and Robert Korajczyk for their helpful comments and suggestions, and for always finding time to discuss my work. I am especially grateful to my committee cochairs, Ravi Jagannathan and Ernst Schumburgh for their constant encouragement and endless discussions. I would like to thank Torben Andersen, Janice Eberly, Mike Fishman, Arvind Krishnamurthy, Robert McDonald, Mitchell Petersen, Christopher Polk, Todd Pulvino, Paola Sapienza, Costis Skiadas, Annette Vissing-Jorgensen for their valuable comments and suggestions. Finally, I would like to thank Jim Guo and Patricia Ledesma for extremely responsive data and programming support.

I am grateful to my coauthors, Zhi Da, Joey Engelberg, Rock Gao, Kevin Huang, and Tongshu Ma for many long hours of enlightening discussions. I am fortunate to meet a lot of great friends in Evanston: Ricardo Alonso and Rachel Compos, Zhiguo He and Renna Jiang, Niels Schuehle and Grace Koo, Yong Wang and Xiaohui Liu, Jing Xue and Jianhong Lv, Benjamin Passty, Alex Tetenov, and Weina Zhang. Special thanks go to Ricardo Alonso, who constantly challenges me to think in economics; Zhi Da, who is not only a wonderful friend but also a great coauthor making my duration of doctoral program fun and productive; and Wallace Mok, who helps me to go through the first year of graduate school, coaches my English and proofreads my papers.

I thank my younger brother, Pengqin Gao, who always has faith in me and who is always there to help me. I am grateful to my parents, Jianying Chen and Baohua Gao, and my late grandmother, Fangqin Sun, for their unconditional love and support, especially during the last ten years when I was studying in the States.

To My Parents and Grandmother

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Clientele Change, Liquidity Shock, and the Return on Financially Distressed Stocks

The pricing of financial distress or default risk is one of the fundamental questions in financial economics. Since defaults are more likely to occur in economic downturns, default risk likely contains a nondiversifiable component, thus requiring a risk premium. In a recent paper, Vassalou and Xing (2004) show that stocks more likely to default indeed earn a higher return than otherwise similar stocks during the first month after they enter the highest default-risk portfolio. However, the magnitude of the risk premium appears rather large – the stocks in the highest default risk decile constructed by Vassalou and Xing (2004) earn about 90 basis points more per month than otherwise similar stocks, with an associated monthly Sharpe ratio of around 0.25 during the period from 1970 to 1999. As Hansen and Jagannathan (1991) point out, such high Sharpe ratio can not be easily explained within the "perfect and complete markets" paradigm. As a comparison, Fama and French (1992, 1993) conjecture that book-to-market ratio (BM) captures relative distress risk and therefore the average HML return also reflects a premium for relative distress. However, during the same period from 1970 to 1999, the monthly return on HML is only 35 basis points with a monthly Sharpe ratio of 0.13.

In this paper we argue that a sharp rise in a firm's exposure to default risk, as measured by the Default Likelihood Indicator (DLI) as in Vassalou and Xing (2004), triggers a clientele change in its underlying stockholders. It is well recognized in the literature that the downgrading of a bond can cause a change in the underlying clientele for that bond. For example, when the a bond's rating falls below investment grade, some institutions that hold the bonds are required to sell it. We believe that a similar clientele change occurs for the stock of a firm that experiences a sharp rise in the probability of financial distress. Institutional investors are often restricted to invest in stocks that are liquid, with considerable market capitalizations and stable dividend payouts (c.f. Almazan, Brown, Carlson and Chapman, 2004). A stock is less likely to satisfy these requirements when its default likelihood goes up, a phenomenon that will trigger selling amongst institutional investors who currently hold such a stock. Consistent with this view, we find that mutual funds significantly decrease their holdings of stocks from firms that experience a sharp rise in their default likelihood measures. In addition, significant institutional selling of such stocks are confirmed by a close examination of a proprietary institutional trading dataset. Additional tests suggests that such clientele change is a result rather than a cause of DLI increase.

A sudden change in the clientele for a stock triggers selling by one group of investors with no simultaneous compensatory increase in the demand from ready buyers. This imbalance results in a liquidity shock. In such situations, market makers will have to step in and provide liquidity. A substantial price concession may have to be offered to the market makers for providing immediacy in those situations. The price will bounce back once outside investors recognize the inherent opportunity and move their capital to that stock.[1] However, as Berndt, Douglas, Duffie, Ferguson and Schranzk (2005) point out,

[1]The trading activities of Midway Airlines (ticker = MDW) during July 9 to Aug 10, 1990 provides a stylized example. Midway Airlines experienced a large increase in its default likelihood during July: the

the flow of capital to the new investment opportunity will take some time. As expected, the liquidity risk of the stock changes during such liquidity shock. We find that the initial price concession and subsequent price recovery for the stock also coincides with changes in its liquidity risk as measured by its exposure to the Pástor and Stambaugh (2003) liquidity factor. We argue that such price recovery explains a large part of the high return on financially distressed stocks documented in Vassalou and Xing(2004).

While a stock may experience a sharp change in its exposure to economy-wide, pervasive risk, any such change is likely to persist for a while. In contrast, we find that most of the high returns on stocks that experience sharp increases in their default likelihood measures accrue during the first month following portfolio formation, and little afterwards. Further, various characteristics of those stocks, such as size, book-to-market ratios and default likelihood hardly change from the first- to the second-month since portfolio formation. Such return pattern supports our interpretation that the first-month high return for stocks that experience a large increase in their default likelihood measure should be considered as reward for those who provide liquidity in the market for those stocks when it is most needed. In addition, we also find that: (1) These stocks experience significant increases in their trading volumes, trading costs and realized spreads around portfolio formation dates; (2) Trading in those stocks are more likely to be seller-initiated during the portfolio formation month when prices are depressed, but are more likely to be

DLI increased from 0.21 at the end of June to 0.49 at the end of July. The increase in DLI was mainly driven by two events: a potential downgrade of the company's preferred stock by S&P announced on July 10 and a large quarterly loss of $11 million dollars announced on July 26. The price of the stock was depressed from $7.875 on July 9 to $6.75 on July 30 accompanied with heavy selling (the order imbalance measures were mostly negative). The price then recovered during Aug as more buyers came into the market (the order imbalance measure became positive). In addition, mutual funds, as a group, decrease their holdings of MDW from 2.4% to 0.6% from June to Sep, indicating a clientele change on its shareholders.

buyer-initiated during the month after as prices recover; (3) The stock's exposure to the Pastor and Stambaugh (2003) liquidity factor increases significantly during the portfolio formation month, coinciding with the price concession. The exposure then returns to its normal level during the month after, coinciding with the price recovery. (4) Past return and its interaction with a liquidity measure (Amihud 2002) drive out DLI in predicting the next-month stock (risk adjusted) returns.

All these observations support our view that a sharp rise in the default likelihood measure of a stock triggers a change in its clientele, which generates a liquidity shock and a temporary price concession, and the subsequent price recovery leads to a higher return on the stock. Not surprisingly, a stock that recently experienced a sharp increase in default risk is likely to be a low-priced past loser. However, additional test shows that just being a low-priced past loser is *not sufficient* to generate the main result in the paper. The simultaneous sharp increase in the default likelihood is also needed to trigger a significant clientele change. Therefore, the theme of our paper— clientele change triggered liquidity shock — and its main empirical evidence are quite distinct from the "simple price reversal effect" which has been well-documented in the literature.

Our findings contribute to a growing literature that examines the relationship between default risk and stock returns by zooming in on the role of liquidity shock. Vassalou and Xing (2004) isolate stocks with greater default risk exposure and find that these stocks earn higher returns during the first month after portfolio formation – too high to be explained by the Fama-French three factor model (1993), which seems to indicate the need for default risk as an additional risk factor. Recent studies by Campbell, Hilscher and Szilagyi (2005), Garlappi, Shu and Yan (2005) and Avramov, Chordia, Jostova and

Philipov (2006) find, however, that higher default risk does not necessarily lead to a higher stock return. These two seemingly contradictory sets of results could be reconciled by the liquidity shock we have identified for the financially distressed stock. The short-term liquidity-induced price reversal plays a very little role in the latter papers, as Campbell, Hilscher and Szilagyi (2005) specifically examine annual return while Garlappi, Shu and Yan (2005) and Avramov, Chordia, Jostova and Philipov (2006) effectively exclude the very illiquid stocks. Consistent with the results in the latter papers, we show that the impact of default risk on stock returns is significantly reduced if second-month returns are used in various asset pricing tests. Therefore, insisting on the necessity of a separate aggregate default risk factor in reduced form asset pricing models may be premature. As a result, we also reconcile the seemingly contradictory results in Vassalou and Xing (2004) and Fama and French (1996). After accounting for the 60 bps compensation for liquidity provision,[2] the remaining return premium of about 30 bps (90 bps - 60 bps) documented in Vassalou and Xing (2004) can be loosely interpreted as compensation for default risk. Such return premium is comparable to the average HML return in magnitude and can be fully explained by the three-factor model.

Our findings also add to the literature that analyzes the impact of liquidity shock on asset prices. Related papers include Grossman and Miller (1988), Campbell, Grossman and Wang (1993), Conrad, Hameed and Niden (1994) and most recently by Avramov, Chordia and Goyal (2005) and Coval and Stafford (2005). These papers theoretically and empirically argue that liquidity shocks have large and persistent impact on asset

[2]The magnitude of such compensation is in line with those documented in the previous literature. For example, Keim and Madhavan (1996) (50 to 100 bps as in Figure 1 of their paper), Coval and Stafford (2005) (79 bps as in Table 5 of their paper).

prices, also confirmed by our findings. The empirical challenge, however, is to identify the economic reason underlying such liquidity shock. In other words, why do agents decide to trade a large quantity of certain asset at the same time? Our paper makes a contribution in this regard by providing one such reason: a sharp increase in default risk. When a stock experiences such increase in default risk, financial institutions with binding investment restrictions have to sell the stock immediately, creating a liquidity shock.

The *liquidity shock* we focus on is distinct from the commonly studied *liquidity risk*: the impact of the liquidity shock is usually temporary but the impact of liquidity risk is permanent because it carries a risk premium.[3] The high return on financially distressed stocks is primarily a result of the liquidity shock since it accrues only during the first month after portfolio formation. However, as one would expect, *liquidity shock* and *liquidity risk* are related empirically. We find that a stock does load more on Pastor and Stambaugh's (2003) aggregate liquidity factor during the liquidity shock and the liquidity factor loading (or the liquidity beta) returns to its normal level soon afterwards. Although the change in liquidity beta is perfectly consistent with the price movement around the event, the aggregate liquidity factor may be insignificant in a standard cross sectional asset pricing test using event-window returns. We therefore contribute to the literature on liquidity risk by illustrating the importance of accounting for the time-varying nature of the liquidity risk.

[3]Acharya and Pedersen (2005) decompose the *liquidity risk* premium on individual stocks into four parts: (1) the part due to the level of stock liquidity (c.f. Amihud and Mendelson (1986), Amihud (2002)), (2) the part due to the covariance between the stock return and the aggregate liquidity in the economy (c.f. Pástor and Stambaugh (2003)), (3) the part due to commonality in liquidity among stocks (c.f. Chordia, Roll, and Subrahmanyam (2000) and Hasbrouck and Seppi (2001), and (4) the part due to the covariance between the level of stock liquidity and the market return.

Finally, our findings also highlight the interesting market microstructure dynamics of a stock around the default event. In this respect, our paper is related to recent work by Odders-White and Ready (2005) that shows various market microstructure measures reliably predicting the bond rating changes.[4]

1.1. Brief Review of Default Likelihood Measures

Previous researchers have identified characteristics that are associated with default or financial distress risk. The most direct measure is financial leverage. A long thread of literature on bankruptcy predictions has consistently found that financial leverage is both economically and statistically significant in predicting the likelihood of bankruptcy, which can be viewed as indirect evidence that financial leverage is related to default risk.[5] Both systematic and idiosyncratic risk increases with financial leverage, *ceteris paribus*, and increases in such risk would be associated with an increase in expected return. Bhandari (1988) finds the expected stock returns are indeed positively related to debt-to-equity ratio, even after controlling for beta and size.

B/M is also believed to be associated with default or financial distress risk. According to Fama and French (1992): "A high B/M says that the market judges the prospects of a firm to be poor relative to firms with low B/M. Thus B/M may capture the relative-distress effect." Since log B/M can also be expressed as the difference between log market leverage and log book leverage, Fama and French interpret B/M as an "involuntary leverage effect".

[4]Our paper differs from Odders-White and Ready (2005) in three important aspects at least. First, we are interested in how market participants transact when a stock becomes financially distressed. As a result, we identify the economic cause of a liquidity shock. Second, we focus on explaining stock return patterns after a default event while Odders-White and Ready (2005) focus on pre-event stock returns. Third, we examine market implied default likelihood rather than ratings assigned by rating agencies, which gives us a much larger stock sample.

[5]See Shumway (2001) for a more comprehensive survey on this topic.

Since small firms are more prone to default, size is also believed to be associated with distress as in Chan and Chen (1991). Other researchers use only accounting bankruptcy measures for distress risk, for instance, O-score and Z-score in Dichev (1998).[6]

A common criticism against using accounting measures argues that the accounting information can only be updated at a lower frequency. To accommodate this problem, Vassalou and Xing (2004) estimate a default likelihood indicator (DLI) within the Black and Scholes (1973) and Merton's (1974) framework for each firm as:

$$DLI = N(-DD) = N\left(-\frac{\ln(V_A/X) + (\mu - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}\right), \tag{1.1}$$

where $N(\cdot)$ is the normal distribution's cumulative density function; DD stands for Distance to Default; X and T are the face value and the maturity of the firm's debt, respectively; V_A is the value of the firm's assets; μ and σ_A are the instantaneous drift and volatility of the firm's assets, respectively. V_A, μ and σ_A are estimated iteratively using daily stock returns of the past year. Vassalou and Xing (2004) are also among the first to analyze the relationship between default risk and equity return. They find: (1) both size and B/M effect can be viewed as default effects, (2) stocks with high DLI (usually also with small size and high book-to-market ratio) have very high returns during the first month immediately after the portfolio formation. and (3) the change in aggregate DLI (denoted by dSV) is priced in cross-sectional stock returns even with the presence of Fama and French's three factors. The main advantage to using DLI is that it works from market price information that is updated more frequently than credit rating and other

[6]See Ohlson (1980) and Altman (1968) for O-score and Z-score , respectively.

accounting default measures, so it is potentially a better measure for predicting bankruptcy. Vassalou and Xing (2004) show that DLI predicts actual defaults well. Hillegeist, Keating, Cram and Lundstedt (2004) compare a slightly modified version of DLI against traditional accounting measures: the Z-score and O-score, and find DLI to provide more information on the probability of default than these two accounting measures. Consistent with previous findings, we show that probability of delisting due to performance-related reasons (CRSP delisting code between 400 and 599) increases monotonically with DLI. Our calculation shows that for stocks in the highest DLI decile, about 12% get delisted due to performance-related reasons during the next one year, compared to only 0.4% for stocks in the lowest DLI decile. For this reason, we decide to use DLI as our default risk measure in this paper.

1.2. Returns on Financially Distressed Stocks Beyond the First Month

Vassalou and Xing (2004) sort all stocks into 10 deciles according to DLI at the end of each month from 1970 to 1999 and compute equally-weighted portfolio return for each decile during the first month after portfolio formation. They find that the stocks in the highest default risk decile earn about 90 basis points more per month than otherwise similar stocks. If such large return premium on financially distressed stocks during the first month is indeed due to exposure to a systematic default risk, we would expect it to persist for a while provided that the characteristics of these stocks hardly change.

Following the portfolio construction in Vassalou and Xing (2004), we sort all stocks into ten deciles according to the DLI measures at the end of every month.[7] We then

[7] The leverage ratios of financial firms are usually high due to the nature of their business, which leads to higher DLI measures but do not necessarily reflect high default risk. For this reason, financial firms are often exlcuded as in Garlappi, Shu and Yan (2005). The results in this paper are qualitatively similar

compute the equally-weighted average stock returns in each of the first six months after portfolio formation.[8] As the default likelihood is directly related to actual default and delisting from major exchanges, delisting returns deserve careful handling in our empirical exercise.[9] The results are provided in Panel A of Table 1.1.

Two interesting observations stand out. First, the large return difference between high-DLI and low-DLI stocks during the first month is primarily driven by stocks in the highest DLI decile. These stocks earn 2.10% in the first month, much higher than the rest. Second, the return of the highest DLI portfolio immediately decreases by more than a quarter from 2.10% in the first month to 1.52% in the second month, and stabilizes afterwards. This drop of 58 bps is highly significant (with a t-value above 10), and is five times higher in magnitude than the average change in the rest of the portfolio returns. Panel B of Table 1.1 reports the average Size, B/M and DLI of the 10 DLI-sorted portfolios one month after portfolio formation. The changes in these characteristics within one month are very small in magnitude. For stocks in the highest DLI decile, these changes are all smaller

---

if we exclude financial firms as reported in an earlier version of the paper. As in Vassalou and Xing (2004), we do not exclude penny stocks as such practice, in the context of the current paper, amounts to excluding a large number of financially distressed stocks – the subset of stocks we are most interested in.

[8]We use equally-weighted returns throughout the paper so our results are comparable to those in Vassalou and Xing (2004).

[9]Shumway and Warther (1999) meticulously examine the delisting returns in CRSP and explore their empirical implications with regards to some well-known “anomalies”. They suggest assigning -0.30 and -0.55 to performance related delistings of NYSE and NASDAQ stocks, respectively. These two delisting returns are widely used in subsequent literature. Nevertheless, such numbers are slightly outdated, considering the recent completion of a historical project in delisting returns, as shown in CRSP white paper (2001). We take a different approach. If delisting returns are available from CRSP, we use CRSP delisting returns in our calculation. Otherwise, in line with Shumway and Warther (1999), we recompute the average delisting returns based on the nature of delisting, as identified by the CRSP delisting code. As a further robustness check, we rerun our empirical exercises using the delisting return suggested by Shumway and Warther (1999), or simply assigning the delisting return as -1, and the results are quantitatively similar. This is not surprising given the small delisting probability during the first month after portfolio formation (less than 1.5% even for stocks in the highest DLI-decile).

than 5%. Therefore, the 58 bp drop in return is unlikely explained by changes in risk associated with these stocks.

Since the risk characteristics of a stock do not change significantly over a month, the second-month returns might be better choices for asset pricing tests.[10] We show that the impact of aggregate default risk on stock returns is significantly reduced if second-month returns are used.

If we run a simple time-series regression of the first month return of stocks in the highest DLI decile on the Fama-French three factors, we obtain a significant positive alpha of 64 bp. The results are reported below with t-value in bracket. They are consistent with the asset pricing test results in Vassalou and Xing (2004) and seem to indicate that the return of high default risk is too high to be explained by the standard Fama-French three factors. A separate default risk factor seems to be needed in the reduced form asset pricing model.

$$\begin{aligned} R_{HDLI,1} - r_f &= 0.0064 + 1.13MKT + 1.85SMB + 0.75HML \\ & \quad\ (2.33) \quad (16.64) \qquad (18.97) \qquad (6.91) \end{aligned}$$

If we use second month return instead, we have:

$$\begin{aligned} R_{HDLI,2} - r_f &= 0.0003 + 1.09MKT + 1.79SMB + 0.75HML \\ & \quad\ (0.13) \quad (16.48) \qquad (19.04) \qquad (7.13) \end{aligned}$$

[10] This is also consistent with standard practice in momentum literature. In addition, it helps to reduce the bias introduced by the bid-ask bounce. In fact, it is often the cited reason for skipping a week or a month between portfolio formation and portfolio holding period in momentum literature. For instance, Jegadeesh and Titman (1993) skip a week to avoid "bid-ask spread, price pressure and lagged reaction effects". Similarly, Fama and French (1996) skip a month to "reduce bias from bid-ask bounce".

The intercept term drops to a number indistinguishable from zero while the slope coefficients hardly change, confirming that risk characteristics of the stock did not change by much during the first month after portfolio formation. The 61 bp decrease in the alpha (from 64 bp to 3 bp) is very close to the 58 bp drop in average return from the first to the second month. This decrease is not likely driven by change in risk as both characteristics and the Fama-French three factor loadings hardly change. Once we use the second month return, the return of high default risk stock can be fully explained by the three factors and we do not need an additional default risk factor.[11] The above evidence shows that the first-month abnormal return on financial distressed stocks is unlikely driven by exposure to systematic risk. The next section provides a liquidity-shock-based explanation of such abnormal return.

1.3. Clientele Change and First-month Returns on Financially Distressed Stocks

1.3.1. Characteristics of high-DLI stocks

To compute DLI, Vassalou and Xing (2004) use three economically sensible inputs: V_A/X, μ and σ_A. Empirically, μ is computed as mean of changes in $\ln V_A$ and is closely related

[11]Several alternative tests confirm the results from the time-series regressions. We first conduct the GMM tests on the 10 DLI-sorted portfolios. Using the first-month returns, an aggregate default risk factor, dSV, computed as the changes in the average DLI across all stocks, is significant even with the presence of the Fama-French three factors. The significance of dSV disappears if the second-month returns are used: dSV ceases to provide any additional explanatory power on top of the three factors. Similar results are obtained when we repeat the GMM tests on the 27 portfolios formed by independent triple sorts on DLI, size and book-to-market ratios. Again, dSV becomes insignificant once second-month returns are used. We also verify that the risk characteristics of the stock did not change significantly after one month for the 27 portfolios. The changes in the default risk factor loadings are small. For the highest-DLI stock portfolio, the factor loading decreases from 1.9 to 1.8, but the size of such change is too small to explain the 58 bps drop in return. These results are available from the authors upon request.

to stock returns (ret).[12] V_A/X is closely related to financial leverage ($lev = D/E$), as $V_A/X \simeq 1 + 1/lev$. Finally, σ_A measures the volatility of the assets over the return estimation horizon, which cannot be directly observed but must be estimated using the return and firm asset value; σ_A, then, is also closely related to the stock return volatility.[13] In summary, DLI can be thought of as an "all-in-one" measure, defined as a nonlinear transformation of leverage with two additional variables, i.e., $DLI = f(lev, ret, \sigma_A)$. To better understand DLI, we can look at the relative importance of these three variables. For this purpose, we carry out a variance decomposition exercise similar to those studied in Vuolteenaho (2002). The details are provided in the Appendix A. In a nutshell, the variance decomposition delineates how much the cross-sectional variations of the DLI can be attributed to the cross-sectional variations of the three variables.

Several observations emerge from the variance-decomposition results in Table 1.2. First, financial leverage contributes to approximately 50 percent of the cross-sectional variation of DLI, regardless whether we focus on the whole sample or the subsample of firms with high DLIs. Consistent with prior empirical evidence in Altman (1968) and Shumway (2001), among others, financial leverage is the most salient proxy for default or financial distress risk. Second, Vassalou and Xing (2004) highlights the importance of firm level volatility as a determinant of default risk. We find even though asset volatility contributes modestly (around 20 percent) to the cross-sectional variations in DLI of the overall sample, its contribution in the high DLI subsample (the sample of interests for this paper) is much less. In the top one-third of the sample (as in Panel B) with the highest

[12] To be more precise, $\mu_E - r = \frac{\partial E}{\partial V}\frac{V}{E}(\mu - r)$ where E denotes equity value and μ_E denotes equity return and $\frac{\partial E}{\partial V}$ measures the sensitivity of equity value with respect to the underlying asset value V.

[13] To be precise, $\sigma_E = \frac{\partial E}{\partial V}\sigma_A$ where σ_E measures the stock return volatility.

DLI, it only contributes about 7 percent, while in the top DLI quintile (as in Panel C), it contributes less than 4 percent. Third, past returns contribute the lion's share to the cross-sectional variation of DLI. In the overall sample, it contribute about 17 percent; but in the top one-third and one-fifth of the sample (the high DLI samples), it contributes 32 percent and 34 percent, respectively.

Our variance decomposition exercise shows that past return contributes substantially to high DLI. Since past return is negatively related to DLI, we expect high DLI stocks to be past losers. To confirm this, we compute the equally-weighted average return during the portfolio formation month for each DLI-decile. The results are provided in Table 1.3, Panel A . The results confirm that in general there is indeed a negative relationship between DLI and past return. In particular, stocks in the highest DLI decile earn an average return of -3.58% during the portfolio formation month; They are clearly recent losers. They also earn the highest return (2.1%) during the first month after portfolio formation. This return pattern is consistent with the short-term return reversal previously documented in the literature.[14]

In fact, short-term return reversals on the highest-DLI stocks are mainly driven by a subset of New High DLI stocks that only recently entered the highest DLI decile. Panel B of Table 1.3 displays the probability transition matrix of a stock moving from DLI decile i during the month immediately prior to the portfolio formation month $(t-1)$, to DLI decile j during the portfolio formation month (t). All probabilities in the same row should therefore add up to 1. As shown in the last column, about 17% of the stocks in the highest-DLI portfolio migrated from other deciles and are associated with larger increases

[14]Stocks with the lowest DLI also demonstrate some degree of return reversal: they earn a high return of 2.48% during the portfolio formation month and a low return of 1.13% during the following month.

in DLI. We label these stocks New High DLI stocks and the remaining 83% of the stocks in the highest-DLI portfolio "old" high DLI stocks. The New High DLI stocks display more pronounced return reversal patterns. On average, they suffer a larger return loss during the portfolio formation month (as in the last column of Panel C) and have higher positive returns during the month after (as in the last column of Panel D).[15] Panel E reports the corresponding Fama-French three-factor risk-adjusted returns during the first month after portfolio formation ($t + 1$). Again, only stocks that experience sharp increase in default risk (stocks moving from DLI decile 9 to decile 10 or from DLI decile 8 to decile 9) have significant positive risk-adjusted first-month returns.[16] Apparently, the first-month high return on the highest-DLI portfolio is mainly driven by the New High DLI stocks as the risk-adjusted return on "old" high DLI stocks is not significantly positive. In addition, abnormal return is only present during the first-month after portfolio formation. For instance, the Fama-French three factor alpha is as high as 138 bps (with a t-value of 5.1) for the New High DLI stocks if first-month returns are used. The alpha drops to -23 bps (with a t-value of -0.79) if the second-month returns are used.

Panel A of Table 1.3 also documents various characteristics of the 10 DLI-sorted portfolios. Consistent with Vassalou and Xing (2004), the highest-DLI stocks are associated

[15]A notable exception is a stock that migrates from decile 1 to decile 10 within a month. However, such stocks are too scarce (28 out of almost $900,000$ stock/month observation) to let us draw any reliable inference.

[16]Stocks that experience sharp decrease in default risk (stocks moving DLI decile 10 to decile 9 or from DLI decile 9 to decile 8) exhibit symmetric return reversals: they are past winners during portfolio formation month but significantly under-perform during the first month after portfolio formation (the risk-adjusted returns are significantly negative). Such return pattern can be explained a similar clientele change. Institutional investors would like to hold a financially distressed stock as the optimal portfolio decision rule suggests but they cannot because of various investment restrictions. Therefore, as a stock's default risk decreases sharply, investment restrictions become non-binding and institutional investors start buying the stock. Such buying pressure pushes up the stock price during portfolio formation and leads to lower return during the first month after. We thank Anthony Lynch for pointing out this explanation.

with the smallest size and highest book-to-market ratios. Not surprisingly, high DLI stocks also trade at low prices. In fact, both mean and median price decreases monotonically with DLI. The highest DLI stocks trade at a mean of \$3.58 and a median of only \$2.37.[17] The low trading price makes the percentage transaction cost much higher for financially distressed stocks, thus making them more illiquid at the same time. We consider the "illiquidity" measure suggested by Amihud (2002):

$$Amihud_t = \frac{1}{T} \sum_{d=1}^{T} \frac{|R_{i,t-d}|}{Vol_{i,t-d}}. \tag{1.2}$$

We average the daily absolute value of the ratio between return and dollar trading volume of individual stocks during the portfolio formation month t to get the Amihud measure for month t – $Amihud_t$. [18] The illiquidity measures of individual stocks are then equally-weighted to obtain the illiquidity measure at the portfolio level. Clearly, Amihud's illiquidity measures increase almost monotonically with the DLI.[19]

[17]One common practice in empirical asset pricing studies is to exclude penny stocks in light of liquidity related concerns. However, this practice, in the context of the current paper, amounts to excluding a large number of financially distressed stocks – the subset of stocks we are most interested in. Therefore, as in Vassalou and Xing (2004), we decide not to apply any price filter. Instead we explicitly examine and control for the liquidity effects associated with these stocks. If we exclude stocks traded less than 5 dollars, the highest DLI stocks in the remaining sample do not earn significantly higher returns even during the first month after portfolio formation, consistent with the evidence reported in Garlappi, Shu and Yan (2005).

[18]In order to construct the Amihud measure, we use the filtering rules suggested by Amihud (2002), except that we do not exclude NASDAQ stocks and stocks traded at less than five dollars. In particular, we require that individual stocks must be traded on the stock exchanges for at least 200 days. Furthermore, to minimize the influence of special liquidity provisions from the market makers during the IPO process (see Ellis, Michaely and O'Hara, 2000), we exclude the first 250 observations when a firm first enters CRSP in our sample. The Amihud measures for NASDAQ stocks are likely to be underestimated due to "double countings" in their reported trading volumes. We verify the positive relation between the Amihud measure and DLI in a subsample of only NYSE/AMEX stocks.

[19]The highest-DLI stocks are also more illiquid according various market-microstructue-based measures as discussed in the later subsections and Table 1.7.

In addition, Panel A of Table 1.3 reports the average idiosyncratic risk measures for stocks in 10 DLI-sorted deciles. For each month and each stock, we regress the daily stock excess returns on the Fama-French three factors over the past six months and take the $1-R^2$ (where R^2 is the adjusted-R^2) as a measure of firm-level idiosyncratic risk. Clearly, the idiosyncratic risk measure increases monotonically with DLI. In particular, for stocks with the highest DLI, nearly 97% of the total risk is idiosyncratic in nature. Finally, we show that high-DLI stocks receive little wall street coverage. As a proxy for Wall Street research coverage, for each stock each month, we check whether analyst earnings forecast is made for the firm's announced past quarter earning and, if so, compute the number of unique analysts. The earning forecast data is obtained from I/B/E/S from 1984 to 1999. For each of the 10 DLI sorted portfolio, we report the average percentage of stocks receiving analyst coverage and the average number of analysts for the stocks receiving coverage at all in Panel A of Table 1.3. As expected, both coverage measures decrease with DLI. Amongst stocks in the lowest-DLI decile, 74% receive analyst coverage – 5.4 analysts on average following each stock, if the stock receives analyst coverage at all. In sharp contrast, amongst stocks in the highest-DLI decile, only 20% receive analyst coverage and there are only 2.5 analysts per stock, if the stock receives analyst coverage at all.

1.3.2. Institutional selling pressure

In summary, the highest-DLI stocks are characterized by small market capitalization, high book-to-market ratio, high idiosyncratic risk, low trading price, low level of liquidity and low Wall Street coverage. Institutional investors such as pension funds and mutual funds are often restricted to invest in stocks that are liquid, issued by high-quality companies, with considerable market capitalizations, low idiosyncratic risk and stable dividend payouts (c.f. Almazan, Brown, Carlson and Chapman, 2004). Table 1.4 lists a few examples of such restrictions by institutional investors. A financially distressed stock will unlikely satisfy these restrictions; it is not surprising then to observe a clientele change for these stock as the institutional investors sell it from their current holding.

1.3.2.1. Selling pressure from mutual funds. Institutional investors may include mutual funds, pension funds and hedge funds, among others. We decide to focus on mutual funds because they constitute a relatively homogenous group of investors and have regular disclosures as required by the Security and Exchange Commission (SEC).[20] It turns out our conjecture about the clientele change is true at least for mutual funds as a group. Mutual funds are likely to be a group of investors facing many potential investment constraints. For example, there is anecdotal evidence that a typical mutual fund in general avoids low priced stocks so as not to be looked as "speculative" or "imprudent".[21,22] For example, between 1980 to 2005, in the sample of stocks held by all mutual funds and which can be matched with CRSP monthly stock file, merely 3.73 percent of stocks are priced less than 5 dollar as of reporting date while 90.38 percent of stocks are priced more

[20]We also obtain qualitatively similar results using CDA/Spectrum Institutional 13F Stock Holdings and Transactions database, where the quarterly transactions and holdings by institutional investors including mutual funds, banks, insurance companies, pension funds and endowment funds are recorded.

[21]Mutual funds may "window dress", i.e., they sell recent losers before reporting their holdings (c.f. Haugen and Lakonishok (1988)). This could be another reason why increase in financial distress could trigger a clientele change and selling by mutual funds, as financially distressed stocks are likely to be recent losers.

[22]The eventual delisting may be very costly to the stockholders and SEC rules preclude most institutions from holding unlisted shares (cf. Macey, O'Hara and Pompilio, 2004). In addition, liquidity tend to dry up when delisted stocks are later on traded in the OTC Bulletin Board and/or the Pink Sheets(cf. Harris, Panchapagesan and Werner, 2004). For the above reasosns, some institutions may want to sell the stocks even before the eventual delisting.

than 10 dollars. We choose to focus on mutual funds as a clientele and we infer their buy and sell decisions by looking at the aggregate mutual fund holdings and holding changes when stocks become financially distressed.

The mutual fund holding data come from the CDA/Spectrum mutual fund holding database, which collects the holding information from the N30-D filings to the SEC. A detailed description of the database can be found in Wermers (1999). As our mutual fund holding database only starts at 1980, we only consider the sample from 1980 to 1999. Although typically stocks are likely to be held by a large number of mutual funds, there are number of stocks which are only held by one or two mutual funds recorded by the CDA/Spectrum database. A possible explanation for this observation is that small holdings are exempted from reporting by SEC regulations, giving us a lower-end truncated sample.[23] Therefore, it is likely the number of mutual fund shareholders are under-stated according to CDA/Spectrum but the likely impact should be relatively small. Without further assumptions, it is not entirely clear how such reporting practice may influence the inference of current empirical study. To assess such bias, we further sort the stocks into three groups based on the breadth of ownership as a robustness check: Low refers to ones for which the underlying shareholders is less than or equal to 2; Medium refers to ones for which the underlying shareholders between 3 and 7 (inclusive); and High refers to ones for which the underlying shareholders greater than or equal to 8. These break points roughly match the 33 percentile and 67 percentile of underlying mutual fund shareholders across all stocks and all years in our sample. We report the statistics from the full sample (1980

[23]For example, N30-D form filing guideline states “A Manager may omit holdings otherwise reportable if the Manager holds, on the period end date, fewer than 10,000 shares (or less than $200,000 principal amount in the case of convertible debt securities) and less than $200,000 aggregate fair market value (and option holdings to purchase only such amounts).”

- 1999), and also two subsamples (1980 - 1989 and 1990 - 1999) to ensure that the results are not driven by later period when the number of mutual funds dramatically increases. A final caveat is in order. Because we only look at the aggregate mutual fund holdings and holding changes in the event of stocks' financial distress, we cannot say much about intra-fund flows of share holdings.

At any quarter, we sum across the reported number of shares held by individual mutual funds and obtain the aggregate holdings of mutual funds. We examine two aspects of the aggregate mutual fund holdings and holding changes of the financially distressed stocks. We first investigate the aggregate mutual fund holdings and holding changes of *all* high DLI stocks. At a given quarter Q, we identify all stocks which fall into the highest DLI decile ranking during any month of the current quarter and record the aggregate mutual fund holdings ($Holding_{i,Q}$). Then we track all high DLI stocks' aggregate mutual fund holdings during the preceding quarter ($Holding_{i,Q-1}$). The aggregate holding change ($\Delta Holding_i$) is defined as

$$\Delta Holding_i = Holding_{i,Q} - Holding_{i,Q-1} \tag{1.3}$$

and we conjecture that mutual funds on average decrease their holdings of the stock ($\Delta Holding_i < 0$) for high DLI stocks if mutual funds on average avoid holding financially distressed stocks.

We also examine the aggregate mutual fund holdings and holding changes of *new* high DLI stocks. That is, at a given quarter Q, we only identify stocks which were not in the highest DLI decile in *all* months during the preceding quarter, but recently migrated into high DLI decile during *any* month in current quarter. We compare the mutual fund holdings before ($Holding_{i,Q-1}$) and after ($Holding_{i,Q}$) the stocks become financially distressed in current quarter, and compute the aggregate mutual fund holding changes ($\Delta Holding_i$) as

$$\Delta Holding_i = Holding_{i,Q} - Holding_{i,Q-1} \tag{1.4}$$

We also conjecture that the mutual funds on average decrease their holdings of the stock ($\Delta Holding_i < 0$) if the stock becomes financially distressed. In addition, we expect the holding decreases to be sharper for *new* high DLI stocks if the clientele change is triggered by a sudden increase in financial distress.

The results presented in Table 1.5 consistently supports our conjecture that when stocks becomes financially distressed, there is a change of clientele, as proxied by mutual fund aggregate ownership, across all sample periods and all levels of the breadth of ownership. On average, mutual funds avoid holding high DLI stocks. In the full sample period, mutual funds decrease their holdings of *all* high DLI stocks by 0.67% of all shares outstanding on average within a quarter; and for *new* high DLI stocks, mutual funds decrease holdings by 0.95% within one quarter.[24] The decrease of holdings is particularly pronounced for high breadth of ownership stocks. In the full sample period, mutual fund decreases holdings of all high DLI stocks with high number of ownerships by 1.87% of all shares outstanding on average within a quarter; and for *new* high DLI stocks with high breadth of ownership, mutual funds decrease holdings by 2.36% within one quarter. All

[24]The mutual fund holding change does not differ significantly across different calendar quarters. For *all* high DLI stocks, the mutual fund change is -0.6%, -0.58%, -0.72% and -0.76% during calendar quarter 1 to 4. For *new* high DLI stocks, the mutual fund change is -1%, -0.7%, -1.1% and -1.0% during calendar quarter 1 to 4. Therefore, the mutual fund holding change result is unlikely to be driven primarily by large year-end selling for tax reasons as documented by Branch (1977).

these reported changes are statistically significant at 1 percent significance level. We also verify that the decrease in mutual fund holding mostly occurs during the quarter when the stock becomes financially distressed (see Panel C of Table 1.5). For *all* high DLI stocks, the absolute quarterly mutual fund holding change is below 0.11% during each of the four quarters immediately following the event quarter (Q). For *new* high DLI stocks, although there are still significant decrease in mutual fund holding during the first two quarters immediately following the event quarter (Q), the magnitude of such decrease is much smaller (0.13%) as compared to the decrease during the event quarter (0.95%). The result on mutual fund selling is not driven by a few outliers. we plot the histogram of changes in individual mutual fund holdings for high DLI stocks. Specifically, for each stock i, mutual fund j, at quarter Q, we compute the holding change $\Delta Holding_{i,j,Q}$ as:

$$\Delta Holding_i = Holding_{i,j,Q} - Holding_{i,j,Q-1} \tag{1.5}$$

and examine the distribution of all the holding changes $\Delta Holding_i$. It turns out that more than 73% of the individual mutual fund change is negative, indicating heavy selling pressure.

1.3.2.2. Evidence of institutional selling pressure at a higher frequency. Given the quarterly mutual fund holding reporting frequency, we cannot rule out the possibility that mutual fund holding changes actually occur during the month prior to the increase in DLI. It would be better to examine the institutional trading activities during the same month when the stock experiences a sharp increase in DLI. This becomes possible with the help of a proprietary institutional trading dataset provided by the Plexus Group, a consulting firm for institutional investors that monitors the cost of institutional trading. Plexus Group's customers consist of over 200 financial institutions that collectively transact over $4.5 trillion in equity trading volume prior to the acquisition by ITG, Inc. The Plexus group data have been used by Keim and Madhavan (1995) and Conrad, Johnson and Wahal (2003) among others.[25] The Plexus group dataset examined in this section is a combination of the one used by Keim and Madhavan (1995) (which covers from Q2 of 1991 to Q1 of 1993) and the one used by Conrad, Johnson and Wahal (2003) (with the coverage from Q1 of 1996 to Q1 of 1998). The dataset records the details (time, size, buy/sell indicator, type of the order among others) of every institutional order for all the institutions that Plexus Group monitors. It also records when and how many orders actually get executed. Therefore, for every stock in our sample during portfolio formation month, we are able to compute the aggregate net buy/sell orders (as percentage of total number of shares outstanding) submitted by institutions and the actual aggregate shares bought/sold (again as percentage of total number of shares outstanding) by institutions at monthly frequency. We can then average these two institutional trading measures first across all stocks at portfolio level and then across time. The results for the 10 DLI deciles and the portfolio of New High DLI stocks are presented in Table 1.6. Though we have made a refined and precise measurement of institutional trading, the trade-off for using the Plexus Group dataset is a short sampling period and the fact that institutions monitored by Plexus group is only a subset of the universe of all institutions.[26]

[25] A detailed description and summary statistics of the Plexus Group data can be found in Conrad, Johnson and Wahal (2003) for example.

[26] By early 2003, Plexus Group analyzed 25% of exchange traded volume worldwide. Early year coverage of Plexus Group data is significantly less in total volumes, but still substantial. Given said, we believe our sample is representative of US equity institutional transactions.

Table 1.6 confirms a significant selling pressure for a stock during the month when the stock's DLI increases. Panel A presents the result for the full Plexus Group dataset. A negative number indicates net selling. For both all high DLI stocks and New High DLI stocks, the institutions submit significantly more sell orders and, on average, sold them. Since the coverage of Plexus Group dataset is significantly smaller during the first sub-sample (from Q2 of 1991 to Q1 of 1993), the institutional trading measures could be considerably noisy especially for New High DLI stocks. For example, the average number of New High DLI stocks with Plexus Group coverage comes out at only 2 for the first sub-sample. The coverage of Plexus Group dataset improves significantly during the second sub-sample (Q1 of 1996 to Q1 of 1998). For example, the average number of New High DLI stocks with Plexus Group coverage is 18 during the second sub-sample. For this reason, we also report the results during the second sub-sample separately in Panel B. The institutional trading measures during the second sub-sample, arguably less noisy, are qualitatively similar to those in the full sample. For both all high DLI stocks and New High DLI stocks, there is significant selling pressure during the portfolio formation month. In addition, the selling pressure is more significant for New High DLI stocks as we would expect.

1.3.3. Lack of ready buyers

The selling of financially distressed stocks by institutional investors such as mutual funds is unlikely to be absorbed by ready buyers without moving the price. The market makers, afraid of the selling being information-driven, will only want to buy the stock with price concession. Outside investors are unlikely to move in their capital immediately as argued by Berndt, Douglas, Duffie, Ferguson and Schranzk (2005). It takes time and human capital for an investor to identify a profitable opportunity and then mobilize capital (capital immobility).[27] We think this is especially true for financially distressed stocks. The success and failure of distressed securities investing depend on the investor's efficiency and effectiveness in uncovering and analyzing all of the variables specific to the distressed company. The investor "will not only know everything about the company and its financials but will have studied the creditors involved in the reorganization as well: their numbers, their willingness to compromise, and the complexity of their claims help indicate how long the reorganization will last, what the asset distributions will be, and whether the expected returns are worth the wait".[28] Gathering and analyzing such firm specific information is a daunting task and very time consuming, requiring a large amount of human capital. The absence of Wall Street research coverage on distressed firms makes this task even harder.[29]

When there is large selling pressure and lack of immediate ready buyers, the stock price will be temporarily depressed. The price concession may attract new buyers including arbitrageurs to enter the market and the price will soon recover. Pástor and Stambaugh (2003) focus on liquidity shocks that play out within the span of a day. Keim and Madhavan (1996) does this as well, showing that the price impact of a block sell order

[27]Consistent with the capital immobility argument, Duarte, Lonstaff and Yu (2005) find that the fixed-income arbitrage strategies requiring more "intellectual capital" to implement tend to produce significant risk-adjusted returns and the risk-adjusted excess returns from these strategies are related to capital flows into fixed-income arbitrage hedge funds.

[28]See "Distressed Securities Investing" by Dion Friedland, Chairman of Magnum Funds.

[29]"The lack of Wall Street coverage is due to the fact investment banks tend not to view companies emerging from bankruptcy as potential clients. Further, these companies are tainted in general by the financial distress and thus do not make it onto the list of companies to which Wall Street investment banks allocate expensive research resources..." – "Distressed Securities Investing" by Dion Friedland, Chairman of Magnum Funds.

lasts on average for just one day. To examine the duration of liquidity shock for financially distressed stocks, we trace out the first 20 daily returns after portfolio formation for stocks in the highest DLI portfolio. We plots these daily returns. Consistent with Pástor and Stambaugh (2003) and Keim and Madhavan (1996), we observe a strong first day return reversal for financially distressed stocks. However, the above average return lasts until the second week after portfolio formation, which indicates a persistence in the liquidity shock.

Panel A of Table 1.3 show that financial distressed stocks are usually penny stocks associated with very high idiosyncratic risks and little Wall Street coverage. These stock characteristics contribute to the persistence of the liquidity shock for financially distressed stocks. From the perspective of the market maker, higher idiosyncratic risk means larger amount of nondiversifiable risk in his stock inventory. In response, the market maker is less willing to provide liquidity temporarily as predicted by Spiegel and Subrahmanyam (1995). Our idiosyncratic risk measure can also be interpreted as a proxy related to the proportion of private information (c.f. Durnev, Morck, Yeung, and Zarowin (2003)). It follows that the high idiosyncratic risk measure in the high DLI portfolio indicate that a large fraction of the information is private in nature.[30] Moreover, the difficulty of collecting and analyzing information specific to distressed stock results in a higher degree of information uncertainty. Market makers, in order to protect themselves from information asymmetry in such a uncertain environment, will impose higher trading costs

[30]Given this interpretation, it is easy to understand why the liquidity shock is particularly pronounced among the high DLI portfolios as private information is usually associated with larger price impact of trade as in Kyle (1985). Bessembinder, Chan, and Seguin (1996) also provide some supporting evidence. They find firm-specific information to have the largest proportional effect on the volume of small firms, which is consistent with the increased turnover we documented for financially distressed stocks.

for the distressed stocks over a longer period of time, as argued in Sadka and Scherbina (2006). From the perspective of an arbitrageur, higher idiosyncratic risk makes it difficult to locate similar stocks to short in the arbitrage portfolio as argued in Wurgler and Zhuravskaya (2002). This difficulty, together with larger percentage transaction costs associated with low-priced financial distressed stocks, keep risk-averse arbitrageurs from investing immediately after the stock becomes distressed, as argued in the "limits-to-arbitrage" literature (c.f. Shleifer and Vishny,1997) and price recovery takes longer.[31]

To summarize the findings so far, a sharp increase in a firms' financial distress risk is likely to trigger a clientele change of its stockholders. Selling off amongst existing institutional investors such as mutual funds, which is unlikely absorbed by ready buyers, generates a liquidity shock. The subsequently temporarily depressed stock price will induce the market maker to step in and take the other side. The liquidity will improve after a while and the prices will bounce back, as outside investors recognize the opportunity and gradually move their capital to the stock. In the next subsection, we examine the trading volume, trading cost, order imbalance and level of liquidity during such time, providing additional evidence to support the presence of liquidity shock.

1.3.4. Changes in liquidity-related characteristics during the liquidity shock

Due to the liquidity shock associated with a financially distressed stock, an investor wishing to sell a significant quantity of it will suffer a price concession, and conversely, an investor ready to buy it (therefore provide liquidity) will be rewarded by the later price recovery. Such liquidity shock has been discussed in the model of Grossman and Miller

[31]Consistent with the limits-to-arbitrage arguement, we show that among "new" high DLI stocks, those with higher arbitrage risk measures (Wurgler and Zhuravskaya, 2002) exhibit larger return reversals.

(1988) and Campbell, Grossman, and Wang (1993). One implication of the model is that large trade by liquidity investors leads to a temporary divergence between price and fundamental value. This implies that price concessions accompanied by high volume will tend to be reversed. Empirically, Conrad, Hameed, and Niden (1994) report that stocks with high trading activity are likely to experience short-term return reversal and stocks with low trading activity short-term return continuation. We examine the trading activity of financially distressed during the liquidity shock and document a similar pattern. Panel A of Table 1.7 compares the trading volume for stocks in various DLI deciles during three two-month-periods: (1) the two months prior to the portfolio formation month ([-2,-1]); (2) the portfolio formation month and the first month after portfolio formation ([0,1]); (3) the second and third month after portfolio formation ([2,3]). The trading volumes are adjusted for changes in the total number of shares outstanding. Finally, all trading volumes are normalized by the trading volume during the two months prior to the portfolio formation month ([-2,-1]). New High DLI stocks are stocks which have just recently entered the highest-DLI decile during the portfolio formation month. Although the normalized trading volumes during month ([0,1]) are in general decreasing in DLI, this pattern reversed for the highest DLI-decile: we observe an increase in trading for stocks in the highest-DLI decile around the liquidity shock. This pattern is mainly driven by New High DLI stocks. For this subset of stocks that have recently become financially distressed, we observe a significant increase in trading activity only around the liquidity shock, and not afterwards, consistent with the implication of the model by Campbell, Grossman, and Wang (1993).

Financially distressed stocks also experience a large increase in trading cost during the liquidity shock. We measure the trading cost using the percentage bid-ask spread, defined as the ratio between the quoted bid-ask spread and the midpoint of the quoted bid and quoted ask. The percentage bid-ask spread is computed using intraday quote data from TAQ (after 1993) and ISSM (before 1993). The sampling period for NYSE stocks is from 1983 to 1999 and the sampling period for NASDAQ stocks is from 1987 to 1999. The average spreads are also reported in Panel A of Table 1.7. As we expect, the spread measure increases monotonically with DLI, verifying that financially distressed stocks are more costly to trade. More interestingly, while the trading cost measure hardly changes during the portfolio formation month for stocks in DLI decile 1 to 9, it increases significantly for the financially distressed stocks in DLI-decile 10. Again, such increase in trading cost is mainly driven by New High DLI stocks whose percentage bid-ask spread increases by more than 1% with an associated t-value above 10. This increase is not surprising given the fact that New High DLI stocks are recent losers.

If heavy selling by institutional investors leads to price concession and subsequent buying by outside investors leads to later price recovery, we would expect more sell-initiated trades during portfolio formation month and more buyer-initiated trades during the month after formation for financially distressed stocks. This is exactly what we find using order imbalance measures developed in Chordia, Roll and Subrahmanyam (2002). The time series of the order imbalance measures start from 1988 and end in 1998. $OIBSH1_t$ is the buyer-initiated shares purchased less than the seller-initiated shares sold on day t. $OIBSH2_t$ is $OIBSH1_t$ scaled by the total number of shares traded on day t. We average both variables first within each month and then within each DLI-sorted portfolio to get monthly order imbalance measures for each portfolio. The results are reported in Panel B of Table 1.7. For stocks in the highest-DLI decile, $OIBSH1$ is negative during portfolio formation month which means more trades are seller-initiated, and $OIBSH1$ is positive during the month after formation which means more trades are buyer-initiated. The change in $OIBSH1$ is positive and significant. In addition, across all DLI-sorted deciles during the formation month, $OIBSH1$ is only negative in the highest-DLI decile. We also observe significantly more buyer-initiated trades after portfolio formation for stocks in the highest-DLI decile with the relative order imbalance measure ($OIBSH2$). Finally, we show that this change in order imbalance is more pronounced for New High DLI stocks only recently entering the highest-DLI decile during the portfolio formation month. Changes in both order imbalance measures are more positive and significant.

Panel C reports the average realized (half) spreads (scaled by traded price) for each DLI decile and the portfolio of New High DLI stocks around portfolio formation month. The realized spread is originally developed by Huang and Stoll (1996) as a direct measure of what the liquidity supplier actually earn. The realized spread is computed using intraday trade and quote data from1983 to 1999.[32] If high-return on the high DLI stocks is related to compensation for liquidity provision, we would expect the average realized spread for these stocks to be much higher around the liquidity shock. Indeed, the average realized spread for high DLI stocks is higher around portfolio formation (month = 0 and 1) and such pattern is again driven by New High DLI stocks. The average realized spread for New High DLI stocks increases significantly during the portfolio formation month when there is a liquidity shock, reflecting an increased compensation for liquidity

[32]See Huang and Stoll (1996) for detailed estimation procedure. The time horizon used for the estimation is 30 minutes to account for infrequent tradings.

provision required by liquidity suppliers. The higher realized spread persists during the first month after portfolio formation and drops to its normal level.

In addition, we expect the liquidity risk of a stock to fluctuate around the liquidity shock. We measure the stock liquidity risk using the liquidity beta proposed by Pástor and Stambaugh (2003). The liquidity beta measures the exposure of the stock to an aggregate economywide liquidity factor. Specifically, the liquidity beta n months after portfolio formation for portfolio i is defined as the slope coefficient (β_i^n) in the following regression:

$$r_{i,t}^n = \alpha_i^n + \beta_i^n L_t + \beta_{i,M}^n MKT_t + \beta_{i,S}^n SMB_t + \beta_{i,H}^n HML_t + \varepsilon_{i,t},$$

where $r_{i,t}^n$ is the excess return nth month after portfolio formation; L_t is the innovation in the aggregate liquidity factor defined by Pástor and Stambaugh (2003); and MKT, SMB and HML are the Fama-French three factors. We examine four liquidity betas: (1) the pre-formation liquidity beta which is the average liquidity betas during the three months prior to the portfolio formation month (month $[-3, -1]$); (2) the liquidity beta during the portfolio formation month (month 0); the liquidity beta during the first month after the portfolio formation month (month 1); and (4) the post-formation liquidity beta which is average liquidity betas during the second to fourth month after portfolio formation (month $[2, 4]$).

We plots the four liquidity betas for High DLI stocks (all stocks in the highest-DLI decile), New High DLI stocks (subset of High DLI stocks that only recently entered the highest-DLI decile during the portfolio formation month) and Old High DLI stocks (the remaining High DLI stocks that also belong to the highest-DLI decile during the portfolio formation month).[33] For High DLI and New High DLI stocks, their liquidity betas display an inverse-V shape around portfolio formation. The liquidity betas increase significantly during the portfolio formation month, which indicates a drop in stock liquidity risk, coinciding with the price concession. The liquidity betas then drop significantly during the first month after portfolio formation and return to their normal levels thereafter. The decreases in liquidity betas indicate an improvement in stock liquidity risk, coinciding with the price recovery. As expected, the inverse-V shape is more pronounced for New High DLI stocks. In contrast, the liquidity betas of Old High DLI stocks do not vary significantly around portfolio formation. Panel D of Table 1.7 reports the four liquidity betas for all 10 DLI-sorted portfolios as well as the New DLI stocks. It also reports in the changes in liquidity betas from period to period. The t-values associated with these changes are computed using the Newey-West standard error estimators with three lags. Across all 11 portfolios, we observe statistically significant fluctuations in liquidity betas around portfolio formation only for the High DLI stocks and the New High DLI stocks.

That stock price reversal coincides with changes in liquidity beta is consistent with the theoretical model and empirical findings by Pástor and Stambaugh (2003). Since liquidity beta carries a positive risk premium, when the liquidity beta of a stock increases, the stock becomes more risky, *ceteris paribus*, and its discount rate goes up, resulting in a price drop. Conversely, as the liquidity beta later drops, the discount rate also decreases and the stock price will recover. The resulting high return on the stock during the first month after portfolio formation is therefore consistent with the dynamic decrease in the liquidity beta. However, an unconditional asset pricing test which ignores the dynamic

[33]This figure is not presented to save space.

nature of the liquidity risk, is likely to produce spurious results. As we can see from Panel D of Table 1.7, stocks in the highest-DLI decile having the smallest liquidity beta earn the highest return while stocks in the lowest-DLI decile having the largest liquidity beta earn a lower return. An unconditional cross-sectional regression where first-month returns are regressed on the liquidity betas will likely produce a negative risk premium on the aggregate liquidity factor, which is counter-factual. This is again because the large first-month return on high-DLI stocks is mainly driven by the price recovery following the temporary liquidity shock, rather than a permanent liquidity risk premium as can be captured by the loading on the aggregate liquidity factor.

1.3.5. Characteristics regression

In this subsection, we want to directly examine how various stock characteristics explain next month stock returns. Since various characteristics are highly correlated with each other at the portfolio level (see Panel A of Table 1.8), sorting stocks into portfolio according to one characteristic will inevitably induce dispersion along the dimensions of other characteristics. Therefore, double-sorting is less effective in controlling for these characteristics. We therefore use a cross-sectional regression approach at individual stock level. If the first-month high return on financially distressed stocks are in fact driven by high default risk and DLI captures default risk better than other stock characteristics, we would expect DLI to be significant in the cross-sectional regression even with the presence of other stock characteristics. On the other hand, if the first-month high return is a result of of the liquidity-induced price reversal, we would expect $Pastret$ to always be strongly significant. Since a larger price concession will be followed by a larger price recovery, *ceteris paribus*, the past one-month return is negatively related to the next-month return in a mechanical way Finally, as financially distressed stocks are typically illiquid, we would also expect the liquidity measure *Amihud* to be significant in the regression, among others.

The cross-sectional regression approach is similar to Brennan, Chordia and Subrahmanyam (1998). We control for systematic factor risk by first computing the Fama-French three factor alpha.[34] The factor loadings at month m are computed using rolling window regression from $m - T - 1$ to $m - 1$. For each month from 1970/01 to 1999/12, we run a cross-sectional regression of the next month alpha on various stock characteristics from the current month. All variables are cross-sectionally demeaned so the intercept term of the regression is zero. In addition, the stock characteristics are standardized so the regression slope coefficient of a variable can be interpreted as the impact on the alpha of a one standard deviation change in the variable. The slope coefficients are averaged across time and reported. The robust t statistic is computed using the Newey-West autocorrelation adjusted standard error with 12 lags. We consider: $Pastret$ (stock return during the month prior to portfolio formation), $Amihud$, DLI, $Size$ (log of market capitalization) and B/M (book-to-market ratio). We exclude stocks with missing characteristics and negative B/Ms.

Panel A of Table 1.8 reports the correlations among these five characteristics in both the full sample and the top DLI-quintile subsample. Then signs of these correlations are all consistent with the pattern reported in Panel A of Table 1.7. DLI is highly correlated

[34]The results are qualitatively similar if the first month returns instead of alphas are used.

with *Size* and *B/M*. *Amihud* and *Pastret*, on the other hand, are less correlated with other characteristics at individual stock level. Panel B of Table 1.8 reports the regression results where factor loadings are computed using monthly returns in a rolling window of 5 years. In the first three regressions (Models 1 to 3), the only regressor is either *DLI*, *Amihud* or *Pastret*. Either *DLI*, *Amihud* or *Pastret* individually is significantly associated with the next month stock return alpha. *Pastret* is strongly significant (t-value of -9.7) and *Amihud* is slightly more significant than *DLI* (t-value of 3.56 for *Amihud* v.s. 3.16 for *DLI*). *DLI*, however, becomes insignificant with the presence of other characteristics (Model 4 and 5). Specifically, *DLI* becomes insignificant once *Pastret* and *Amihud* are included (Model 4). In addition, since all three characteristics are correlated with *Size* and *B/M*, both of which are shown to have explanatory power on alpha, Model 5 controls for the *Size* and *B/M* characteristics by including them in the regressions. In Model 5, *DLI* is not significant and assumes the wrong sign but *Pastret* and *Amihud* are still significant. Finally, since the liquidity-shock-induced price reversal is likely to be more pronounced for illiquid stocks, we would expect a interaction term between *Pastret* and *Amihud* to be negative and significant. This is indeed the case as in Model 6. The interactive term is highly significant and subsumes the explanatory power of *Amihud*. We also repeat the regressions in the sample we are more interested in – the group of stocks in the highest DLI quintile. The results are qualitatively identical.[35]

[35]Since risk characteristics may change when a stock becomes financially distressed, factor loadings estimated using a rolling window of 5 years may not reflect the risk characteristics of the stock at portfolio formation. As a robustness check, we estimated the factor loadings using a "sum-beta" method with daily return in a much shorter rolling window of 6 months and obtain qualitativey similar results. These results are available from the authors upon request.

1.4. Robustness

In the previous section, we establish that firms that have recently become financially distressed (new high DLI stocks) are more likely to experience return reversal, which explainw the abnormal first-month returns on their stocks. Relatedly, a battery of market liquidity attributes changes are consistent with a clientele change effect. Throughout, we argue that the change of default characteristics of the stock is a necessary condition in driving clientele change, and is ultimately related to the return reversal effect. However, since the underlying debt levels of these new high DLI stocks fluctuate only modestly at the monthly frequency, most of the action is coming from firms that have experienced significant recent price declines. The fact that Default Likelihood Indicator (DLI) is a function of stock price and past return raises two concerns. First, is our result purely driven by low-priced past losers and therefore just a relabeling of the short-run return reversal ("simple return reversal effect") that has been well documented in previous literature (c.f. Jegadeesh (1990), Lehman (1990), and most recently Avramov, Chordia and Goyal (2005))? Second, in this paper, we argue that a sharp increase in default risk measured by DLI leads to institutional selling, resulting in a further price depression ("default-driven return reversal"). Could the causal relationship go the other way? In other words, could some exogenous institutional selling activity depress the stock price, translating into to a sharp increase in DLI? In this section, we address these two concerns.

1.4.1. Simple return reveral or default-driven return reveral?

To differentiate simple return reveral from default-driven return reveral, we construct and compare two portfolios in the following fashion. During our sample period from 1970 to 1999, at the end of each month, we first sort all stocks according to DLI into deciles 1 (low-default risk) to 10 (high-default risk) as in Vassalou and Xing (2004), and then examine two stock portfolios:

(1) New High-DLI Stocks: a portfolio of high-default-risk stocks which have become financially distressed only recently (i.e., they are in the 10th DLI-decile during the current month but not in the previous month).

(2) Characteristics-matched Low-DLI Stocks: a portfolio of stocks with similar past returns and trading prices but relatively low default risk. This portfolio is constructed as follows: at the end of each month, we focus on low-DLI stocks (stocks in DLI-decile 1 to 9) and further sort them on their past one-month returns and trading prices into 36 portfolios. We choose a 6 by 6 double sort to ensure that the number of stocks in each of the 36 portfolios is close to that of the New High-DLI stock portfolio. Among the 36 portfolios, we choose the portfolio such that the past one-month return and trading price are closest to those of the New High-DLI stocks on average.

The main characteristics are summarized in Panel A of Table 1.9. The two portfolios have similar past returns (-12.75% vs. -15.50%) and trading prices (\$4.85 vs. \$4.14). If the return reversals documented in our paper are purely driven by low-priced recent losers, we would expect the two portfolios to have similar next one-month returns. In fact, since the Characteristics-matched Low-DLI stocks have slightly lower past returns and trading prices, we would expect them to have slightly higher returns than New High-DLI stocks next month. However, the opposite is observed: the New High-DLI stocks have much higher returns than the Characteristics-matched Low-DLI stocks (2.93% vs. 2.15%) have.

The difference of 78 bps per month is highly significant (t-value is 3.24) and not casued by difference in risk exposure to common factors. The Fama-French three-factor risk-adjusted return difference is 58 bps per month with an associated t-value of 2.37. Interestingly, the 58 bps difference is also in line with the magnitude of the liquidity shock documented in the paper. To directly controls for the standard short-run return reversal effect, following Gatev, Goetzmann and Rouwenhorst (2006), we also augment the three factor model wth an additional "reversal factor".[36] The resulting four-factor risk-adjusted return difference hardly changes (56 bps per month with an associated t-value of 2.30) as the factor loading on the reversal factor is close to zero. We therefore conclude that the higher return on New High-DLI stocks is not due to the standard short-run return reversal effect.

The dimension in which these two portfolios differ significantly is default likelihood. The new high-DLI stocks recently experience a sharp increase in their default likelihood (the portfolio DLI jumps from 0.099 to 0.233 during the portfolio formation month). This is not the case for the characteristics-matched Low-DLI stocks (the portfolio DLI increases only slightly from 0.02 to 0.033). In the paper, we argue that a sharp increase in default likelihood will likely trigger a larger clientele change and lead to heavier institutional selling pressure. As a result, the stock price will be further depressed and lead to a larger price recovery during the next month, explaining the higher return on the New High-DLI stocks.

[36]The short-term reversal factor, DMU , is constructed as the return on a zero investment portfolio which is long last months losers and short last months winners. It is obtained from Professor Ken French's website.

Finally, we verify that there is indeed more selling pressure on the New High-DLI stocks (relative to the characteristics-matched low-DLI stocks) using both the actual institutional transaction data provided by the Plexus Group and the mutual fund holding data. To faciliate the comparison, the numbers in Panel B of Table 1.9 corresponding to the New High-DLI stocks are reproduced from Table 1.5 Panel A and Table 1.6 Panel B. The column definitions can also be found there. As a percentage of total number of shares outstanding, the New High-DLI stocks experience much larger and more significant selling pressures. Using actual institutional transaction data, we observe large and significant institutional selling pressure on the New High-DLI stocks during the month when they experience the sharp increase in DLI. The selling pressure on the characteristics-matched low-DLI stocks, on the other hand, is much smaller and insignificant. At lower frequency, we also find consistent evidence using the mutual fund holding data. The mutual fund holding change is -0.948% during the event quarter for the New High-DLI stocks. Although the mutual fund holding change is also significantly negative for the characteristics-matched low-DLI stocks, its magnitude is only -0.045%, not a likely representation of a significant clientele change.[37]

These set of results suggest that just being a low-priced past loser is insufficient to generate the strong return reversals and systematic insitutional sellings. The stock has

[37]Consistent evidence is also obtained at even higher frequency (daily) when examining institutional-trade order imbalance using intraday data from TAQ and ISSM. Following Lee and Radhakrishna (2000), we use dollar cutoffs to determine whether a trade is an institutional trade. In particular, we use $\$50,000$ as our baseline cutoff, and for all trades with dollar value greater than the cutoff value, we classify them as institutional trades. We then compute daily institutional-trade order imbalance (number of institutional-trades that are buyer-initiated minus number of institutional-trades that are seller-initiated as a percentage of total number of shares outstanding) for stocks in the two portfolios. We find that the selling pressure is consistently higher for New High-DLI stocks than for the characteristics-matched Low-DLI stocks.

to simultaneously experience a sharp increase in the default likelihood, which pushes it to the highest-DLI portfolio, so as to trigger a significant clientele change. Therefore, the theme of this paper – clientele change triggered liquidity shock – and the main empirical evidence in this paper tell a different story than the simple price reversal effects shown in the literature.

1.4.2. A discussion of causality

The portfolio characteristics in Panel A of Table 1.9 also allow us to address the second concern: could the sharp increase in DLI be a result of (rather than the cause of) institutional selling as DLI itself is a function of stock price? We observe that the new high-DLI stocks experience an average increase in DLI from 0.099 to 0.233. On the other hand, the size of the transitory price concession caused by institutional selling is about 60 bps, as documented in the paper. A simple back of the envelope calculation reveals that the relatively small price concession is not sufficient to cause such a large increase in DLI.[38]

To address the concern more directly, using actual institutional transaction data from Plexus group, we sort stocks according to institutional selling pressure (defined as number of shares sold as percentage of total number of shares outstanding) into deciles every month and examine the associated changes in DLI. The results are reported in Panel C of Table 1.9. For the portfolio that of the heaviest institutional selling pressure, the average percentage shares sold by institutions is about 1.33% – ten times larger than that

[38]Based on equation (1), an increase in DLI from 0.099 to 0.233 implies a decrease in Distance to Default (DD) of about -0.56 [$= N^{-1}(0.099) - N^{-1}(0.233)$]. At monthly frequency, book value of debt (X) and drift term in the asset value (μ) do not change much. Assuming an asset volatility (σ_A) of about 0.65 (the sample mean for the "new" high-DLI stocks), the decreases in DD implies a reduction in asset value (V_A) by 36% ($= 0.65 \times 0.56$). The reduction in asset value seems to be implausibly large for a rather tiny change in the equity value (-0.6%).

on the New High DLI stocks (see Panel B of Table 1.9). The associated increase in DLI, however, is only 0.07%, almost negligible compared to that for the New High DLI stocks ($23.3\% - 9.9\% = 13.4\%$ as in Panel A of Table 1.9). All of the above evidence leads us to believe that institutional selling is not the main cause of the sharp increase in DLI we observed for the New High DLI stocks.

1.4.3. Additional Discussions

In Appendix B, we show that the first-month high return on financially distressed stocks is unlikely driven by bias through random bid-ask bounce or the high level of uncertainty associated with the distress event. In addition, we gauge the economic significance of the return and find outside investors unable to take advantage of the high return on financially distressed stocks after transaction cost. This finding has two implications. First, the high return and large Sharpe Ratio earned by high default risk stocks do not constitute a violation of efficient market hypothesis. Second, only market makers, generically defined, are compensated for providing liquidity when it is most needed. Finally, we show that when market making became more competitive after the mid-1997's, the return earned by financially distressed stocks also dropped, indicating a reduced compensation for liquidity provision.

1.5. Conclusion

Vassalou and Xing (2004) show that stocks of firms under financial distress, on average earn a large positive abnormal return during the first month after portfolio formation, even after adjusting for risk using standard asset pricing models. In this paper, we show that a sharp rise in a firm's exposure to financial distress risk triggers a clientele change for its stock, resulting in temporary selling pressure. For example, mutual funds significantly decrease their holdings of stocks from firms that experience sharp rises in their default likelihood measures. When the liquidity of the stock later improves, the stock price recovers, which contributes to the high return on financially distressed stocks during the first month after portfolio formation. Changes in various market microstructure attributes of a stock, such as trading volume, percentage bid-ask spread, realized (half) spread and order imbalance measures, are all consistent with there being such liquidity shock. Therefore, a major part of the high return on these stocks can be interpreted as reward for liquidity provision when it is most needed.

Consistent with this view, we find that the high returns on financially distressed stocks accrue during the first month following portfolio formation, but little during the months afterwards, although various risk characteristics hardly change. This result supports the claim in Campbell, Hilscher and Szilagyi (2005), and Garlappi, Shu and Yan (2005), that high default risk itself does not necessarily translate to high return in the future. In addition, we find that although the first month high return on the high-DLI stocks cannot be explained by the standard Fama-French three factors, when we skip a month, the second month return can be well explained by the three factors, and an aggregate default risk factor ceases to be significant in various asset pricing tests using the second month returns on portfolios sorted on DLI. Collectively, evidence so far suggests that there is no need for a separate aggregate default risk factor in reduced form asset pricing models. Our findings also highlight the time-varying nature of a stock's exposure to liquidity risk. A stock's exposure to Pastor and Stambaugh's (2003) aggregate liquidity factor increases significantly during the liquidity shock and then returns to its normal level afterwards, coinciding with the initial stock price concession and subsequent price recovery.

In this paper, we measure the default or financial distress risk using Default Likelihood Indicator (DLI) proposed by Vassalou and Xing (2004), which has the advantage of incorporating market price information that is more frequently updated. However, since DLI is estimated at monthly frequency, we still cannot identify the exact time at which a firm experiences a sharp increase in default risk. Our approach, which is essentially a calendar-time approach, only identifies the *average* impact of default risk on stock liquidity at a portfolio level. A complimentary event-time approach which focuses on large credit rating downgrades for individual firm, could potentially provide a sharper identification of such impact. In addition, if the bond of the firm is also traded, we can make use of the information embedded in the bond price change to better isolate that component of the stock price change which is due to the liquidity shock. These are potential venues for future research.

CHAPTER 2

Herding, Information Aggregation and Momentum Effects

There is a large empirical literature documenting momentum effects in stock prices. For example, Jegadeesh and Titman (1993) find price momentum - past winners continue to win and past losers continue to lose. They suggest market systematically under-reacts to firm-specific news. Ball and Brown (1968) find post-earnings announcement drift in stock prices - prices of stocks with positive earnings surprise continue to drift up and that of stocks with negative earnings surprise continue to drift down. Chan, Jegadeesh and Lakonishok (1996) examine earnings momentum and price momentum effects and find one does not subsume the other. They provide empirical evidence consistent with the idea of slow information diffusion. Chan (2003) finds similar price drift patterns after newspaper coverage. Portfolio strategies that exploit these patterns show abnormal risk adjusted returns (Fama and French, 1996), posing a serious challenge to standard theory that assumes perfect and informationally efficient securities markets. In this paper, I provide an explanation for momentum effects.

Financial analysts provide valuable information to stock market participants. Barber, Lehavy, McNichols and Trueman (2001), and Green (2006) show markets quickly react to analyst recommendations. The profits from trading on the analyst recommendations are either overwhelmed by transaction costs or accrue within a very short window. Jegadeesh, Kim, Krische and Lee (2004) find that the change of the ananlyst recommendations contains value relevant information beyond twelve other known stock characteristics having forecast power of future returns. Using survey data, Cheng, Liu and Qian (2006) reveal that more than 90 percent of the sophisticated institutional investors make use of research by the sell side analysts in the investment decisions. When both buy-side and sell-side analyst researches are available, these managers put an average weight of 10 to 35 percent on those generated by sell side analysts (see figure 1 in their paper). Boni and Womack (2006) establish that analysts have skills at industry level in form of stock recommendations, and aggregated analyst information within an industry helps to explain industry momentum effects first documented in Moskowitz and Grinblatt (1999).

However, the way analysts process and reveal information to the public will be influenced by their incentives and possibly their cognitive constraints. One such bias is the tendency to herd among sell-side financial analysts in their quarterly earnings forecasts. Specifically, security analysts at times may choose to bias their quarterly earnings forecasts towards the public consensus and away from their own best estimates. Herding tendency, by its very definition, implies inefficient aggregation of value relevant private information possessed by the analysts (Banerjee, 1992; Bikhchandani, Hirshleifer and Welch, 1992; Trueman, 1994). Therefore, when analysts have such tendencies, some of the valuable firm-specific information might not reach the investors in a timely manner, if the market participants do not adequately account for such bias. The stronger the herding tendency among the analysts is, the less the amount of private information gets aggregated, and the more pronounced the momentum effects in stock prices.

The testable implication is that the cross-sectional differences in analysts' herding tendency generate predictable patterns in stock returns, because different strengths of herding tendencies are associated with different levels of information aggregation. First, price momentum effects should be more pronounced among the stocks when analysts covering that stock have a stronger tendency to herd, because stronger herding tendency further reduces the information aggregation. In my sample of large market capitalization stocks well-covered by financial analysts, winners (those stocks that are in the highest past 12-month return quintile) earn only 63 basis points per month more than the losers (those stocks that are in the lowest past 12-month return quintile) over the following 3 month holding period. Consistent with the findings in Hong, Lim and Stein (2000), this simple momentum strategy return is statistically insignificant. However, among the subset of stocks where analysts have a strong herding tendency, winners earn 114 basis points per month more than the losers, and the price momentum strategy return is statistically significant at 1 percent level. In sharp contrast, when there is little herding tendency among financial analysts, the price momentum strategy earns only 29 basis points per month, and is not statistically different from zero. The difference in the momentum strategy's returns of high and low herding stocks ranges from 74 basis points per month for 3-month holding horizons and 45 basis points per month for the 12-month holding horizons and these differences are statistically significant at the 1 percent level. These findings are consistent with my view that the bias of the information producers and disseminators – the sell side financial analysts – may impact asset prices.

Second, conditional on the past returns, a stock should exhibit more pronounced price momentum at those points in time when the analysts following that stock tend to herd more. Therefore, among the stocks with good news in the past, for the subset of stocks where analysts covering them exhibit stronger herding tendency, the upward drift in price is more pronounced than the subset with weaker herding tendency. Similarly, among the stocks with bad news in the past, for the subset of stocks with stronger herding tendency among analysts covering them, the downward drifts in prices should be more pronounced than the subset with weaker herding tendency. My evidence shows that, after the risk adjustment, the strategy of going long on the loser stocks (low past twelve month returns) with low herding tendency and shorting the loser stocks with high herding tendency earn about 6.48 percent per year (significant at the 1 percent level). Similarly, taking a long position in the winner stocks (high past twelve month returns) with high herding tendency and shorting the winner stocks with low herding tendency earn about 3.60 percent per year (significant at the 10 percent level). These asymmetric effects in returns are consistent with how the bias of analysts related to the information aggregation. When the performance of a stock is poor, financial analysts may have an incentive to either withhold value relevant information for fear not doing so may alienate themselves from the management or dropping the coverage, which in turn could reduce the speed with which information gets reflected in stock prices when investors do not account for the information in “the dogs that do not bark”.

A natural question is why analysts herd on some stocks at some points in time but not on other stocks at other points in time. While I do not have a complete answer to this question, I do find that the characteristics of stocks that analysts herd on are consistent with what was conjectured in the previous literature.[1] I show that analysts of small stocks

[1]Several explanations of such bias of information producers and disseminators have been advanced, including reputation concerns (see Graham, 1999; Lamont, 2002; Lim, 2003); investment banking business generation incentive (Michaely and Womack, 1999; Chen and Jiang, 2006); career concerns (Hong, Kubik and Solomon, 2000; Zitzewitz, 2001; Hong and Kubik, 2003); trading volume generating incentives (Chen and Jiang, 2006); access to the management (Lim, 2001). Regarding the determinants of herding tendencies, Shiller (1995) and Hirshleifer and Teoh (2003) conjecture that the tendency to herd is related to the complexity of decision environment. Shiller (1985, 1987), and Shiller and Pound (1989) suggest that media may be at least partially responsible for fuelling the herding tendency.

or growth stocks indeed tend to herd more during those time periods when there is more news media coverage. My findings about herding tendency and firm characteristics have important implications for the recent literature on information uncertainty and expected returns (Jiang, Lee and Zhang, 2006; Zhang, 2006). There is some evidence that stocks with higher level of information uncertainty about fundamental values earn significantly more momentum profits. So far this observation is almost exclusively interpreted as being supportive of investor overconfidence hypotheses. The usual argument is that individuals tend to be more overconfident in settings where feedback on their information or decisions is slow or inconclusive than where the feedback is clear and rapid. In this paper, I provide an alternative channel. Notice that smaller growth stocks are inherently harder to value, and the analysts following them may have strong tendency to herd. But the stronger herding tendency leads to less efficient information aggregation, which in turn generates more pronounced momentum effects.

This paper contributes to our understanding in several threads of literatures. First, this paper is related to several belief-based theoretical models of momentum effects, including Barberis, Shleifer and Vishny (1998; BSV), Daniel, Hirshleifer, and Subrahmanyam (1998; DHS), as well as Hong and Stein (1999; HS).[2] The empirical evidence lends support to characterize momentum effects as underreaction to news as emphasized in BSV and HS. Moreover, the empirical evidence presented in this paper suggests a plausible mechanism of underreaction currently unexplored in these information-based models. If

[2]Barberis and Thaler (2003) provide a comprehensive survey on behavioral finance, including several belief-based models on causes of momentum effects. To date, there is still no formal preference-based model on momentum effects. As a first step towards the preference-based model of momentum effects, Barberis and Xiong (2006) make a theoretical attempt to integrate prospect theory and mental accounting to generate so-called "disposition effects" (Shefrin and Statman, 1985) - the investors hold their losing stock too long and sell their winning stocks too fast.

there are biases of financial analysts due to incentive or cognitive constraints but market fails to make sufficient adjustment for such biases for rational or behavioral reasons, the biases may impact asset prices.

Second, this paper is related to and benefits substantially from the literature on the measurement of herding tendency among sell-side financial analysts. Welch (2000) extracts herding tendency from analysts' stock recommendations. Zitzewitz (2001) develops a novel method of estimating the degree of herding versus exaggeration of differences (the opposite of herding; or anti-herding for short) among analysts. Measurement of herding tendency and its statistical foundations are further carefully examined in the recent work of Bernhardt, Campello and Kutsoati (2006) and Chen and Jiang (2006).[3] However, I am not aware of other papers that investigate herding tendency among sell-side financial analysts and asset prices, in particular the momentum effects.

Finally, this paper makes a contribution in estimating panel data commonly encountered in finance when the Fama-MacBeth (1973) procedure is subject to Petersen's critique (Petersen, 2006). Petersen (2006) points out that when the time-series length is short in the panel data, the t-statistics computed from the Fama-MacBeth procedure are usually biased. I deploy a type of spatial heteroskedasticity and autocorrelation consistent (SHAC) estimator in the cross-sectional time-series asset pricing tests, built on the theoretical work of Conley (1999). The SHAC estimator nests several commonly used covariance estimators as special cases. In conjunction with an "economic distance" I construct from firm characteristics, this type of SHAC estimator is particularly powerful in

[3]Clearly, this paper is also related to a broader category of studies on the alleged "herd like" behaviors among a wide spectrum of other direct market participants including mutual fund managers (Grinblatt, Titman and Wermers, 1995; Nofsinger and Sias, 1998; Chevalier and Ellison, 1999; Wermers, 1999), speculative traders (Brunnermeier and Nagel, 2004; Temin and Voth, 2004).

panel data with relatively short time length but large cross sections. The application of the SHAC estimator has implications for empirical research beyond what is examined in this paper. Using the traditional Fama-MacBeth regressions and the time-series cross-sectional regressions with the novel SHAC estimator, I am able to show that effects of the herding tendency on returns are distinct from earnings momentum effects, information uncertainty effects and liquidity effects, among others.

2.1. Momentum Effects: Synthesis and Hypotheses Development

2.1.1. Risk-based Explanation of Momentum Effects

Jegadeesh and Titman (2001) conclude that time-series dependence in returns contributes most to the price momentum effects. Theoretically, Berk, Green and Naik (1999) attempt to rationalize momentum effects, and they are able to show past returns contain valuable information about future returns, and momentum effects could be attributed to risks. However, as summarized by Johnson (2002), momentum strategies "*do not appear to be especially dangerous.*"

Empirically, the multifactor model does not help to explain momentum effects (Fama and French, 1996). In fact, the *unconditional* multifactor model usually exacerbates the mispricing (Fama and French, 1996; Grundy and Martin, 2001; Liu, Warner and Zhang, 2004). It is not entirely surprising that the traditional risk-based models have difficulty in explaining momentum effect. Momentum strategy is associated with high Sharpe ratio. For example, the simplest momentum strategy can be constructed by buying past twelve month winners and selling past twelve month losers, skipping one month, and holding the winners and losers portfolios by one month (UMD factor as in Fama and French, 1996). This momentum portfolio earns about 84 basis points per month between 1962 and 2005 with a monthly Sharpe ratio of 0.21; and it earns about 83 basis points between 1984 and 2005 (during my sample period) with a monthly Sharpe ratio of 0.18.[4] In a frictionless market, MacKinlay (1995) shows such a high Sharpe ratio is difficult to be explained by a multifactor model, and Hansen and Jagannathan (1991) argue that a high Sharpe ratio implies a high volatility of the stochastic discount factor (SDF) and high level of risk aversion, both are difficult to be justified empirically. Recent development of unconditional multifactor model includes the Fama-French three factor model augmented with liquidity risk factor in Pástor and Stambaugh (2003) and Sadka (2006). They show that the momentum profits attenuate after adjusting liquidity risk. Chordia and Shivakumar (2002) and Griffin, Ji, and Martin (2003) consider a set of common factors related to business-cycle risk, and achieve different degrees of success in explaining momentum effects. Conditional multifactor models (Jagannathan and Wang, 1996; Lettau and Ludvigson, 2001; Santos and Veronesi, 2006) are shown to be able to explain value/growth premium to some extent- particularly in cross-sectional regressions yet they cannot explain momentum effects [for the cross-sectional tests, see Wu (2002); for time-series tests, see Lewellen and Nagel (2006)].

2.1.2. Information-based Explanation of Momentum Effects

Several information-based models, some of them relying on investor psychological biases and bounded rationalities, have been advanced to explain momentum effects. These

[4]For comparison, the value/growth strategy (HML factor, Fama and French, 1993) earns about 47 basis points between 1962 and 2005 with a monthly Sharpe ratio of 0.16; and it earns about 33 basis points between 1985 and 2005 (my sample period) with a monthly Sharpe ratio of 0.10. These numbers are computed based on the UMD and HML factors obtained from Ken French's website.

models share the common feature that information does not aggregate efficiently into asset prices, even though the exact underlying mechanism of such information misaggregation differs from one model to another. Barberis, Shleifer and Vishny (BSV, 1998) suggest the presence of bias in investor's updating due to dissonance bias and over-extrapolation can generate momentum effect, and they characterize the momentum effects as "under-reaction". Daniel, Hirshleifer and Subrahmanyam (DHS, 1998) argue that the investor's overconfidence and the biased self-attribution contribute to the momentum effect. They argue that the momentum effect is primarily driven by "delayed overreaction" to public information, as opposed to "under-reaction" to firm-specific news advanced by Jegadeesh and Titman (1993). Not explicitly relying on any particular type of behavioral bias, but bounded rationalities, Hong and Stein (HS, 1998) suggest that the interaction of slow information diffusion and investor heterogeneities causes the momentum effect.

Empirically, Hong, Lim and Stein (2000) use the analyst coverage (i.e., size adjusted number of analysts) a proxy for the speed of information diffusion, and find that momentum effects are rather weak among stocks covered by analysts. Recent evidence from Vega (2006) suggests that the analyst coverage is not associated with information diffusion once the stock is covered by the analysts. In fact, analyst coverage is *negatively* correlated with the speed of information diffusion. Grinblatt and Han (2005) suggest that the disposition effect may generate the price momentum effects. In particular, the aggregate unrealized capital gains seem to be related to the profitability of momentum strategies. Using different proxy variables for the unrealized capital gain measure, Frazzini (2006) and Grinblatt and Han (2005) illustrate the role of disposition effects in generating drift patterns in stock prices. Motivated by the adjustment and anchoring bias (Kahneman, Slovic,and Tversky, 1982), George and Hwang (2004) show an investment strategy based on the 52-week high price outperforms past return based momentum strategies.

2.1.3. Persistence of Momentum Effects

Another perplexing aspect of momentum effect is its persistence. Chen, Stanzl and Watanabe (2001), Korajczyk and Sadka (2004), Lesmond, Schill and Zhou (2004) estimate the transaction costs associated with momentum strategies, and suggest that the large trading and price impact costs seriously impede the implementability of such strategies, hence the limited-arbitrage (Shleifer and Vishny, 1997) contributes to the persistence of momentum phenomenon.

2.1.4. Herding Among Analysts and Information Aggregation

The informational herding models of Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992) provide an intuitive way to think about how to relate herding tendency to the efficiency of information aggregation. In these models, there is so-called information cascade phenomenon. In a sequential decision setting, the agents completely and rationally disregard their own private signals at late stages; and the *ex post* collective outcome may be suboptimal due to the inefficient information aggregation. This is a strong form of informational herding. If agents do not completely disregard their private signals, then it is not information cascade but herding, which is a weak form of informational herding.[5] For the purpose in this paper, whether the form of informational herding

[5]Many studies point out that once the assumptions of simple information structure and constrained action space are relaxed, then information cascade described in Banerjee (1992), Bikhchandani *et al.* (1992) disappears (Avery and Zemsky, 1998). However, as Chari and Kehoe (2004) illustrate, if the *exogenous timing* assumption is also relaxed, then information cascade emerges again.

is strong or weak is inessential. The main point is that the value relevant private information does not quickly get impounded in the analyst forecasts due to the herding tendency among the analysts.

Herding behaviors can occur through various channels. One such channel, which is particularly relevant to understand the herding tendency among financial analysts, is reputational and career concerns. The key insight is that one agent can benefit from mimicking other agents, so she intentionally misweights her own private information. This channel is investigated in Scharfstein and Stein (1990), Trueman (1994), Prendergast and Stole (1996), Graham (1999) and Lim (2003). The second channel, which has direct implication for the market aggregation of information in equilibrium, is the information acquisition channel. A general feature of this type of models is that some information may not be acquired, because payoffs from different pieces of information depend on the information acquisition strategies of other players. Theoretical models in Brennan (1990), Froot, Scharfstein and Stein (1992), Dow and Gordon (1994) have this feature. Third, there are situations where allegedly some elements of irrationalities lead to observed herding behaviors. Ehrbeck and Waldman (1996) consider a reputational model of herding among forecasters, but reject rationality and favor behavioral explanations.[6] In summary, regardless of rational or behavioral reasons, herding tendency among analysts implies

[6] A complete list of herding literature is beyond the scope of this paper. Ivo Welch's website (http://welch.econ.brown.edu/cascades/) has an excellent collection of herding literature. Bikhchandani, Hirshleifer, and Welch (1998), and Hirshleifer and Teoh (2003) provide comprehensive suvery of herding behaviors in capital markets. Chapter 6 of Brunnermeier (2001) gives a textbook treatment of theoretical development in the herding literatures. Because the purpose of this paper is to investigate the pricing implication of herding behaviors, I do not attempt to differentiate whether incentives or cognitive constraints are driving the herding tendencies among analysts. It could well be the case that both effects influence analysts.

failure of complete information aggregation in the consensus forecast, particularly the aggregation of private information of individual analysts.

2.1.5. Biases in Analysts Forecasts and Asset Prices

Financial analysts collect "new information on the industry or individual stock from customers, suppliers, and firm managers" and incorporate "this new information into their analysis to form earning estimates and recommendations" (see Womack and Michaely, 1999). There is no doubt that quarterly earning forecasts of the financial analysts contain material private information not current in the information set of the general public before analysts issue their estimates and opinions. The herding tendency among financial analysts implies that the aggregation of private information is inefficient in their forecast earnings, recommendations and other forecast products. The forecasts by financial analysts are just one piece of information used in investors' valuation model, and they are not equivalent to asset prices. As long as the market participants completely adjust for such bias when they use the forecast information to formulate the prices, such bias will not impact asset prices.

Does market sufficiently adjust for these biases when forming asset prices? On the one hand, evidence suggests that market is aware of information content of analyst forecasts. Cooper, Day and Lewis (2001) find the market reacts strongly to forecasts by lead analysts, who generate timely information. Gleason and Lee (2003) and Clement and Tse (2003) illustrate that bold forecast revisions are associated with greater return response coefficients than those for herding forecast revisions. Consistent with Clement and Tse (2005), bold forecast revisions imply more complete aggregation of private information.

Ivković and Jegadeesh (2004) show the precision of analysts' private information increases between two consecutive earning announcement dates. The market's reactions to analyst recommendations indicate the market factors in such changes in forming asset prices. However, at least some investors fail to adjust for some of these most widely publicized and easily identified biases. Malmendier and Shanthikumar (2005) find less sophisticated investors do not adjust for bias due to investment banking relationship when they act on the recommendations issued by non-independent analysts. Bradley, Jordan and Ritter (2006) show that the market does not react differently towards recommendations issued by the affiliated versus non-affiliated analysts after the IPO's "quiet period".

Specifically, I consider the following joint hypothesis: If market does not adjust biases of analyst forecasts sufficiently, and to the extent that price momentum effects reflect inefficient aggregation of private information of the analysts, there will be short-term price drifts. There are two testable implications here. The first one is what I call the "concentration of momentum effects hypothesis". The momentum effects should concentrate in the portfolios where the aggregation of private information is least efficient. The second testable hypothesis is what I call the "within-news" herding effect. Conditional on the sign of the public news, the directional drifts in prices should be more pronounced among stocks with stronger herding tendency among analysts.

2.2. Data Description

The data come from several sources. Stock returns and accounting data are obtained from CRSP and Standard and Poor's COMPUSTAT databases respectively. To link these two databases accurately, CRSPLINK file produced by Center for Research in Security

Prices (CRSP) is used. I only consider common shares (share code is either 10 or 11) traded on NYSE, AMEX or NASDAQ (exchange code is 1, 2, and 3). To mitigate market microstructure noise induced confounding effects, I exclude all stocks traded below 5 dollars at portfolio formation month (Jegadeesh, 1990; Jegadeesh and Titman, 1995).

Sell-side financial analyst quarterly earning per share forecasts are obtained from Institutional Brokerage Estimation System (I/B/E/S) detailed historical files and summary historical files. I use the unadjusted I/B/E/S detailed individual analyst earning per share forecast and actual earning per share data, and split-adjust the earning per share values over time using the adjustment factors provided by I/B/E/S without any rounding.[7] In order to implement some of the estimation procedures at individual stock level, I further impose the constraint that there are at least seven forecasts made by at least two different analysts during a quarter prior to a quarter's earning announcement for the stocks to be included in the sample.

For actual earning per share data, I use the actual values from I/B/E/S. The actual earning per share from I/B/E/S is preferable to the actual earning per share from COMPUSTAT for a number of reasons. First, I/B/E/S excludes "all discontinued operations, extra-ordinary charges, and other non-operating items", so the actual earning per share recorded by I/B/E/S is comparable to the analyst forecast earning per share forecast

[7]To ensure the per share amounts are comparable over time, I/B/E/S retrospectively split-adjusts the earning per share forecast and actual values, rounded to the nearest penny, and these adjustments are reflected in the regular I/B/E/S release. However, these adjustments made by I/B/E/S render the regular I/B/E/S release unsuitable for my research for two reasons. First, the regular release of I/B/E/S introduces a potential "look-ahead" bias discussed in Diether, Malloy and Scherbina (2002). As multiple stock splits samples are likely to be successful firms experiencing large stock appreciation, the portfolios formed over time may be subject to this type of look-ahead bias if the regular I/B/E/S release is used. Second, benchmarked against unadjusted files, Payne and Thomas (2003) uncover substantial misspecification errors when regular I/B/E/S release is used in to classify earning surprise.

(I/B/E/S Glossary, 2003), but COMPUSTAT earning per share data are not directly comparable. Excluding special items from COMPUSTAT earning per share mitigates the discrepancy but does not completely resolve it. Second, Della Vigna and Pollet (2005) provide evidence that I/B/E/S is more accurate in recording the dates of earning announcement . Third, COMPUSTAT earning per share follows the generally accepted accounting principle (GAAP), and it contains transitory components; while I/B/E/S earnings (the "street earnings") discard such transitory components, and reflect fundamentals better (Frankel and Roychowdhury, 2004). Consistent with Hong, Lim and Stein (2000), I/B/E/S - CRSP intersection usually contains large firms, so the reduction of number of unique stocks is substantial. After imposing these constraints, my sample has $1,702$ unique stocks from April 1985 to December 2005.

To count the news media coverage, I rely on a news media coverage database built on Factiva's collection of newspaper headlines. The database is constructed by systematically collecting newspaper headlines from Factiva then matched with CRSP based on the ticker symbols at time of news report (identified by Factiva) and those reported by CRSP daily event file (CRSP-DSE). Several quality control measures are in place to ensure accurate matching between CRSP and Factiva.[8] The database contains more than 2.1 million newspaper headlines for total of $8,371$ unique stocks in CRSP from January 1990 to November 2005. Each year, this database covers about $5,000$ unique stocks, or about 80% of the market capitalization of all publicly traded stocks in the U.S. Presumably,

[8] Appendix A in Engelberg and Gao (2006) provides details about construction of this database.

this news media coverage database is comparable to or slightly more comprehensive than similar databases used in prior studies.[9]

Finally, I obtained individual stock's dates of initial public offerings (IPOs) and dates of seasoned equity offerings (SEOs) from SDC Platium database, combined with Jay Ritter's IPO database for IPOs during 1975 to 1984 (Loughran and Ritter, 1995).

2.3. A Measure of Herding Tendency Among Analysts

2.3.1. Estimation of Herding Tendency Among Analysts

To capture the herding tendency among financial analysts, I adopt a measure developed in Bernhardt, Campello and Kutsoati (2006), Chen and Jiang (2006), and Zitzewitz (2001). The basic intuition underlying this measure is simple.[10] If the analyst issues unbiased forecast corresponding to the posterior mean, then the probability that her forecast exceeds or falls short of true earnings should be the same, conditional on *anything* in her information set. In particular, the probability that her forecast exceeds or falls short of true earnings should be the same regardless of whether her forecasts exceeds or falls short of consensus. This simple intuition implies $\varphi = \Pr[sign(FE) = sign(DEV)] = 1/2$ should hold, where FE is the forecast errors of individual analysts, or the forecasted earnings minus the actual earnings, DEV is the forecast of analyst minus the consensus forecast immediately before she announces her forecast, and $sign\left(\cdot\right)$ is the sign function.

[9]For example, the database used in Chan (2003) and Vega (2006) covers about 4000 stocks during 1980 to 2000; and it covers about 766 at the beginning of 1980 (first year of their sample), and 1,500 at the end of 2000 (last year of their sample). The database used in Antweiler and Frank (2005) covers $245,429$ news stories during 1973 to 2001.

[10]More rigorous statistical foundations can be obtained from their papers. For example, section 1 in Chen and Jiang (2006) provides an intepretation from Bayesian weighting of public and private information. This paper follows their notations.

The systematic deviation from this equality captures the extent of herding tendency. The above probability can be estimated by its sample analog:

$$\hat{\varphi} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{1}\left[sign\left(FE\right) = sign\left(Dev\right)\right] \tag{2.1}$$

where $\mathbf{1}[\cdot]$ is the indicator function such that $\mathbf{1}[\cdot] = 1$ if and only if $sign\left(FE\right) = sign\left(Dev\right)$, and $\mathbf{1}[\cdot] = 0$ otherwise, and K is the number of forecasts under consideration.

There are several implementation issues. First, to estimate the above model, it is necessary to specify the *immediate* market consensus forecast before the analyst's forecast. Consider an analyst j , who is going to announce the $(N+1)$-th forecast. Immediately before she makes the forecast, she observes two sequences of numbers $F = \{f(1), f(2), ..., f(N)\}$ and $D = \{d(1), d(2), ..., d(N)\}$, where the first sequence (F) contain all the forecasts made sequentially before hers, and the second sequence (D) contains the time elapse after the n-th forecast was made $(n = 1, 2, ..., N)$. In a sequential Bayesian updating framework, and under some assumptions about the updating rules, it can be shown that the latest forecast $f(N)$ immediately before the $(N+1)$-th forecast summarizes all previous public and private information. Nevertheless, the market consensus forecast used in generating the $(N+1)$-th forecast is still a function of all previous forecasts, denoted as $G(F)$.[11] Another way to characterize the consensus is to exogenously assume the consensus forecast be a function of all previous forecast as an approximation of $G(F)$, denoted as $\tilde{G}(F)$. Consistent with Chen and Jiang (2006), I adopt the second approach and specify a version of "time-decay weighted" average consensus. The "time-decay weighted" average consensus can be computed as $c(N) = \sum_{n=1}^{N} f(n)\, w(n)$ where

[11] Along this line of logic, Friesen and Weller (2003) formalize this notion in a sequential setup.

the weighting factor is defined as $w(n) = \left(\frac{1}{d_n}\right) / \sum_{n=1}^{N} \frac{1}{d_n}$. The main idea underlying this specification is that the further away the forecast, the less weight is put on to generate the consensus, since recent forecasts contains more updated information.[12]

Second, we need a proxy for the market consensus for the very first analyst's earning forecast at any given quarter. I experiment with three alternative specifications: (1) Previous period quarterly earning realization; (2) Seasonally differenced random walk model (estimated from full sample or estimated from past realized earnings); (3) Dropping the first forecast, and taking the first forecast as the market consensus for the second forecast. It turns out that the exact specification does not matter for my main conclusions. To avoid potential model misspecification, in the subsequent analysis I take the most conservative approach by simply dropping the very first forecast.

Third, instead of relying a pooled cross-sectional time-series regression across all stocks and all analysts to obtain the herding tendency as in Chen and Jiang (2006), I estimate the probability $\hat{\varphi}$ in (2.1) stock by stock and quarter by quarter. For each stock i at quarter q, the estimate from (2.1) is denoted as $\hat{\varphi}_{i,q}$, where the subscript i denotes the i-th stock, and subscript q denotes the q-th quarter. I call the $\hat{\varphi}_{i,q}$ *raw* information weighting factor.

There are two additional issues related to estimating the analyst herding tendency. First, the analyst herding tendency for a given stock is necessary a concept related to the overall history of analyst forecasts of that stock. However, using overall history of

[12]This specification is also consistent with the empirical evidence in Welch (2000), who shows the analysts take advantage of immediate preceding analyst revision of recommendations. Chen and Jiang (2006) show that the alternative linear weight averaging does not change their estimations. Therefore, I choose this "time-decaying" specifications in my analyses.

analyst forecasts may introduce some forward-looking bias. Second, there are also time-series autocorrelations of $\hat{\varphi}_{i,q}$. In order to fully use the time-series correlation of the information weighting factor, I implement a predictive regression to encapsulate such herding tendency for a given stock. The procedure is achieved in three steps. In the first step, I estimate the following autoregressive (AR) model with four lags of $\hat{\varphi}_{i,q}$ at individual stock level,

$$\hat{\varphi}_{i,q} = \alpha_{i,0} + \beta_{i,1} \times \hat{\varphi}_{i,q-1} + \beta_{i,2} \times \hat{\varphi}_{i,q-2} + \beta_{i,3} \times \hat{\varphi}_{i,q-3} + \beta_{i,4} \times \hat{\varphi}_{i,q-4} + \varepsilon_{i,q} \tag{2.2}$$

where $\left\{\hat{\varphi}_{i,q}, \hat{\varphi}_{i,q-1}, \hat{\varphi}_{i,q-2}, \hat{\varphi}_{i,q-3}, \hat{\varphi}_{i,q-4}\right\}$ are the raw information weighting factor computed using (2.1). When available, I use information from past 12 quarters to estimate the AR model, but restrain the minimum number of observations to be 6; otherwise, I simply take the historical average. I choose 4 quarters as the number of lags because it covers an entire fiscal cycle of earnings. In the second step, I use only past four quarter information and the estimated coefficients from the first step to generate the *predicted* information weighting factor ($\tilde{\varphi}_{i,q+1}$),

$$\tilde{\varphi}_{i,q+1} = \hat{\alpha}_{i,0} + \widehat{\beta}_{i,1} \times \hat{\varphi}_{i,q} + \widehat{\beta}_{i,2} \times \hat{\varphi}_{i,q-1} + \widehat{\beta}_{i,3} \times \hat{\varphi}_{i,q-2} + \widehat{\beta}_{i,4} \times \hat{\varphi}_{i,q-3} \tag{2.3}$$

where $\left\{\widehat{\beta}_{i,1}, \widehat{\beta}_{i,2}, \widehat{\beta}_{i,3}, \widehat{\beta}_{i,4}\right\}$ are the coefficients estimated from the AR regressions. The final step is to normalize the predictive information weighting factor by (i) subtracting 0.5; and (ii) truncate them at -0.5 and 0.5 level to preserve the probability interpretation, and define the normalized predicted analysts' information weighting factor as,

$$IWF = \begin{cases} 1/2 & \text{if } \tilde{\varphi}_{i,q+1} > 1; \\ \tilde{\varphi}_{i,q+1} - 1/2 & \text{if } 0 < \tilde{\varphi}_{i,q+1} < 1 \\ -1/2 & \text{if } \tilde{\varphi}_{i,q+1} < 0. \end{cases} \tag{2.4}$$

The normalized predicted analysts' information weighting factor estimated from (2.3) will be used in the subsequent portfolio exercises.[13] Hereinafter, I short-hand the normalized predicted analysts' information weighting factor as "information weighting factor" (IWF).[14]

2.3.2. Results from the Estimation of Analysts' Information Weighting Factor

Based on the specification in equation (2.1) and the individual analyst's quarterly earning forecasts, I first estimate the *raw* quarterly information weighting factors stock by stock and quarter by quarter. Table 2.1 presents the descriptive statistics of these estimates, and two features emerge. First, the mean value of the information weighting factor values (0.52) computed from the time-series and cross-sections is slightly (thought statistically significant) above 1/2, the value implied by the optimal forecast. These estimates are consistent with those in Bernhardt, Campello and Kutsoati (2006), Chen and Jiang (2006),

[13] Clearly, preserving probability intepretation by using (2.4) is not essential for forming portfolios. Without resorting to procedures described in (2.2) to (2.4), the results are qualitatively similar. Additionally, because I require past three-year data to estimate the AR regression in (2.2), this is not feasible for 1985, 1986 and 1987. For year 1985 and 1986, I simply use the raw values in (2.1). For year 1987, I do not exactly use three-year worth of data. Instead, I will use the amount of forecast avaiable, as long as (2.2) can be estimated.

[14] The "information weighting factor" is probably a misnomer. It's clear that it is not a risk factor, but a variable of stock characteristic reflecting the herding tendencies among analysts following the stock.

and Zitzewitz (2001). These authors find analysts *on average* overweight private information or bias the forecasts away from the public consensus.[15] Second, there are large cross-sectional variations of the information weighting factor values within a year, but the magnitudes of variations are comparable across years. The lower and upper quartile values are 0.36 and 0.69 respectively, average across years. Note also that the inter-quartile ranges are also similar across years, fluctuating around 0.34. In sum, the cross-sectional heterogeneity of the *raw* information weighting factors at stock level justifies the estimation procedure at individual stock level. Third, I split the sample ($1984-2005$) into two sub-periods to investigate whether there are substantial changes of the analysts' information weighting factor before and after Regulation Fair Disclosures (Regulation FD). There is some evidence that Regulation FD may change how analysts aggregate information. During the pre-FD period, on average analysts systematically overweight public information ($\varphi = 0.5345$; statistically significant at 1% level); but during the post-FD period, on average the analysts seem to optimally weight public and private information ($\varphi = 0.5001$; statistically significant only at 10% level). These are consistent with Gomes, Gorton, and Madureira (2004), who suggest that the information environment of analysts changes after passage of Regulation FD.

There are also statistically significant time-series correlations between current and lagged raw quarterly analyst information weighting factors at stock level, and such correlations can be useful to characterize the herding tendency among analysts at individual stock level. The results from the estimation are shown in Table 2.2. Panel A summarizes

[15]Recent work of Bernhardt, Campello and Kutsoati (2006) suggests the measure of herding tendency used in Chen and Jiang (2006) may be biased against finding herding tendency. In turn, such bias would bias my results against finding herding tendency is related to the hypotheses examined here.

the stock-level autoregressions in the form of equation (2.2) by averaging the regression coefficients across stocks and quarters. The cross-sectional average of time-series regression coefficients are negative and statistically significant at the first four lags. The results from the full sample period, the pre-Regulation FD, and post-Regulation FD are qualitatively similar, but the autoregression coefficients from post-Regulation FD sample are slightly larger in absolute values. Moreover, there are heterogeneities among the autoregression coefficients, and such heterogeneities are quite different in pre-Regulation FD and post-Regulation FD eras. In the full sample, the inter-quartile range of first four lags are 0.35, 0.34, 0.32 and 0.33 respectively (Panel A); in the pre-Regulation FD sample, the inter-quartile range of first four lags are 0.35, 0.34, 0.32 and 0.34 respectively (Panel B); and in the post-Regulation FD sample, the inter-quartile range of first four lags are 0.48, 0.52, 0.43 and 0.44 respectively (Panel C). During the full sample and two subsample periods, the cross-sectional average of the these regression coefficients are statistically different from zero at conventional levels. The intercept terms from the regressions are statistically significantly different from 0.5 in all sample periods.

2.3.3. Analyst Herding Tendency and Stock Characteristics

2.3.3.1. Determinants of Analyst Herding Tendency Measure: Evidence from Univariate Correlations. Table 2.3 presents the time-series average of the pairwise correlation coefficients and the associated t-statistics between the analyst herding tendency measure (*IWF)* obtained in equation (2.4) and a set of firm characteristics. The frequency of the time-series is quarterly. Details on the construction of these variables can be found in Appendix A. The choice of these characteristic variable is mainly based on a set of characteristics shown to be related to future returns. Panel A uses all available data from $Q2/1985$ to $Q4/2005$. Panel B is similar to Panel A, except that it uses data from $Q1/1991$ to $Q4/2005$. The past 12-month news media coverage is only available after 1991 since the news media coverage database itself starts from January 1990. The information weighting factor is positively correlated with analyst forecast dispersion (0.11 in the full sample and 0.13 in the sub-sample), standardized unexpected earnings in percentiles (about 0.01 in both samples), the book to market equity ratio (0.09 in the full sample and 0.10 in the sub-sample), and firm size (0.11 in both samples). The information weighting factor is negatively correlated with stock turnover ratio (-0.09 in the full sample, and -0.11 in the subsample), Amihud illiquidity measure (-0.07 in the full sample, and -0.05 in the sub-sample), past 12-month cumulative return (-0.01 in the full sample, and -0.03 in the subsample), and past cumulative return in year 2 and year 3 prior to portfolio formation (-0.04 in the full sample, and -0.07 in the subsample). The information weighting factor is negatively correlated with the past 12-month news media coverage, with the correlation coefficient of 0.04.

Three observations are worth to note. First, past 12-month returns, past 2 to 3 year returns, and standardized unexpected earnings are not strongly correlated with the information weighting factor itself. Thus, some of the return spreads shown in the portfolio sorting and cross-sectional regression analysis cannot be entirely attributed to mechanically sorting on a variable strongly related to returns. Second, for a set of characteristic variables shown to be related to future returns, the correlations between these variables and the information weighting factor is not particularly strong. Third, the past 12-month news media coverage is strongly correlated with size (with a correlation coefficient of 0.43

- by far the largest correlation coefficient among all variables). Therefore, the positive correlation between the news media coverage and the information weighting factor may be spurious because both variables reflect the common size effect.[16]

2.3.3.2. The Determinants of Analyst Herding Tendency Measure: Evidence from Multivariate Regressions. Table 2.4 explores the relationship between the analyst herding tendency measure (IWF) and other variables of interests in a set of multivariate regressions. In all models, the standard errors are clustered at firm level (Rogers, 1993) following suggestions of Petersen (2006). In addition to the variables considered in the univariate correlation analysis, investment banking business related incentives, SEO and IPO history indicator variables are also included in all the multivariate regressions. These indicator variables take value of one if the stock has an IPO or SEO during past 36 months; and zero otherwise. Model 1 uses full sample from $Q2/1985$ to $Q4/2005$.[17] Following Chen and Jiang (2006), Models 2 adds the future six month turnover ratio to control for incentives related to generating trading volumes for the brokerage business. Model 3 uses the subsample data from $Q1/1991$ to $Q4/2005$ in which I also impose the constraint that the stocks must be covered by both CRSP and the news media database. In the interpretation of these regressions, we should note that lower information weighting factor values imply stronger herding tendency. The multivariate regression models reveal that the past 12-month average turnover ratio, the future 6-month average turnover ratios,

---

[16]For example, Engelberg and Gao (2006) find that size by itself explain nearly 20% of the cross-sectional variations in news media coverage.

[17]Fama-MacBeth regression delivers similar pattern but slightly more significant t-statistics. Petersen (2006) points out the t-statistics obtained from the Fama-MacBeth procedures are likely to be overstated when the time-series length is short. Because there is a truncation of the raw information weighting factor at $-1/2$ and $1/2$, I also redo the regression with Tobit model with upper and lower truncation set at $1/2$ and $-1/2$, and the results are qualitatively similar in all regression models considered in this section.

and the past 12-month news media coverage are positively related to the herding tendency (all significant at the 5% level or higher). Book to market equity, size, analyst forecast dispersions and SEO indicator variable are negatively related to the herding tendency (all significant at the 5% level or higher). Finally, the Amihud illiquidity measure, the past 12-month returns, and the past 2 to 3 year returns are all negatively associated with information weighting factor, but none of them is statistical significant at conventional levels. In summary, the analyst's tendency to herd is stronger in the population of smaller growth non-SEO stocks with more media coverage and higher share turnover ratio.

Insert Table IV About Here

2.3.3.3. Discussion of the Determinants of Analyst Herding Tendency Measure. Variables such as size and book to market equity ratios are sometimes interpreted as proxy variables capturing information uncertainty about fundamental values of the assets. Herding behaviors are more likely to occur when the decision-making setting is complex rather than simple (Blacke, Helson and Mouton 1955; Shiller, 1995; Hirshleifer and Teoh, 2003). Presumably, small capitalization and growth oriented stocks are harder to value, and so herding behaviors can be more prevalent for small growth stocks. Related to the above observation, there is some recent evidence showing that stocks with higher information uncertainty about fundamental values of the assets earn significantly more momentum profits (Jiang, Lee and Zhang, 2006; Zhang, 2006). So far this observation is exclusively interpreted as being supportive of the investor overconfidence hypotheses. The usual argument is that investors tend to be more overconfident in settings where feedback on their information or decisions is slow or inconclusive than where the feedback is clear and rapid.[18] As I have illustrated, these set of variables are also correlated with the herding tendency measure. Stronger herding tendency leads to less efficient information aggregation, which in turn generates more pronounced momentum effects. Hence, through the channel of biased financial analysts, the information uncertainty effects should be related to the momentum effects.

The relationship between the past 12-month news media coverage and the herding tendency measure is consistent with the conjectures in Shiller (1995), and Shiller and Pound (1989). Shiller and his coauthor emphasize the role of media in forming the herding tendency: when the news media coverage is extensive, it fuels the herding behaviors of investors including financial analysts. This conjecture is confirmed by the data.

Theoretical model developed by Froot, Scharfstein and Stein (1992) establishes a link between the horizon of information and horizon of investors' holding. If the horizon of the investors is short, then herding on short-term information is more likely. If the production and dissemination of information by the financial analysts cater the horizons of information and holding, which are of interests to the investors, then the analysts are likely to herd on the same set of short-run value relevant information. Empirically, one may interpret the turnover ratio as a measure inversely related to the holding horizon

[18]See Einhorn (1980) and Griffin and Tversky (1992) for early psychological evidence. However, recent work from experimental economics literature provides some new insights. For example, Biais, Hilton, Mazurier, and Pouget (2005) find their experimental subjects systematically exhibit the tendency to overestimating precision of public information, or the tendency to herd. Using psychometric measures of judgment biases and actual transaction data, Glaser and Weber (2004) find overconfidence, as defined in terms of miscalibration, is unrelated to excess trading; but as defined as "above average" (see Odean, 1998), is related to excessive trading. Hoelzl and Rustichini (2005) illustrate that when the stake is not high and the task is easy, experimental subjects exhibit overconfidence. But when task is hard and unfamiliar or when the money is at stake, experimental subjects exhibit herding tendencies. Kirchler and Maciejovsky (2002) show that in a sequential trading experiment setting, subjects only exhibit modest overconfidence in early stages of the trading but exhibit well-calibrated beliefs and sometimes herding tendencies in later stages of the trading. Presumably, earnings forecasts are complex, difficult and high stake jobs involving significant experiences over time.

of underlying shareholders (Sapienza and Polk, 2006), and infer about the horizon of value relevant information. Thus a larger the turnover ratio, the shorter the holding horizon, and the more likely both investors and financial analysts herd on the short-run information. Empirically, the turnover ratio is strongly related to the tendency to herd among analysts.

Interestingly, there exists a positive relationship between earnings forecast dispersions and information weighting factor, where the lower the value of information factor, the stronger the tendency to herd. Therefore, the higher the dispersion of analyst earning forecasts, the lower the tendency to herd. At first glance, this seems to contradict to early argument that herding is more likely during complex decision-making environment. However, this observation is actually consistent with one of the most robust findings in the opinion conformity literature of social psychology (see Blacke, Helson and Mouton, 1957; Wilder, 1977). According to this literature, forecasters evaluates the variance among the forecasts. When there is small amount of divergence of opinions, the forecasters are more likely to adopt to the majority rule, and exhibit stronger herding tendency. To the extent the forecast dispersions capture divergence of opinions as argued in Diether, Malloy and Scherbina (2002), the herding tendency and the divergency of opinions should exhibit the pattern as suggested by the opinion conformity literature.[19]

The sign of SEO indicator variable suggests the herding tendency among analysts is weaker if the firm conducts at least one seasoned equity offering during the past 36 months. Thus among recent SEO stocks, there seems to be systematic bias for the analysts to herd

[19] Another possible intepretation is that, due to the herding tendency among analysts, the dispersion of forecasts is mechanically small.

against consensus. This could be consistent with the view that analysts are more likely to rely on their private information obtained from the prior investment banking relationship.

In sum, the univariate and multivariate analyses illustrate that information weighting factor constructed from the analysts' sequential quarterly earning forecasts is closely related to several characteristic variables. Though not conclusive, it seems plausible that the information weighting factor captures the tendency to herd among the analysts.

2.4. Price Momentum Strategies and Analyst Information Weighting Factor

2.4.1. Portfolio Formation

The portfolio formation procedure is illustrated in Figure 2.1. I first compute the analyst herding tendency measure (IWF). The value of the information weighting factor will be used for all the months in quarter $(q+1)$.[20] Second, I accumulate returns (with dividends) from month $(t-11)$ to month t, and obtain the cumulative returns for individual stocks. I restrict the stocks to be traded on NYSE/AMEX/NASDAQ and the stock prices to be no less than five dollars at the end of month t to avoid the market microstructure issues induced confounding effects. Then all the stocks where I can obtain the herding tendency measure and past 12-month returns are sorted into five equally-spaced quintiles based on past returns. Within each quintile, I further sort the stocks into terciles based on the value of information weighting factor and the breakpoints for the IWF values are 0.30 and 0.70.[21] To further alleviate the liquidity issues and bid-ask bounce effects in the asset

[20] Firms may not always have regularly spaced earning announcement months so there are cases there are more than one information weighting factor values for quarter $(q+1)$. In this case, the most recent ones will be retained.

[21] Using independent sort procedure generates similar results, except some of the extreme portfolios contain fewer stocks during early period of the sample.

pricing tests, I skip a month between portfolio formation month and return accumulation month. To increase the power of the tests, the price momentum strategies I examine include portfolios with overlapping holding periods as in Jegadeesh and Titman (1993). In any given month τ, the strategies contain a series of portfolios that are selected in the current month τ, as well as previous $J-1$ months, where J is the holding period. The portfolios are constructed by the past J month return ranking. That is, in each month τ, the strategy closes out the long/short positions initiated in month $\tau - J$. Therefore, the managed portfolio revises its weights on $1/J$ of the securities in the entire portfolio in any given month, and carries the rest from previous months. In this study, I focus on the return accumulation period of 12 months ($K = 12$) and portfolio holding horizon of 3 and 6 months ($J =$ 3,6), but also examines portfolio holding horizon of 9, 12, 24 and 36 months.

Insert Figure 1 About Here

Finally, for simplicity, I introduce the following notations: $R1$ (losers) and $R5$ (winners) are the portfolio quintiles with the lowest and highest past returns; $IWF1$ (strong herding tendency) and $IWF3$ (weak herding tendency, or anti-herding tendency, or exaggeration of differences) denote the portfolio terciles with the lowest and highest values of the information weighting factor values. $\{R5 \cap IWF1\}$, $\{R1 \cap IWF1\}$, $\{R5 \cap IWF3\}$, $\{R1 \cap IWF3\}$ denote winner/high herding tendency, loser/high herding tendency, winner/ low herding tendency and loser/low herding tendency portfolios respectively.

2.4.2. Summary Statistics of Basic Portfolio Characteristics

Table 2.5 reports some summary statistics of basic portfolio characteristics. Panel A shows that there is significant difference in past cumulative excess returns between past winners ($R5$) and losers ($R1$) conditional on each level of information weighting factor values ($IWF1$ to $IWF3$). On the other hand, conditional each return level ($R1$ to $R5$), there is significant difference in information weighting factor values, so the double sorting procedure generates sufficient variations in the past returns and information weighting factor values. Furthermore, conditional on past returns, sorting on information weighting factor values generate a small spreads among past winners ($R5$) and past lowers ($R1$), but such differences are relatively small - about 2.92 percent among past losers and 4.1 percent among past winners (both numbers refer to the difference in 12 month cumulative returns). In contrast, the difference between $R5$ and $R4$ ($R4$ being the portfolio of stocks adjacent to past winners), or $R1$ and $R2$ ($R2$ being the portfolio of stocks adjacent to past losers) are well above 20 percent.

The time-series average stock price in each of the portfolio is 20.88 dollar per share (past loser/high herding stocks), and the maximal average price is 51.02 dollar per share (past winner/low herding stocks). The time-series average of the percentage market capitalization for each of these 15 double sorted portfolios are reported in the fourth column of Panel A. The monthly percentage market capitalization is computed as the total monthly market capitalization of stocks belonging to each of the 15 portfolios divided by the total monthly NYSE/AMEX/NASDAQ common stocks total market capitalization. Overall, the extreme portfolios only contains a small percentage of total market capitalization - approximately 7 percent for past losers and 13 percent for past winners.

Finally, note that the portfolios are relatively well diversified. On average, it has 33 stocks in the extreme portfolios. Panel B reports the time-series average of individual stock's market capitalization (in percentile ranking and dollar values respectively) for each portfolio. The stocks used in this paper are rather large: the minimal average market capitalization is 150 million, which is above the 50th percentile of NYSE size percentile breakpoints. As my sample of stocks are well above last two size deciles of NYSE market capitalization breakpoints, short-sale constraints play a rather limited role according to D'Avolio (2002).

Figure 2.2 illustrates the percentage of the market capitalization of the sample of stocks in the CRSP - I/B/E/S intersection to the total market capitalization (as measured by all CRSP common shares traded on NYSE, AMEX and NASDAQ). Due to better coverage of I/B/E/S over time, the total number of stocks in the CRSP - I/B/E/S intersection increased substantially. The early year coverage is relatively small, only 26.18% in 1985, but quickly increase to about 70% since early 1990's.

2.4.3. Simple Price Momentum Strategy

I first investigate the magnitude and statistical significance of simple price momentum strategy in my sample of firms where I have both past returns and estimates of analysts herding tendency measures. Panels A and B of Table 2.6 confirm that there is momentum effect in my sample period between April 1985 and December 2005, particularly for holding horizon of three to six months.

During the period of three months after portfolio formation, the winner portfolio (top quintile) on average earns about 1.74% per month ($J = 3$) and 1.67% per month ($J = 6$)

, while the loser portfolio (the bottom quintile) on average earns about 1.10% per month ($J = 3$) and 1.22% per month ($J = 6$) . The hedged portfolio by taking long positions on the winners and taking short positions on the losers earns about 63 basis per month (t-statistic = 1.76) for three month holding horizon, and 45 basis points (t-statistic = 1.30) for six month holding horizon. Jegadeesh and Titman (2001) report that small capitalization (below the median NYSE market capitalization) and large capitalization (above the median NYSE market capitalization) momentum portfolios earn around 1.65% and 0.88% per month between 1990 and 1998. Therefore, even these returns from my sample are slightly smaller, they are still in line with large capitalization momentum portfolio returns reported by Jegadeesh and Titman (2001).

Consistent with Jegadeesh and Titman (1993), there are strong January seasonalities in momentum portfolio returns in my sample. Though the January returns are not statistically different from zero (t-statistic = -0.86 for $J = 3$ and t-statistic = -1.23 for $J = 6$) in my sample, the magnitudes are quite large (-1.21% per month for $J = 3$ and -1.34% for $J = 6$) and the signs are noticeably negative, which drag down the overall monthly average returns. Excluding January returns improves the profitability of momentum strategies. Between 1985 and 2005, the average February to December returns of momentum portfolio is 79 basis points (t-statistic = 2.17) for three month holding horizon, and 60 basis points for six month holding horizon.[22]

[22]These magnitudes are also similar to what is reported in Jagadeesh and Titman (1993 and 2001). In table IV of Jagadeesh and Titman (1993), they report the January momentum portfolio return of -7.97%, -3.47, and -1.61% for the bottom, middle, and top one-third market cap stocks between 1965 and 1989. There is some evidence of dissappearing “January effect”. For example, In table II of Jagadeesh and Titman (2001), they report the average January return attenuates to -1.24% across all NYSE/AMEX/NASDAQ stocks, compared to relatively pronouced “January effect” in early period around -4.35% ($= \frac{(-7.97\%)+(-3.47\%)+(-1.61\%)}{3}$), based on table IV of Jagadeesh and Titman (1993)) across all NYSE/AMEX/NASDAQ stocks.

The weaker momentum profits in my sample are primarily due to the difference between my sample and Jegadeesh and Titman (1993, 2001) sample. The firms in my sample are large and well-covered stocks, and they are in later periods (between 1985 and 2004). The empirical evidence in this paper is also consistent with the main finding in Hong, Lim and Stein (2000) who show that simple price momentum strategy does not work well with the portfolio of large stocks well-covered by financial analysts.

Panel A, Table 2.7 reports the factor model adjusted returns. Model 1 is the Fama-French three-factor model,

$$R_t^{MOM} = \alpha + m \times MKTRF_t + s \times SMB_t + h \times HML_t \tag{2.5}$$

where $MKTRF_t$, SMB_t, and HML_t are the market excess return, small-minus-big and high-minus-low factors respectively. R_t^{MOM} is the monthly returns from the momentum portfolios. The factor-adjusted returns of the winner minus loser portfolios with three month holding horizon are about 95 basis points per month and statistically significant (t-statistic $= 2.58$), which are about one-third higher than the mean portfolio spreads reported early. For six month holding horizon, it is about 83 basis points per month, and remain highly significant. The momentum portfolio loads little on the SMB factor and not statistically significant. The lack of loadings on SMB factor is likely due to that my CRSP-I/B/E/S sample contains relatively large market capitalization stocks. The portfolio loads modestly on the market factor but quite significantly on HML factor. For example, for three month holding horizon, MKTRF is -0.2442 with t-statistic $= -2.66$, -0.4589 on HML with t-statistic $= -3.41$ (untabulated). The factor loadings on HML factor reveals the momentum strategy is a contrarian strategy relative to the value/growth strategy, and it is also contrarian relative to the market portfolio returns.[23] Overall, the low book to market equity stocks earn higher momentum returns, which is consistent with the evidence in Asness (1997), and Daniel and Titman (2000). In my sample, Fama-French three factor regression does not completely eliminate the January seasonality but reduce its magnitudes. The regression-based "alpha" measure of abnormal return is even slightly bigger (1.11% and 0.98% per month for $J = 3$ and 6) for non-January months.

Pástor and Stambaugh (2003), and Sadka (2006) show liquidity risk adjustment attenuates returns from the price momentum strategy. Therefore, I consider the Fama-French three factor model augmented with the liquidity risk factor, where the liquidity risk factor is obtained from Sadka (2006),

$$R_t^{MOM} = \alpha + m \times MKTRF_t + s \times SMB_t + h \times HML_t + l \times LIQ_t \tag{2.6}$$

where $MKTRF_t$, SMB_t, HML_t and LIQ_t are the market excess return, small-minus-big, high-minus-low, and liquidity factors respectively. R_t^{MOM} is the monthly returns from the momentum portfolios.[24] Panel A, Table 2.5 shows the four-factor adjusted abnormal return attenuates from 95 basis points to 63 basis points for 3-month holding horizon ($J = 3$), and from 83 basis points to 54 basis points for 6-month holding horizon($J = 6$). Similar factor adjusted returns of simple momentum strategy also show up in non-January months. The abnormal return attenuates from 1.11 percent per month to 77 basis points

[23]Cooper, Gutierrez, and Hameed (2004) find that the momentum profits is positive after positive market returns, and negative after negative market returns. This is consistent with the notion that the simple price momentum strategy is a contrarian strategy relative to the market.

[24]Specifically, I use the variable components of the liquidity risk factor in Sakda (2006). Using both variable and fixed component of the liquidity risk factor generates similar results. I also experimented liquidity risk factor in Pástor and Stambaugh (2003), and the results are similar.

per month for 3-month holding horizon ($J = 3$), and from 98 basis points to 65 basis points for 6-month holding horizon ($J = 6$).

In summary, my sample of stocks tilt towards large market capitalization and well-covered stocks by construction. However, the sample exhibit main characteristics of the typical price momentum portfolio, even though the momentum effect is weak.

2.4.4. Price Momentum Strategy Interacted with Herding Tendency Measure

2.4.4.1. Concentration of Momentum Effects. Now we consider the simple price momentum strategy interacted with the herding tendency measure (IWF). Depicted in Table 2.6, several interesting patterns emerge that hold up throughout my subsequent analysis. Let us begin by first focusing on three month as the holding horizon. First, the bulk of the momentum profits come from the momentum portfolios ($R5 - R1$) with the highest herding tendency ($IWF1$). For the holding period of three months, the high herding tendency momentum portfolio, $R5 \cap IWF1 - R1 \cap IWF1$, on average earns 1.14 percent per month with a t-statistic of 2.86. In sharp contrast, the low herding tendency momentum portfolio, $R5 \cap IWF3 - R1 \cap IWF3$, on average only earns 0.29 percent per month with a t-statistic of 0.79. The median IWF momentum portfolios, $R5 \cap IWFW2 - R1 \cap IWF2$, on average earn somewhere in between, 0.50 percent per month with a t-statistic of 1.30. The economic magnitude is clearly important. The returns from the high herding tendency momentum portfolios are roughly 2 times and 4 times of those in the median and lowest herding tendency momentum portfolios respectively. In contrast to the simple price momentum strategy, the price momentum strategies interacted with the herding tendency measure exhibit little January seasonality. All of the price momentum portfolios earn negative return during January months (not reliably different from zero), but the momentum portfolio with the highest herding tendency loss much less (-33 basis points, t-statistic $= -0.22$). For the rest of the year, the monthly returns preserve the pattern of the average returns reported in all month returns. The momentum portfolios with the highest herding tendency, $R5 \cap IWF1 - R1 \cap IWF1$, earn at least twice as much as the rest of the portfolios (1.26 percent per month for $R5 \cap IWF1 - R1 \cap IWF1$, versus 0.70 percent per month for $R5 \cap IWF2 - R1 \cap IWF2$, and 0.44 percent per month for $R5 \cap IWF3 - R1 \cap IWF3$.) The difference in returns between the momentum portfolio with strong herding tendency and momentum portfolio with anti-herding ranges from 74 basis points for 3-month holding horizon (t-statistics $= 3.08$) to 45 basis points for 12-month holding horizon (t-statistics $= 2.34$). Looking at other holding horizons ranging from $J = 9$ to $J = 36$ months, the momentum profits seems to be concentrated within the first six months after portfolio formation even for the momentum portfolio interacted with the herding tendency measure.

At this point, it is also useful to look at the intercept term and the factor loadings on the market, SMB, HML and liquidity risk factors. First, similar to the simple price momentum strategies discussed above, Panel B of table 2.7 suggests that the portfolio returns from the simple price momentum portfolios interacted with the herding tendency measure cannot be explained by the three factor model. Including returns throughout the year, the different momentum portfolios interacted with the herding tendency measure earn 1.42 percent ($R5 \cap IWF1 - R1 \cap IWF1$), 0.89 percent ($R5 \cap IWF2 - R1 \cap IWF2$) and 0.54 percent ($R5 \cap IWF3 - R1 \cap IWF3$) per month after adjusting returns by the Fama-French three-factor model. These time-series regression results are also consistent with the early results based on the raw returns - the momentum profits are concentrated in the portfolios where the analysts on average exhibit herding tendencies. The non-January portfolio returns are also consistent with the results from all month portfolio returns. Excluding January returns, the different momentum portfolios interacted with the measure of herding tendency earn 1.58 percent ($R5 \cap IWF1 - R1 \cap IWF1$), 1.07 percent ($R5 \cap IWF2 - R1 \cap IWF2$) and 0.68 percent ($R5 \cap IWF3 - R1 \cap IWF3$) per month. The price momentum portfolios interacted with the herding tendency measure load significantly on HML factor, with t-statistics of -2.78, -3.86 and -2.86 for $R5 \cap IWF1 - R1 \cap IWF1$, $R5 \cap IWF2 - R1 \cap IWF2$, and $R5 \cap IWF3 - R1 \cap IWF3$ portfolios respectively. The magnitudes of loadings are similar among these different momentum portfolios.

In addition to those three factors suggested by Fama and French (1993), the model with an additional liquidity factor (Pastor and Stambaugh, 2003; Sadka, 2006) reduce the momentum portfolios interacted with the herding tendency measures. In all cases, the magnitude of returns decreases in each of the three spread series, but only remain economically meaningful and statistically significant in the momentum portfolio where herding tendency among analysts is strong, or ($R5 \cap IWF1 - R1 \cap IWF1$). When liquidity risk factors are included, the abnormal returns of the momentum portfolios with medium and high information weighting factor values are within the range of 30 to 60 basis points per month and statistically insignificant. But, the momentum portfolio with strong herding tendencies among the analysts following the stock still earns 1.08 percent per month for 3-month holding horizon - in contrast to 1.42 percent per month using Fama-French three-factor model as adjustment and the abnormal returns are highly significant (t-statistic $= 2.90$). The factor loadings on liquidity factors are positive and generally significant with comparable magnitudes across different momentum portfolios interacted with herding tendency measure.

Is there any reversal effects of momentum strategy returns? Jegadeesh and Titman (2001) investigate momentum strategy's return at longer horizon - as long as five years. I choose to focus on three-year horizon because at longer horizon, the long-run reversal effects may mask the intermediate-term momentum effects (see Moskowitz and Grinblatt, 1999 for related discussions). Let us first consider the momentum portfolios with the strongest herding tendencies among analysts. This is the portfolio with the highest momentum return. The average monthly return, *including January returns*, is 1.14%, which implies three-month cumulative return of $(1 + 1.14\%)^3 = 103.45\%$ measured from portfolio formation month. Similarly, the cumulative returns for 6-, 9-, 12-, 24- and 36-month holding horizons, *including January returns*, are 103.45%, 105.19%, 105.39%, 102.31%, 103.59% and 100.97% respectively. Even though, as reported in Panel D of Table 2.6, the returns from the momentum strategy within the high herding tendency portfolio are almost exclusively confined within the first 12 months, and statistically significant only within the first 6 months, the cumulative returns do seem to revert to initial value by the end of the third year. However, such reversal effects are mainly driven by January returns. The cumulative returns, *excluding January returns*, are 103.84%, 106.04%, 107.01%, 106.94%, 106.64% and 109.73% for 3-, 6-, 9-, 12-, 24- and 36-month holding horizons respectively. Interestingly, even for the momentum portfolio with the anti-herding tendency among analysts, where there is no price momentum effects during the first 6 months, there are also long-run reversal effects. The cumulative returns, *including January returns*, are 100.87%, 100.78%, 100.04%, 99.22%, 91.07% and 98.27% for 3-, 6-, 9-, 12-, 24- and 36-month holding horizons respectively. The cumulative returns, *excluding January returns*, are 101.32%, 101.80%, 100.04%, 100.18%, 94.37%, and 94.22% for 3-, 6-, 9-, 12-, 24- and 36-month holding horizons respectively. Given the above evidence, two comments are in place: First, there are strong return seasonalities in momentum returns, even when we focus on large capitalization stocks. Second, at least in my sample, the conclusions about long-run returns reversal effects of momentum strategy returns depend on whether January returns are included or excluded.[25]

2.4.4.2. "Within-News" Herding Effects. Table 2.8 reports the existence of different future price movements among stocks which share similar prior news but differ in the extent of analyst herding tendencies. This is the second key finding of this paper. For the subset of stocks with good news in the past and strong herding tendency among analysts (low IWF portfolio), the upward price continuation is stronger than the subset of stocks where analysts do not herd (high IWF portfolio). In contrast, for the subset of stocks with bad news in the past and strong herding tendency among analysts who follow the stock, the downward drifts in prices are more pronounced than the subset of stocks where the analysts do not herd. Such tendency is particularly strong among the loser portfolios. In the lowest past return quintile ($R1$), high information weighting factor portfolio ($R1 \cap IWF3$) *outperforms* the low information weighting factor portfolio ($R1 \cap IWF1$) by 54 basis points per month in the next three months, and the spread is highly significant (t-statistic $= 2.74$). In the highest past return quintile, high information

[25]It seems that some of the long-run effects documented in the literature may be related to the return seasonalities and tax avoidance behaviors studied in Grinblatt and Moskowitz (2004).

weighting factor portfolio ($R5 \cap IWF3$) *underperforms* the low information weighting factor portfolio ($R5 \cap IWF1$) by 31 basis points per month in the next three to six months and the spread is marginally significant (t-statistic $= -1.73$ for 3 months and t-statistic $= -1.97$ for 6 months). There are some January seasonalities in the returns from the "within-news" strategy: Among the past losers, return from the "within-news" strategy remains positive (though not statistically significant) for such strategy with three and six month holding horizons. It nevertheless becomes negative beyond holding horizon of six months. Among the past winners, the January returns are particularly negative, but again statistically insignificant for holding horizon within a year.

Based on the patterns of raw returns, one may suspect that the "within-news" strategy is driven by some liquidity effects, so the adjustments by liquidity or reversal factors become relevant. These factors help to understand the nature of the returns for the "within-news" strategy. Table 2.9 reports the Fama-French three-factor model (Model 1), three-factor model with a liquidity risk factor (Model 2), and three-factor model with short-run reversal factor (Model 3). The short-run reversal factor is constructed based on the testing portfolios in Jegadeesh (1992) and Lehmann (1992). The factor models with short-term reversal factor are specified as

$$R_t^{MOM} = \alpha + m \times MKTRF_t + s \times SMB_t + h \times HML_t + v \times STREV_t \tag{2.7}$$

where $STREV$ is the "short-run reversal factor". It is constructed based on the testing portfolios construction in Jegadeesh (1992) and Lehmann (1992). Specifically, it is the return spreads between stocks in the top one third of the previous month's return terciles minus the bottom one third of the return tercile. Using model specifications 1 to 3, all these regressions basically lead to the same conclusion. First, the factor regression adjusted returns from the "within-news" strategy among past losers ($R1 \cap IWF3 - R1 \cap IWF1$) are about 50 basis points per month, and the factor regression adjusted returns from the "within-news" strategy among past winners ($R5 \cap IWF1 - R1 \cap IWF3$) are about 29 basis points per month. Second, the the reversal factor's factor loadings on these "within-news strategies" are small. Third, the liquidity factor's factor only loads statistically significant with modest magnitudes on the returns from "within-news strategy" among past losers ($R1 \cap IWF3 - R1 \cap IWF1$) but not past winners.

In summary, there is some evidence suggesting that the herding tendency measure is related to future returns. The winners and losers portfolios are not heterogeneous: the future return continuation is contingent on the extent of herding tendency. Finally, there seems to be an asymmetric effect of herding tendency and future returns. The "within-new" strategy returns are much larger among losers. This is also related to the analyst herding tendency. For example, analysts certainly have incentives not to alienate themselves further from the management of the company if the performance of the company's stocks has been poor. If they have valuable information which should warrant their disclosures, they may choose to withhold it.[26]

2.5. Further Discussion of Portfolio Characteristics

2.5.1. Construction of Portfolio Characteristics

In this section, I present a number of additional portfolio characteristics and evaluate potential alternative interpretations of my early empirical evidence. The details of the

[26] I look at the "within-news" effects after Regulation Fair Disclosures, and did not find there are significant changes in returns for past losers portfolio though.

data construction can be found in Appendix C. For each stock in the portfolio, I update the characteristic information monthly if possible. I average across stocks to compute portfolio level characteristics, and then average the resulting measures across time. Table 2.10 reports these characteristics associated with each portfolio. Instead of explaining each characteristic separately, I formulate several alternative explanations of my early empirical findings and organize these portfolio characteristics to evaluate these alternative explanations.

2.5.2. Discussions of Related Hypotheses Based on Portfolio Characteristics

Panel A of Table 2.10 shows the profitable momentum portfolio indeed has the lowest level of liquidity. In past loser portfolio, the low and high analyst information weighting factor stocks on average have past 12-month average Amihud illiquidity measure values of 0.10 versus 0.06. In the past winner portfolio, the low and high information weighting factor stocks on average have past 12-month average Amihud illiquidity measure values of 0.05 versus 0.03.[27] Past winners generally have higher liquidity level than past losers, which could help to explain why liquidity risk factor loads more significantly for the "within-news" herding strategy returns among the past losers but not past winners.

Second, momentum portfolios earn higher returns when analysts have stronger herding tendencies. The stocks with stronger analyst herding tendencies are also stocks with higher level of information uncertainty effects. They have smaller market capitalizations

[27]Results are similar, if the illiquidity is measured at portfolio formation month. In past loser portfolio, the low and high analyst information weighting factor stocks on average have Amihud illiquidity measure values of 0.1160 versus 0.0687. In the past winner portfolio, the low and high analyst information weighting factor stocks on average have Amihud illiquidity measure values of 0.0337 versus 0.0254 (not tabulated).

(Panel B, Table 2.10), lower book to market equity ratios (Panel C, Table 2.10), higher return volatilities (Panel E, table 2.10), but lower analyst dispersion (Panel B, Table 2.10). These portfolio characteristics are consistent with those results shown early in the univariate correlation and multivariate regression analyses of the herding tendency measure and firm characteristics. Differences among these information uncertainty variables are probably not large enough to completely explain return differences, and the herding tendency measure contains much more information than just information uncertainty.

Third, the concentration of momentum portfolio returns within the high herding tendency stocks is not driven by the earnings surprises. The herding tendency measure is constructed by comparing the probability that the forecast errors and deviation from consensus forecasts have the same sign, and filter these probabilities through an autoregressive (AR) models based on historical information, so there is no reason *a priori* to believe such measure picks up earnings momentum. Chan, Jegadeesh and Lakonishok (1996) show that returns and earning momentum effects are closely related to future momentum profits, but neither of these two subsumes one or the other. Panel B reports the median earning surprises are *minus* 0.61 cents for past losers / high information weighting factors, and *minus* 1.63 cents for past losers / low information weighting factors. In contrast, the median earnings surprise is 1.04 cents for past winners / low information weighting factor, and 1.26 cents for past winners / high information weighting factor.

Closely related to the above discussion of earning surprises and momentum effect, there is a concern about whether the information weighting factor constructed here picks up some sort of "optimism" or "pessimism" biases of the analysts. The "optimism" bias is the tendency to over-forecast the actual earnings. The "pessimism" bias is the tendency to under-forecast the actual earnings. If we take the earning surprise measures as proxy variables for optimism and pessimism biases, there is little evidence such biases are related to the results in the paper. Although the optimism and pessimism biases are similar across momentum portfolios cut by the information weighting factor, these portfolios generate dramatically different momentum profits. It also is not clear why stocks with similar level of optimism and pessimism bias would experience different future price changes. The empirical evidence here is also consistent with the findings in Jagannathan, Ma, and Silva (2005), who investigate how "optimism" or "pessimism" biases are related to momentum strategy returns among other anomalies. They also do not find any evidence showing the "optimism" or "pessimism" biases are related to momentum strategy returns.

Fourth, share turnover ratio and size adjusted past 12-month news coverage are shown to be correlated with the herding tendency measure. Since news media coverage is highly correlated with market capitalization, I construct a size adjusted news media coverage variable. At each month, I first sort all stocks in the news media database and CRSP into ten portfolios based on their NYSE market capitalization decile ranking. Then I compute the monthly average news coverage for each of size portfolio. Because news coverage can be extremely skewed, when I compute the size portfolio and individual stock's news coverage, I use the logarithm of one plus the number of news coverage. At last, I subtract the monthly average news coverage of each size portfolio from individual stock's news media coverage that month (after taking the logarithms). To further eliminate any one firm's impact, I consider the median coverage within each portfolio. Therefore, the time-series average of news media coverage is the time-series average of the median of news coverage adjusted by size. At portfolio level, such characteristics bear out (see the last column of Panel A, and Panel C). The momentum portfolios with the highest level of herding tendency have relatively large turnover ratio and size adjusted news coverage.[28]

Fifth, if the momentum profits are concentrated in a subset of stocks where the market aggregates the public and private information poorly, is it necessarily the case that arbitrages on these subset of stocks become easier? The answer is negative. Panel E shows that the momentum portfolio with lowest analyst information weighting factor has the highest values of idiosyncratic volatilities, and total volatilities, proxied by ARBRISK and VOLA respectively. In past loser portfolio, the low and high analyst information weighting factor stocks on average have arbitrage risk (ARBRISK, or the residual variance from the market model) values of 0.84 versus 0.68. In the past winner portfolio, the low and high analyst information weighting factor stocks on average have arbitrage risk values of 0.75 versus 0.61. The overall stock return volatilities follow a similar pattern. In past loser portfolio, the low and high analyst information weighting factor stocks on average have overall volatilities of 0.9954 versus 0.8184. In the past winner portfolio, the low and high analyst information weighting factor stocks on average have arbitrage risk values of 0.8831 versus 0.7236. Based on several portfolio characteristics, several factors may prevent risk-averse arbitrageurs from investing in, so there is a good reason to believe that the "limits of arbitrage" argument (Shleifer and Vishny, 1997) is at work. Indeed, risk associated with investing in the subset of high momentum profit stocks is substantial and highly idiosyncratic, and the cost of trading is relatively high.

[28] One should note that the size adjusted news media coverage can be negative by construction. Interestingly, Panel C in Table 2.5 also reveals a bias in media coverage. Past winners (portfolio R5) is associated with *negative* size adjusted news media coverage, while the past losers (portfolio R1) is associated with *positive* size adjusted news media coverage. The difference is statistically significant at 1 percent level. Engleberg and Gao (2006) also finds such pattern, where news media coverage is strongly related to the direction of past returns rather than magnitudes of past returns.

2.6. Characteristic Regressions

As discussed above, there are many variables potentially correlated with the herding tendency measure, and some of them are shown to be able to predict future returns. De Bondt and Thaler (1985) find long-run reversal effect in returns, and long-run reversal effect may attenuate the intermediate term past return net contribution to future returns. In addition, analysts' forecast dispersions may capture the difference of opinions, so high dispersion stocks may earn lower returns if optimistic investors take the upper hands (Diether, Malloy and Scherbina, 2001). Momentum strategy returns are related to market capitalization (Jegadeesh and Titman, 2001; Hong, Lim and Stein, 2000), share turnover ratios (Lee and Swaminathan, 2000), book to market equity ratios (Asness, 2001; Daniel and Titman, 2000), level of liquidity (Korajczyk and Sadka, 2004; Lesmond, Schill and Zhou, 2004), analysts' forecast dispersions (Zhang, 2006). To control for potential earnings momentum effects, I add the earnings momentum, the standardized unexpected earnings (SUE) as a control variable. Some of the analysts' biases could be driven by the incentives to generate investment banking businesses (Lin and McNichols, 1998; Michaely and Womack, 1999) or trading commissions (Chen and Jiang, 2006). Finally, to mitigate these possible incentive related reasons for analysts' herding tendency, I add the history of the seasoned equity offerings (SEOs), past and future share turnover ratios as additional control variables.[29]

[29] I also consider the history of initial public offering (IPO) by constructing an IPO indicator variable in the characteristic regression. Since majority of the firms in my sample are seasoned firms (age since the initial public offering greater than 60 months), it's not surprising that this variable is statistically insigificant in all regressions.

2.6.1. Characteristics Regressions: Concentration of Momentum Strategy Returns

It remains to answer whether information weighting factor makes marginal contribution to the price momentum strategy returns after controlling for these known effects. There are two obstacles to separate out the marginal contribution of any particular predictive variable. First, sorting procedure allows nonlinearity in relating predictive variables to future returns, but if there are many potential variables to control for, multiple-way sorting procedure becomes infeasible with limited data. Second, linear regression approach may mask important nonlinearity in the data, which can overstate or understate the marginal contribution of one particular predictive variable. To tackle these issues, I adopt the characteristic regression approach in Brennan, Chordia and Subrahmanyam (1998). To further accommodate nonlinearity, I use indicator variables to allow the slope coefficients of past 12-month return interacting with the information weighting factor to be different across different tercile of information weighting factor values. Essentially, I attempt to combine the portfolio sorting approach with the characteristic regression approach.

Model 1: The baseline specification of the cross-sectional regression model (denoted as Model 1) is in (2.8),

$$R_{i,t+2,t+2+J} = \begin{array}{l} \beta_1\, Past\ Ret + \beta_2\, Past\ LT\ Ret + \beta_3\, BM + \beta_4\, Past\ TO \\ +\beta_5\, Size + \beta_6\, \overline{ILLQ} + \beta_7\, DISP + \beta_8\, SUE + \beta_9\, SEO \end{array} \tag{2.8}$$

where all independent variables are winsorized at 99.5% and 0.5% levels (across all months and all observations), then I demean and standardize them by each variable's cross-sectional standard deviations (month by month). The dependent variable is also demeaned and standardized similarly. For notation simplicity, I drop the subscript i for individual stock, and t for individual month for the dependent variables. To avoid market microstructure induced noise, there is one month lag between stock return accumulation month and portfolio formation month. $R_{i,t+2,t+2+J}$ is the cumulative returns with J-month holding horizon between month $(t+2)$ and month $(t+2+J)$, where month (t) is the portfolio formation month. *Past LT Ret* is the cumulative returns between month $(t-35)$ and month $(t-12)$. *Past Ret* is the cumulative returns between month $(t-11)$ and portfolio formation month (t). BM is the book to market equity ratio, where the book value of equity is computed based on the definition in Fama and French (2001), and the market equity is the market capitalization of the stock during portfolio formation month. *Past TO* is the share turnover ratio between month $(t-11)$ and month (t). $Size$ is the logarithm of the market capitalization (shares outstanding in $1000's \times$ share price) at the portfolio formation month. $\overline{ILLQ}$ is the NASDAQ volume-adjusted Amihud illiquidity measure, where the volume adjustment procedure follows Atkins and Dyl (1997, 2005). $DISP$ is the average analyst forecast dispersions at the most recent quarter as of month t. SUE is the latest standardized unexpected earnings in percentile ranking up to the portfolio formation quarter.[30] SEO is an indicator variable taking value of one if there is a seasoned equity offering (SEO) by the firm during any month between $(t-35)$ and (t).

Model 2: In Model 2, I consider each herding tendency measure stratified portfolio's momentum profits. This model allows different slope coefficients for past returns (*i.e.*

[30]Using alternative form of earning surprise, such as measuring SUE with a drift term or constructing SUE as the actual minus the analyst consensus forecasts of quarterly earnings do not change the conclusions.

$\{\phi_j\}_{j=1}^{j=3}$), conditional different levels of herding tendency measures. To assess whether information weighting factor is directly related to future returns, I separately estimate information weighting factor's predictive power on future returns at different levels of its own value (*i.e.* $\{\eta_j\}_{j=1}^{j=3}$). Finally, the regression model also controls for the "level effects" of future returns due to possible difference in herding tendencies (*i.e.* $\{\gamma_j\}_{j=1}^{j=3}$). The regression model is specified as (2.9),

(2.9)

$$R_{i,t+2,t+2+J} = \begin{array}{l} \sum_{j=1}^{j=3} \phi_j R_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) + \sum_{j=1}^{j=3} \eta_j IWF_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) \\ + \sum_{j=1}^{j=3} \gamma_j \mathbf{I}(IWF_{i,t} = j) \end{array}$$

where $\mathbf{I}(IWF_{i,t} = j)$ is an indicator variable, which takes the value of 1 if portfolio formation month's information weighting factor value belongs to the j-th tercile (where j = 1, 2 and 3), and takes the value of 0 otherwise. The tercile ranking of individual stock's information weighting factor is obtained from *independent* sorting on all stocks at the portfolio formation month. Under the "momentum concentration hypothesis", the null hypothesis is that momentum profits concentrate in "herding stocks", *i.e.*, $\phi_1 \neq 0$ but $\phi_3 = 0$. The information aggregation story suggests that there is momentum effects if there is inefficient aggregation of public and private information, though in the absence of any news, the information weighting factor should not predict future returns. Under this prediction, the null hypothesis is that $\eta_j = 0$ (no incremental predictive power of herding tendency on future returns) and $\gamma_j = 0$ (no predictive power from herding tendency in terms of level of herding on future returns), where $j = 1$, 2 and 3.

Model 3: Model 3 in (2.10) is similar to Model 2, but it controls for other characteristics shown to be related to future returns:

(2.10)

$$R_{i,t+2,t+2+J} = \begin{array}{l} \beta_1 \; Past\ LT\ Ret + \beta_2 \; BM + \beta_3 \; Past\ TO + \beta_4 \; Size + \beta_5 \; \overline{ILLQ} + \\ \beta_6 \; DISP + \beta_7 \; SUE + \beta_8 \; SEO + \sum_{j=1}^{j=3} \phi_j R_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) \\ + \sum_{j=1}^{j=3} \eta_j IWF_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) + \sum_{j=1}^{j=3} \gamma_j \mathbf{I}(IWF_{i,t} = j) \end{array} .$$

Model 4: Model 4 in (2.11) adds one extra variable, future share turnover ratio ($Future\ TO$) corresponding to the return accumulation period in the the regression model (2.10).

(2.11)

$$R_{i,t+2,t+2+J} = \begin{array}{l} \beta_1 \; Past\ LT\ Ret + \beta_2 \; BM + \beta_3 \; Past\ TO + \beta_4 \; Size + \beta_5 \; \overline{ILLQ} + \beta_6 \; DISP + \\ \beta_7 \; SUE + \beta_8 \; SEO + \beta_{10} \; Future\ TO + \sum_{j=1}^{j=3} \phi_j R_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) \\ + \sum_{j=1}^{j=3} \eta_j IWF_{i,t-11,t} \times \mathbf{I}(IWF_{i,t} = j) + \sum_{j=1}^{j=3} \gamma_j \mathbf{I}(IWF_{i,t} = j) \end{array}$$

2.6.2. Characteristics Regressions: "Within-News" Strategy Returns

Model 5: Model 5 tests the "within-news" herding effects in returns. Specifically, we test whether conditional the level of on past returns, stocks with different levels of herding tendencies experience different future returns after controlling other variables known to predict future returns, especially past returns. I consider the regression model (2.12) of the form,

(2.12)

$$
R_{i,t+2,t+2+J} = \begin{array}{l} \beta_1 \; Past\ LT\ Ret + \beta_2 \; BM + \beta_3 \; Past\ TO + \beta_4 \; Size + \beta_5 \; \overline{ILLQ} + \beta_6 \; DISP + \\ \beta_7 \; SUE + \beta_8 \; SEO + \beta_{10} \; Future\ TO + \sum_{j=1}^{j=5} \kappa_j IWF_{i,t-11,t} \times \mathbf{I}(R^*_{i,t-11,t} = j) \\ + \sum_{j=1}^{j=5} \varphi_j R_{i,t-11,t} \times \mathbf{I}(R^*_{i,t-11,t} = j) + \sum_{j=1}^{j=5} \xi_j \mathbf{I}(R^*_{i,t-11,t} = j) \end{array}
$$

where $\mathbf{I}(R^*_{i,t-11,t} = j)$ is an indicator variable taking value of one if the past 12-month return ranking $(R^*_{i,t-11,t})$ belongs to the j-th quintile ranking; and zero otherwise. The $IWF_{i,t-11,t}$ is the portfolio formation month information weighting factor value. Other variable definitions are similar to early characteristic regressions. Empirically, I test whether $\kappa_1 > 0$ and $\kappa_5 < 0$, the hypotheses that among past winners, the stocks with stronger herding tendencies of the analysts should outperform the stocks with weaker herding tendencies; but among past losers, the stocks with stronger herding tendencies should underperform the stocks with weaker herding tendencies. Finally, aimed at controlling for the "level effects" of past return on future returns, the indicator variables $\mathbf{I}(R^*_{i,t-11,t} = j)$ are included in the regression as control variables, where $j = 1, 2, 3, 4$ and 5.

2.6.3. Calculation of Standard Errors

The calculation of standard errors of the parameters in these characteristic regressions deserve some comments. A standard approach to compute standard errors associated with the estimated parameters is to use the Fama-MacBeth (1973) method with correction for serial correlation following Newey and West (1987). Petersen (2006) makes the first attempt to thoroughly compare various methods of calculating standard errors using

Monte Carlo simulations. He points out that the standard Fama-MacBeth method has serious limitations in the presence of firm specific time invariant fixed effects in panels of large cross sections with short time series length. Skoulakis (2006) suggests a modification of Fama-MacBeth regressions by first demeaning the variables at the firm level to remove firm specific time invariant effects, and then applying the Fama-MacBeth procedure. With the presence of firm specific time invariant effects in large cross sections with short time series panels, he suggests estimating time-series slope coefficients for each firm, then calculating the t-statistics from the collection of firm level estimates available.[31]

In the set of characteristic regressions relating future returns to past overlapping returns and firm characteristics I examine in this paper, there are time-series serial-correlations, time-series cross-autocorrelations, cross-sectional correlations and conditional heteroskedasticity. Note that the length of the time-series (April 1985 to December 2005, 249 months) is relatively short – near the minimum length required for applying the law of large numbers ("large T" case in Skoulakis (2006)) to ensure the consistent HAC estimates of standard errors. Cochrane (2005) succinctly summarizes one of the main challenges in the empirical asset pricing literature:

> "Our econometric techniques all are designed for large time series and small cross sections. Our data has a large cross section and short time series. A large unsolved problem in finance is the development of appropriate large-N small-T tools for evaluating asset pricing models." (p.226)

[31] With more complicated dependence structures common in panel data, he references the work of Conley (1999) on the spatial method.

In this paper, I take this challenge and deploy the necessary econometric techniques. Building on the theoretical foundations of Conley (1999), I introduce a spatial heteroskedasticity and autocorrelation consistent (SHAC) estimator to compute the standard errors in the time-series cross-sectional panel data. One appealing feature of the SHAC estimator is that its asymptotic properties rely on large $N \times T$, making it is suitable for relatively large cross section and relatively short time series.

To date there are few applications of the SHAC covariance estimator in the finance literature and none in modeling asset returns to my best knowledge. There are three obstacles to overcome when applying this approach. First, the SHAC estimator involves the choice of an "economic distance" metric which must capture the dependence structure among covariances, but there is no such "economic distance" that will work in all situations. My construction of "economic distance" in an unbalanced panel structure complements the empirical application of the SHAC estimator in a context where a natural distance may exist.[32] Second, it is important to examine and validate whether a given economic distance indeed captures the covariance structure.[33] Third, estimation of the spatial HAC estimator is computationally intensive when the panel size is large. In my unbalanced panel dataset, the total pairwise economic distances are over 39 million. Even though the main conclusions from the multivariate regressions are robust to the standard

[32]For example, Conley and Dupor (2003) apply the concept of "upstream" and "downstream" industries to study sector complimentarities.

[33]There are three papers I am aware of applying spatial methods. Pulvino (1998) decides the weighting schemes based on the identifies of buyers and sellers, and transaction time. Silva (2001) uses two-digit SIC code as a measure of closeness, and he assigns weight of one within each industry, and industry return correlations between industry. However, both papers do not evaluate whether such characterization actually captures the dependence structure. Gao (2006) applies "economic distance" to study return comovement of stocks in Standard and Poor's 500 index, and shows such covariance structure characterization is robust both in-sample and out-of-sample. His evidence suggests spatial approach as a viable alternative to factor model in modelling large portfolio's variance-covariance structure.

errors calculated from the Fama-MacBeth procedure with Newey-West adjustment or the time-series cross-sectional regression with the SHAC estimator, these two approaches do yield some noticeable differences for a subset of regressors. In some cases, the use of Newey-West correction for serial correlations in Fama-MacBeth procedure reduces the t-statistic of the standard Fama-MacBeth estimates by a factor of 2, and the use of the SHAC estimator reduces it further by another factor of 2. In Appendix B, I provide background information on the SHAC estimator. I show that the SHAC estimator nests the popular covariance estimators discussed in Petersen (2006) and Skoulakis (2006). In particular, I show how to construct a distance measure using the Euclidian norm of the Z-scores of size, book to market equity and past returns, and provide evidence that this distance measure indeed captures the dependence structure in my dataset remarkably well.

2.6.4. Results from Characteristic Regressions

This section outlines the results from these characteristic regressions. I focus on the concentration of momentum effects and within-news herding effects.

Concentration of Momentum Profits: I consider a holding horizon of six months ($i.e.$ $J = 6$) in these characteristic regressions in (2.8) to (2.11), because six months seem to be the most profitable holding horizons for the price momentum strategies. Looking at other horizons, say one month or three months, generates qualitatively similar results. Tables 2.11 and 2.12 report the results from the Fama-MacBeth and the pooled time-series cross-sectional regressions respectively. In the Fama-MacBeth regressions, the t-statistics are calculated using Newey-West HAC estimators with five lags to take into account the overlapping nature of the returns. In the time-series cross-sectional regressions, the standard errors are computed using the SHAC estimators. In the construction of the economic distance index for SHAC estimator, the number of lags on the time-dimension is chosen based on the return's overlapping lengths, and the cross-sectional distance is based on the Euclidean norm of the Z-scores of size, book to market equity and past 12-month returns.

The point estimates of the regressors are similar in both the magnitudes and statistical significance. For robustness, I rely on both types of t-statistics to draw the conclusions. In my sample, there is considerable momentum effects: the past 12-month returns reliably predict future 6-month returns (Model 1, Panel A) regardless of how we calculate the standard errors. The statistical significance of such predictability slightly strengthens after controlling for other known characteristics shown to be able to predict returns (Model 2, Panel A). Adding future turnover ratios as an additional control does not significantly change the magnitude and statistical significance of the momentum effects (Model 3, Panel A). The signs of these control variables are generally consistent with the evidence from the prior literatures. Due to the differences in sample selection and time period, some of the results of statistical significance are attenuated. However, the issuance effect is rather pronounced (Loughran and Ritter, 1995; Brav, Geczy and Compers, 2001). Panel B in tables 2.11 and 2.12 shows that the price momentum effects only exist among the "herding stocks", $\mathbf{I}(IWF_{i,t} = 1) = 1$. When there is anti-herding, i.e. $\mathbf{I}(IWF_{i,t} = 3) = 1$, there is no price momentum effects. For the stocks belonging to the intermediate information weighting factor tercile, i.e. $\mathbf{I}(IWF_{i,t} = 2) = 1$, the price momentum effect is rather weak and statistically insignificant according to the Fama-MacBeth regressions.

Moreover, consistent with the early evidence of the profitability of various momentum portfolios interacting with the herding tendency measures, the magnitudes of the coefficients estimates decrease monotonically from herd stocks to anti-herd stocks. Specifically, after controlling for other variables shown to predict future returns, the Fama-MacBeth regression coefficients and portfolio characteristics implied 6-month returns of different information weighting factor stratified portfolios range from 2.98% for herding stocks, 0.93% for anti-herding stocks, and 1.83% for the rest of the stocks in the portfolio. Panel C illustrates that adding the set of control variables including future stock turnover ratios does not change the conclusion that the price momentum effects only exist among the "herding stocks". The results from Panels B and C (in Tables 2.11 and 2.12) confirm no incremental predictive power of information weighting factors on future returns, conditional on the different levels of herding tendency. The levels of herding tendency, captured by the indicator variables $\mathbf{I}(IWF_{i,t} = j)$, where $j = 1$, 2, and 3, have no predictive power of future returns and the magnitudes of coefficient estimates are small (not reported in the table).

"Within-News" Effects: The return accumulation horizon for the characteristic regression in (2.12) somewhat matters, especially for herding and anti-herding past winners. Therefore, different holding horizons ranging from one month to six months are reported in tables 2.13 and 2.14. In the Fama-MacBeth regressions, the t-statistics are calculated using Newey-West HAC estimators with zero, two and five lags to take into account the overlapping nature of the returns. In the time-series cross-sectional regressions, the standard errors are computed using SHAC estimators, where the construction of the economic distance metric is similar to models 1 to 4.

Among the past winners, the coefficients from the Fama-MacBeth regressions in table 2.13 and the sample characteristic implies that the high herding stocks outperform the low herding stocks about 23 bpts in the first month. The difference between high herding and low herding stocks are not statistically significant for holding horizon beyond the second month after portfolio formation, even the outperformance of herding stocks increase to 47 bpts during the second to the fourth months, and further to 57 bpts during the second to the seventh month. Among the past losers, the high herding stocks underperform the low herding stocks by about 23 bpts during the second month after portfolio formation ($J = 1$), about 78 bpts during the second to the fourth months after portfolio formation ($J = 3$). Subsequently, when the holding horizon increase to six months, the underperformance of high herding stocks reaches 90 bpts ($J = 6$) but not statistically significant. The magnitudes of "within-news" herding effects implied by the Fama-MacBeth regressions are consistent with the portfolio sorting exercises.

The results from the time-series cross-sectional regressions in Table 2.15 are consistent with those from the Fama-MacBeth regressions. The exception is about the herding effects within the past winners. The time-series cross-sectional regressions suggest the "within-news" herding effects are strongest among the stocks with past return's quintile ranking equal to four rather than the extreme past losers (past return's quintile ranking equal to five). Similar to the regression evidence from Fama-MacBeth regressions, the "within-news" herding effects among winners are relatively short-lived, about during the second month after portfolio formation only. The levels of past returns, captured by the indicator variables $\mathbf{I}(R^*_{i,t-11,t} = j)$, where $j = 1$, 2, 3, 4 and 5, have no predictive power of future returns and the magnitudes of coefficient estimates are small (not reported).

2.7. Additional Robustness Checks

2.7.1. Further Control of Earnings Momentum Effects

After directly controlling for the standardized unexpected earnings (SUE) in the cross-sectional regressions, to further control for possible vintage of earnings momentum effects, I exclude all stocks with SUE in the top and bottom deciles during the portfolio formation quarter, and redo the Fama-MacBeth regressions. This is a very stringent control, because the earnings momentum effects mainly present in these two extreme decile portfolios; and outside these extreme portfolios, the earnings momentum effects are rather weak. For example, Sadka (2006) constructs the earnings momentum portfolios by sorting the universe of stocks into 25 portfolios. Except for the bottom three and top two of these 25 SUE portfolios, the intercept terms from the Fama-French three-factor regressions are not reliably different from zero. In my sample, none of the SUE control variable in the cross-sectional regressions are significant, and its magnitude further attenuates. Panel A in Table 2.16 reports the concentration of momentum effects as specified by regression model (2.12), and Panel B reports the within-news herding effects as specified by regression model (2.11). The point estimates and statistical significance of these variables are similar to the early results where the sample including the extreme SUE stocks.

2.7.2. Subsample Period Evidence

To ensure the results are not driven by early years when (1) there might be imprecise recording of the date of individual analysts forecasts, or (2) imprecise recording of actual earnings announcements, or (3) the coverage of I/B/E/S is relatively small, I redo all the cross-sectional regressions starting from January 1990 to December 2005. The results are reported in Panels C and D in table 2.16. The sub-sample period results are consistent with the full sample period evidence. Additionally, to ensure the results are not driven by the "bubble" period, I redo all the regressions by excluding all years from 1998 to 2001 (inclusive). The sub-sample period results are consistent with the full sample period evidence (not reported). In the sub-sample period analyses, since the time length is shortened, I also use the pooled time-series cross-sectional regression with the spatial HAC estimator to adjust for standard errors. These estimates and their statistical significance of the regression coefficients are largely consistent with the overall sample evidence presented early (not reported).

2.8. Conclusion

Financial analysts as a group generate important value relevant information and disseminate that information to investors. Like any other group of investors, analysts too are subject to psychological biases; and the incentives they face can also induce biases. One such bias, the tendency to herd when making quarterly earnings forecasts, is one of the most important biases of the financial analysts discussed in the literature. I characterize the tendency to herd among analysts as an inefficient aggregation of private information. I view this paper as a first attempt to explore the implication from the herding tendencies among sell-side financial analysts to asset prices, and in particular the momentum effects.

The first goal of this paper is to investigate which stock characteristics are related to the herding tendencies among analysts. The empirical evidence suggests that financial analysts exhibit stronger herding tendencies among smaller and growth-oriented stocks with higher share turnover ratio, and concurrently at times when there is more news media coverage but smaller degrees of divergence of opinions among analysts. These results are consistent with the idea that herding tendencies are related to the decision environment complexity, the role of media, the information horizon, and opinion conformity, as conjectured by previous literatures.

The second goal of this paper is to investigate whether market takes these biases into account when forming asset prices. In particular, I consider the following joint hypothesis: If the market does not adjust biases of analyst forecasts sufficiently, and to the extent that price momentum effects reflect inefficient aggregation of private information of the analysts, then asset prices may exhibit more pronounced drifts. This hypothesis gains considerate empirical support. First, I show that the price momentum effects concentrate among the stocks when analysts covering the stocks exhibit strong tendencies to herd. In fact, among otherwise similar stocks without herding tendencies among analysts, there are no price momentum effects. Second, conditional on the past returns, a stock exhibits more pronounced price momentum at those points in time when the analysts following that stock tend to herd more. Therefore, among the stocks with good (bad) news in the past, for the subset of stocks where analysts covering them exhibit stronger herding tendency, the upward (downward) drift in price is more pronounced than the subset with weaker herding tendency. A comprehensive set of diagnostics illustrate that the relationship between herding tendencies of the analysts and returns is distinct from the earnings momentum effects, information uncertainty effects and liquidity effects, among other effects shown to be related to momentum effects. Collectively, the evidence suggests a strong relationship between the herding tendencies among analysts and asset prices.

These results have strong implications both for the literature on behaviors of analysts, and the literature in asset pricing. Several prominent theoretical models have been proposed to explain the momentum effects, including BSV, DHS and HS. While all these models imply misaggregation of information of some sort, the reason why the misaggregation happens depends on the model. In contrast to these models, I identify a particular channel for information misaggregation that I am able to empirically verify its relationship to the momentum effects. For instance, BSV assume that representative agents suffer from cognitive biases, viz, conservatism and representativeness. As a result, these agents either do not update enough (conservatism) or extrapolate too much (representativeness bias). DHS assume that investors are prone to overconfidence and self-attribution bias. These investors update their beliefs with non-Bayesian weights on public and private information. It is possible that these biases have a common driver that also causes herding, but herding can also result from rational behavior on the part of analysts in response to the incentives.[34] HS hypothesize that private information diffuses slowly into the financial markets. While findings are certainly consistent with the HS hypothesis, in the slow information diffusion world envisioned by HS, my findings pose the following paradox: without financial analysts, the information flow would be slow; however, with biased financial analysts, the information flow may be even slower for some stocks at certain points in time!

[34]Using data from the football wagering market, Durham, Hertzel and Martin (2005) find that market participants behaviors are consistent BSV. Among recent experimental studies, Bloomfield and Hale (2002) find behaviors of the subject are consistent with regime-switching characterization in BSV. However, Asparouhova, Hertzel and Lemmon (2005) find contrary evidence using different experiment designs. Çelen and Kariv (2004) suggest that overconfidence is related to the information cascade behaviors.

The most intriguing aspect of the empirical evidence is why markets appear to fail to adjust for such bias. Information acquisition costs could be one reason. Some investors (especially small retail traders) may face substantial information acquisition costs. If the costs to "de-bias" are high enough, it may be optimal for them not to adjust for such bias. I can only conjecture this possibility, but this seems to be a potential venue to rationalize why investors fail to adjust for such bias.[35] Without detailed data on individual's information acquisition costs, it is difficult to tell whether the acquisition costs are so high that individual investor chooses to systematically live with such bias. The short term immobility of intermediation capital could be another reason. For sophisticated investors such as financial intermediaries, they need to specialize in certain segment of the markets, and understand what is going on before moving capital to take advantage of profitable opportunities.[36] These possibilities remain to be explored in future research.

[35] Nieuwerburgh and Veldkamp (2006) consider a model of information acquisition costs and correlated learning. An implication from their model is that individuals may choose to hold highly concentrated portfolio with fewer assets than what is implied optimal by the classical portfolio choice model.

[36] Recent work, including Berndt, Douglas, Duffie, Ferguson and Schranzk (2005), Da and Gao (2006), Da and Schaumburg (2006), Gabaix, Krishnamurthy and Vigneron (2006) provide empirical evidence along this line in CDS markets, distressed equity markets, constituents of SP500 index, and MBS markets.

Table 1.1: Returns of portfolios sorted on DLI

For each month from 1971/01 to 1999/12, we sort all stocks into 10 deciles according to their DLIs. Panel A reports the equally-weighted returns of these portfolios during each of the first six months after portfolio formation. Panel B reports the average size, book-to-market ratio and DLI at the end of the first month after portfolio formation and the changes in these characteristics from the previous month.

Panel A: first six month return of the DLI-sorted portfolios

Return (1 mth)	Return (2 mth)	Return (3 mth)	Return (4 mth)	Return (5 mth)	Return (6 mth)
0.0113	0.0120	0.0122	0.0120	0.0121	0.0125
0.0107	0.0154	0.0158	0.0164	0.0156	0.0152
0.0138	0.0148	0.0148	0.0139	0.0125	0.0139
0.0133	0.0143	0.0143	0.0145	0.0139	0.0138
0.0138	0.0148	0.0148	0.0139	0.0146	0.0144
0.0140	0.0155	0.0145	0.0142	0.0138	0.0128
0.0123	0.0132	0.0142	0.0137	0.0130	0.0137
0.0126	0.0137	0.0133	0.0131	0.0142	0.0141
0.0118	0.0123	0.0131	0.0131	0.0146	0.0146
0.0210	0.0152	0.0137	0.0149	0.0143	0.0165

Panel B: Characteristics of the DLI-sorted portfolios

MktCap ($million) 1 mth	? MktCap ($million)	B/M 1 mth	? B/M	DLI (%) 1 mth	? DLI (%)
2189.97	25.05	0.62	0.00	0.01	0.01
1328.26	24.47	0.73	0.01	0.02	0.02
941.52	14.68	0.75	0.00	0.03	0.03
653.11	8.47	0.79	0.01	0.07	0.06
459.87	7.08	0.83	0.00	0.16	0.12
343.50	4.28	0.90	0.01	0.41	0.24
229.26	3.40	0.99	0.01	1.05	0.43
143.43	2.17	1.13	0.01	2.85	0.70
81.87	1.14	1.33	0.01	8.60	0.75
40.67	1.07	1.89	-0.03	34.89	-1.55

Table 1.2: Variance decomposition of Default Likelihood Indicator (DLI) based on leverage, past-return and asset volatility

This table reports the percentage of total cross-sectional variation in DLI explained by financial leverage, past one-year return and asset volatility in a variance decomposition framework. We have preformed the decomposition on the full sample (Panel A), the top 1/3 of the sample with the highest DLI (Panel B) and the top 1/5 of the samples with the highest DLI (Panel C). The sampling period is from 1971/01 and 1999/12. Details are provided in the Appendix A.

	Leverage	Past One-year Return	Asset Volatility	Approximation Errors
Panel A: Full Sample				
Average	0.69	0.02	0.56	
Sensitivity of -DD	1.54	-1.79	3.44	
Beta with respect to -DD	0.34	-0.10	0.06	
WLS Standard Errors	0.00	0.00	0.00	
Percentage of Variance Explained	51.82%	17.03%	20.23%	10.92%
Panel B: top 1/3 DLI sample				
Average	1.47	-0.20	0.74	
Sensitivity of -DD	0.37	-1.35	1.09	
Beta with respect to -DD	1.39	-0.24	0.07	
WLS Standard Errors	0.01	0.00	0.00	
Percentage of Variance Explained	51.70%	31.76%	7.13%	9.41%
Panel C: top 1/5 DLI sample				
Average	2.01	-0.34	0.80	
Sensitivity of -DD	0.21	-1.26	0.60	
Beta with respect to -DD	2.34	-0.27	0.06	
WLS Standard Errors	0.02	0.00	0.00	
Percentage of Variance Explained	48.70%	34.26%	3.69%	13.35%

Table 1.3 10 DLI-sorted portfolios, their migration matrix and the associated returns during portfolio formation month and the first-month after portfolio formation

At the end of each month from 1970/12 to 1999/12, we sort all stocks into 10 deciles according to their DLIs (decile 1: Low DLI and decile 10: High DLI). Panel A reports the equally-weighted return during and one month after portfolio formation, and various characteristics of these portfolios. The Amihud illiquidity measures are multiplied by 1000. The average analyst coverage is estimated from 1984/01 to 1999/12. Panel B reports the transition probability of a stock moving from DLI decile i during the month immediately prior to the portfolio formation month (t-1) to DLI decile j during the portfolio formation month (t). Panel C and D report the associated equally-weighted returns during the portfolio formation month (t) and one month after portfolio formation month (t+1), respectively. Panel E reports the corresponding Fama-French three-factor risk-adjusted returns during the first month after portfolio formation (t+1). Risk-adjusted-returns that are statistically significant (at 5% confidence level) are highlighted in bold. The sampling period is from 1970 to 1999.

Port ID	Return one month after formation	Return during formation month	Characteristics (mean)							
			DLI (%)	MktCap (in million)	Book-to-market	Price	Amihud	Idio risk	% covered by analysts	# of analyst
Low DLI	0.0113	0.0248	0.00	2164.92	0.62	52.12	0.47	86.3%	73.5%	5.39
2	0.0107	0.0231	0.00	1303.78	0.73	29.37	0.92	86.5%	76.7%	4.98
3	0.0138	0.0270	0.00	926.84	0.75	24.48	0.87	88.2%	67.4%	4.55
4	0.0133	0.0268	0.01	644.64	0.78	20.06	1.29	89.0%	62.6%	4.20
5	0.0138	0.0240	0.04	452.80	0.83	17.02	1.56	89.9%	57.0%	3.80
6	0.0140	0.0208	0.17	339.21	0.89	14.52	2.51	90.8%	51.7%	3.42
7	0.0123	0.0167	0.61	225.86	0.99	11.51	3.52	91.9%	44.9%	3.11
8	0.0126	0.0086	2.15	141.27	1.12	8.77	6.24	93.3%	36.6%	2.87
9	0.0118	-0.0022	7.85	80.72	1.32	6.12	11.54	94.8%	29.1%	2.60
High DLI	0.0210	-0.0339	36.45	39.60	1.92	3.58	31.75	96.6%	20.3%	2.50

Panel B: Transition probability from month t-1 to t (in %)

Decile # at t-1	Decile # at t									
	1	2	3	4	5	6	7	8	9	10
1	81.68	7.20	6.63	2.14	0.95	0.63	0.42	0.23	0.10	0.02
2	20.36	50.55	19.93	5.18	1.86	1.05	0.55	0.33	0.16	0.02
3	13.71	11.06	42.29	21.93	6.78	2.36	1.08	0.52	0.22	0.05
4	3.68	2.60	21.85	39.69	21.74	6.77	2.31	0.96	0.34	0.07
5	1.41	0.83	6.36	23.29	37.96	21.11	6.41	1.95	0.56	0.12
6	0.68	0.33	1.96	7.17	22.88	38.70	21.08	5.67	1.33	0.20
7	0.34	0.17	0.63	1.91	6.63	22.85	41.26	21.43	4.31	0.46
8	0.12	0.09	0.26	0.54	1.58	5.68	22.61	46.95	20.52	1.65
9	0.07	0.04	0.07	0.15	0.36	1.02	4.06	20.60	57.81	15.83
10	0.02	0.01	0.02	0.03	0.05	0.14	0.37	1.57	15.09	82.70

Panel C: Average monthly return during month t (in %)

1	1.75	-1.83	-2.38	-2.69	-2.74	-2.27	-2.91	-4.93	-4.23	-6.57
2	5.06	1.63	-1.85	-4.32	-5.26	-3.63	-5.15	-5.94	-8.90	-6.24
3	6.37	5.21	1.83	-1.89	-4.22	-4.58	-4.35	-6.33	-9.21	-11.88
4	6.55	7.99	5.98	1.69	-2.69	-5.29	-5.86	-6.38	-12.09	-18.51
5	5.30	8.24	8.88	6.51	1.37	-3.57	-6.39	-8.85	-12.11	-21.01
6	5.87	6.72	9.15	10.48	7.18	1.03	-4.57	-8.53	-12.53	-15.67
7	7.04	5.51	9.21	12.02	12.60	7.81	0.58	-6.26	-12.87	-21.01
8	6.84	7.08	6.03	12.29	15.44	16.01	9.06	0.16	-8.67	-19.50
9	3.84	5.00	6.19	10.87	18.02	20.23	20.42	11.40	-0.38	-12.70
10	2.10	7.73	4.31	7.00	8.72	14.46	33.00	35.95	16.86	-1.18

Panel D: Average monthly return during month t+1 (in %)										
1	1.09	1.21	1.51	1.71	1.39	1.82	1.47	3.37	1.33	0.69
2	1.16	1.07	1.58	1.13	1.48	3.79	0.36	2.98	4.21	3.99
3	1.34	1.00	1.47	1.45	1.34	2.02	1.51	3.09	0.22	5.68
4	0.96	1.12	1.23	1.26	1.52	1.67	1.81	3.76	2.05	4.37
5	1.77	-0.75	1.24	1.22	1.35	1.66	2.15	1.94	3.74	1.64
6	1.34	1.00	0.96	1.02	1.33	1.37	1.69	2.14	2.62	2.19
7	1.26	0.85	-0.07	1.17	1.34	1.22	1.20	1.53	2.23	2.14
8	0.95	0.53	3.08	2.04	1.02	0.48	0.77	1.32	1.98	2.00
9	-0.36	2.85	0.50	2.92	-0.15	0.88	-0.15	0.40	1.13	3.03
10	-5.81	-4.60	-1.29	4.21	1.25	2.78	0.01	0.37	-0.31	1.93
Panel E: Average three-factor risk-adjusted monthly return during month t+1 (in %)										
1	0.06	0.25	0.42	0.93	0.49	0.75	0.62	2.17	0.26	2.27
2	-0.08	-0.17	0.27	-0.29	-0.12	1.76	-0.34	2.00	-0.54	10.01
3	0.43	-0.05	0.35	0.34	0.10	1.03	0.37	2.32	-0.39	1.58
4	0.12	0.19	0.12	0.13	0.30	0.46	0.55	1.48	0.93	3.08
5	0.78	-1.61	0.14	0.04	0.19	0.36	0.81	0.48	2.46	-0.60
6	0.36	-0.05	-0.20	-0.14	0.13	0.17	0.43	0.75	0.98	0.83
7	-0.17	0.26	-0.70	-0.15	0.25	-0.02	-0.12	0.23	0.69	1.60
8	0.69	0.43	1.33	1.11	-0.55	-0.74	-0.50	-0.04	**0.63**	0.42
9	-0.74	2.56	-0.17	2.05	-0.92	-0.46	-1.35	**-0.90**	-0.30	**1.37**
10	-6.61	-0.04	-0.48	-0.47	0.97	1.17	-0.71	-1.19	**-1.67**	0.35

Table 1.4 Examples of investment restrictions on institutions

Institutions	Restrictions	Keywords
San Francisco State University Foundation	"Equity investment should have adequate liquidity and a market capitalization of at least $500 million."	Liquidity, Market Capitalization
The Mayer Fund	"no company in which we invest shall have a market capitalization less than $100 million; and at least three Wall Street analysts must cover the stock."	Market Capitalization, Wall street coverage
Kingsville foundation	"investments shall be primarily in well-seasoned, quality companies whose securities enjoy marketability adequate for the portfolio, Industry and company investments shall be based upon demonstrable analysis of prospects for above average return over a three-year period."	Quality of the company, above-average return
Florida College Investment Plan	"a coefficient of determination to the benchmark Index of not less than .80 over any rolling five-year time horizon calculated using monthly data."	Tracking error
University of Wisconsin System trust fund	"portfolio positions should be issues that are publicly traded in sufficient volume to facilitate, under most market conditions, prompt sale without severe market price effect."	Easiness of trading

Table 1.5 Aggregate mutual fund holdings and mutual fund holding changes of all and recent High-DLI stocks

This table illustrates the quarterly aggregate mutual fund holding and holding changes of all high DLI stocks and recent high DLI stocks. Panel A examines the aggregate mutual fund holdings and holding changes of all high DLI stocks when they become financially distressed during any month of the quarter. Panel B examines the aggregate mutual fund holdings and holding changes of new high DLI stocks when they become financially distressed during any month of the quarter. Panel C reports the quarterly mutual fund holding changes during the four quarters after the event quarter (Q) for all high DLI stocks. Panel C reports the quarterly mutual fund holding changes during the four quarters after the event quarter (Q) for all high DLI stocks and new high DLI stocks, respectively. For Panel A and B, all high DLI or new high DLI stocks are further sorted into three groups based on the number of underlying mutual fund shareholders. "Low" refers to ones for which the underlying shareholders is less than or equal to 2, "Medium" refers to ones for which the underlying shareholders between 3 and 7 (inclusive), and "High" refers to ones for which the underlying shareholders greater than or equal to 8. The ranking approximately matches the 33rd percentile and 67th percentile of underlying mutual fund shareholders across all stocks and all years. "All" refers to the full sample irrespective of the number of the underlying mutual fund shareholders. N is the number of stocks across all quarters. The sampling period is from 1980-1999. All holdings and holding changes are reported in percentage.

Panel A: Aggregate mutual fund holdings and mutual fund holding changes of *all* High-DLI stocks

	Statistics	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)
		1980 - 1999			1980-1989			1990 - 1999		
Low	Mean	1.599	1.458	-0.141	1.883	1.74	-0.143	1.326	1.188	-0.139
	t-statistics	61.40	60.25	-5.53	50.86	49.77	-4.07	36.92	36.26	-3.77
	N	5711	5711	5711	2798	2798	2798	2913	2913	2913
Medium	Mean	3.707	2.641	-1.066	4.182	3.136	-1.047	3.452	2.376	-1.076
	t-statistics	75.34	69.44	-22.95	55.48	49.91	-15.00	54.47	50.53	-17.71
	N	4496	4496	4496	1570	1570	1570	2926	2926	2926
High	Mean	5.921	4.055	-1.866	5.052	4.236	-0.816	6.192	3.998	-2.194
	t-statistics	44.71	52.19	-14.92	25.58	29.95	-4.83	38.38	43.53	-14.28
	N	1033	1033	1033	246	246	246	787	787	787
All	Mean	2.839	2.17	-0.669	2.834	2.348	-0.486	2.843	2.046	-0.797
	t-statistics	95.48	97.69	-25.78	71.26	70.84	-14.42	67.43	68.91	-21.43
	N	11240	11240	11240	4614	4614	4614	6626	6626	6626

Panel B: Aggregate mutual fund holdings and mutual fund holding changes of *new* High-DLI Stocks

	Statistics	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)	Aggregate Mutual Fund Holdings at (Q-1) (%)	Aggregate Mutual Fund Holdings at (Q) (%)	Quarterly Holding Changes (%)
		1980 - 1999			1980 - 1989			1990 - 1999		
Low	Mean	1.676	1.595	-0.081	1.963	1.857	-0.107	1.351	1.299	-0.053
	t-statistics	31.37	29.64	-1.46	28.14	24.76	-1.46	16.86	17.21	-0.62
	N	1308	1308	1308	694	694	694	614	614	614
Medium	Mean	4.391	3.228	-1.163	4.659	3.719	-0.941	4.228	2.93	-1.298
	t-statistics	45.33	43.62	-12.96	33.74	31.20	-7.13	32.30	31.51	-10.84
	N	1416	1416	1416	534	534	534	882	882	882
High	Mean	7.463	5.099	-2.364	5.789	4.858	-0.931	7.994	5.175	-2.818
	t-statistics	35.75	46.79	-12.42	20.54	23.28	-4.24	31.35	40.67	-11.90
	N	586	586	586	141	141	141	445	445	445
All	Mean	3.862	2.914	-0.948	3.409	2.892	-0.517	4.181	2.929	-1.252
	t-statistics	55.59	60.74	-16.52	41.86	41.13	-7.58	40.59	45.01	-14.81
	N	3310	3310	3310	1369	1369	1369	1941	1941	1941

Panel C: Aggregate mutual fund quarterly holding change after event quarter for *all* High-DLI stocks and *new* High-DLI stocks

		Quarterly Holding Changes (Q+1 - Q) (%)	Quarterly Holding Changes (Q+2 – Q+1) (%)	Quarterly Holding Changes (Q+3 – Q+2) (%)	Quarterly Holding Changes (Q+4 – Q+3) (%)
all High-DLI stocks	Mean	-0.032	-0.102	0.002	0.019
	t-statistics	-1.54	-4.78	0.08	0.83
	N	9193	7901	7410	6982
new High-DLI stocks	Mean	-0.129	-0.115	-0.045	0.056
	t-statistics	-3.08	-2.81	-1.08	1.29
	N	2463	2203	2078	1970

Table 1.6 Institutional trading at monthly frequency

Based on the institutional trading dataset provided by Plexus Group, for each stock in our sample during portfolio formation month, we first compute the aggregate net buy/sell orders (as percentage of total number of shares outstanding) submitted by institutions and actual aggregate shares bought/sold (again as percentage of total number of shares outstanding) by institutions at a monthly frequency. We then average these two institutional trading measures first across all stocks at portfolio level and then across time. A negative number indicates net selling.

Portfolio	Aggregate net buy/sell order (as % of total # of shares outstanding) by institutions (in %)	*t*-value	Aggregate shares bought/sold (as % of total # of shares outstanding) by institutions (in %)	*t*-value
Low DLI	0.02	2.91	0.01	1.91
2	0.01	0.28	-0.01	-0.47
3	0.03	2.01	0.02	2.42
4	0.06	3.31	0.03	3.03
5	0.06	3.98	0.04	3.39
6	0.04	1.11	0.05	1.80
7	0.03	0.55	0.04	1.24
8	-0.09	-1.22	-0.02	-1.41
9	-0.10	-2.45	-0.08	-2.41
High DLI	-0.18	-3.50	-0.14	-3.12
New High DLI	-0.14	-2.26	-0.11	-2.32

Panel B: Second sub-sample (Q1 of 1996 to Q1 of 1998)

Portfolio	Aggregate net buy/sell order (as % of total # of shares outstanding) by institutions (in %)	t-value	Aggregate shares bought/sold (as % of total # of shares outstanding) by institutions (in %)	t-value
Low DLI	0.03	2.34	0.02	1.95
2	-0.01	-0.14	-0.03	-0.40
3	0.04	1.71	0.04	2.20
4	0.07	3.87	0.06	3.21
5	0.09	4.79	0.08	4.27
6	0.12	1.70	0.09	1.68
7	0.13	1.67	0.10	1.66
8	-0.02	-0.51	-0.02	-0.68
9	-0.10	-1.72	-0.10	-1.76
High DLI	-0.15	-2.14	-0.15	-2.05
New High DLI	-0.13	-2.65	-0.12	-2.69

Table 1.7 Changes in liquidity-related characteristics during the liquidity shock

This table reports various stock characteristics during liquidity shock for 10 DLI-sorted deciles and also New High DLI stocks. New High DLI stocks are stocks which just enter into the highest-DLI decile during the current portfolio formation month.

Panel A reports the trading volume during three two-month periods: (1) the two months prior to the portfolio formation month ([-2.-1]); (2) the portfolio formation month and the first month after portfolio formation([0,1]); (3) the second and third month after portfolio formation ([2,3]). The trading volumes are adjusted for changes in the total number of shares outstanding. Finally, the trading volumes are normalized by the trading volume during the two months prior to the portfolio formation month ([-2.-1]). The sampling period is from 1970 to 1999.

Panel B reports the percentage bid-ask spread one month before portfolio formation, at portfolio formation and one month after portfolio formation. The percentage bid-ask spread is defined as (ask – bid) / mid. It is computed using intraday quote data from TAQ (after 1993) and ISSM (before 1993). The sampling period for NYSE stocks is from 1983 to 1999 and the sampling period for NASDAQ stocks is from 1987 to 1999.

Panel C reports two order imbalance measures during and one month after the portfolio formation month. Both measures are developed in Chordia, Roll and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004). OIBSH1 measures the buyer-initiated shares purchased less than the seller-initiated shares sold and OIBSH2 is OIBSH1 scaled by the total number of shares traded. The sampling period is from 1988 to 1998.

Panel D reports the average realized spreads (scaled by traded price) for each DLI decile and the portfolio of "new" high DLI stocks around portfolio formation month. The realized spread is computed using intraday quote data from1983 and 1999. The detailed estimation procedure is described in Huang and Stoll (1996). The time horizon used for the estimation is 30 minutes.

Panel E reports the liquidity betas during the portfolio formation month (0) and the first month after the portfolio formation month (1). It also reports the average liquidity betas during two three-month periods: the pre-formation - the three months prior to the portfolio formation month ([-3.-1]) and the post-formation – the second to fourth month after portfolio formation ([2,4]). As in Pastor and Stambaugh (2003), the liquidity beta during month n is defined as the slope coefficient (β_n) in the following regression:

$$r_{i,t}^n = \alpha_i^n + \beta_{i,L}^n L_t + \beta_{i,M}^n MKTRF_t + \beta_{i,S}^n SMB_t + \beta_{i,H}^n HML_t + \epsilon_{i,t}$$

where r^0 denotes the excess return during portfolio formation month and $MKTRF$, SMB and HML are the Fama-French three factors. L is the innovation in the aggregate liquidity measure defined in Pastor and Stambaugh (2003). The changes in the liquidity betas over these periods are also reported, and the associated t-values are computed using Newey-West standard error estimators with three lags. The sampling period is from 1970 to 1999.

Panel A: Normalized trading volume

Portfolio	(1) Normalized Volume during Month [0,1]	*t*-value associated with (1)-1	(2) Normalized Volume during Month [2,3]	*t*-value associated with (2)-1
Low DLI	1.035	4.33	1.061	6.04
2	1.036	3.13	1.064	4.25
3	1.036	3.89	1.055	4.67
4	1.026	2.91	1.042	3.75
5	1.024	2.48	1.040	3.24
6	1.017	1.57	1.037	2.42
7	1.013	1.08	1.016	1.14
8	1.000	0.04	1.025	1.56
9	0.993	-0.57	1.025	1.31
High DLI	1.037	1.60	1.052	2.71
New High DLI	1.045	2.18	1.040	1.75

Panel B: Percentage trading cost

Portfolio	(1) Percentage spread 1 month prior to the formation (%)	(2) Percentage spread at formation (%)	(3) Percentage spread 1 month after formation (%)	*t*-value associated with (2)-(1)
Low DLI	1.36	1.35	1.36	-0.94
2	1.71	1.71	1.71	0.05
3	1.92	1.90	1.92	-1.51
4	2.30	2.27	2.30	-1.24
5	2.71	2.70	2.71	-0.55
6	3.18	3.17	3.18	-0.60
7	3.84	3.82	3.86	-0.53
8	4.95	4.94	4.98	-0.21
9	6.71	6.76	6.82	0.94
High DLI	10.68	11.01	11.04	3.63
New High DLI	7.22	8.31	8.23	10.53

Panel C: Order imbalance measures

Portfolio	(1) OIBSH1 during formation month	(2) OIBSH1 one month after formation	(2) - (1) Change in OIBSH1	*t*-value associated with (2) - (1)	(3) OIBSH2 during formation month (in %)	(4) OIBSH2 one month after formation (in %)	(4) - (3) Change in OIBSH2	*t*-value associated with (4) - (3)
Low DLI	10756.2	9885.6	-870.6	-3.87	0.91	0.51	-0.40	-5.42
2	11319.7	11466.3	146.6	0.17	-0.28	0.07	0.35	1.66
3	10128.5	10293.3	164.8	0.39	0.05	0.00	-0.05	-0.41
4	10314.2	10161.5	-152.7	-0.30	0.00	-0.12	-0.12	-0.92
5	9586.6	9539.7	-46.9	-0.09	-0.73	-0.77	-0.04	-0.25
6	8058.1	8747.4	689.3	1.20	-1.63	-1.32	0.31	1.87
7	6313.8	7416.9	1103.1	1.75	-2.56	-2.42	0.14	0.76
8	4312.7	5949.2	1636.5	2.52	-3.60	-3.40	0.20	0.88
9	1562.3	3698.0	2135.7	3.23	-5.76	-4.98	0.78	2.90
High DLI	-710.3	818.1	1528.4	2.03	-6.79	-6.05	0.74	2.35
New High DLI	-2456.6	2958.4	5415.0	2.87	-7.81	-4.26	3.55	5.67

Panel D: Huang and Stoll's realized spread (scaled by traded price)

Portfolio	(1) Spread (in bps) one month before [month = -1]	(2) Spread (in bps) during formation [month = 0]	(3) Spread (in bps) one month after [month = 1]	(4) Spread (in bps) two month after [month = 2]	(2)-(1)	*t*-value for (2)-(1)	(4)-(3)	*t*-value for (4)-(3)
Low DLI	-0.5	-0.6	-0.6	-0.6	-0.1	-0.29	0.1	0.13
2	2.0	0.9	0.9	2.4	-1.1	-0.88	1.4	1.15
3	10.3	9.1	9.4	9.4	-1.2	-1.50	0.0	0.02
4	12.8	12.1	13.0	11.9	-0.7	-0.81	-1.0	-1.23
5	15.5	16.1	15.0	15.2	0.6	0.66	0.3	0.29
6	16.4	15.7	15.6	15.3	-0.7	-0.66	-0.3	-0.25
7	19.1	18.3	18.3	17.5	-0.8	-0.63	-0.8	-0.67
8	24.0	22.6	23.0	22.4	-1.4	-0.94	-0.6	-0.43
9	31.4	31.0	29.8	29.9	-0.5	-0.26	0.2	0.08
High DLI	58.1	59.8	58.9	57.5	1.7	0.75	-1.4	-0.61
New High DLI	31.7	49.9	53.8	43.8	18.3	3.76	-10.0	-2.01

Panel E: Pastor and Stambaugh's liquidity betas

Portfolio	(1) average liquidity beta during month=[-3,-1]	(2) liquidity beta during formation month = 0	(3) liquidity beta one month after month = 1	(4) average liquidity beta during month = [2,4]	(2)-(1)	*t*-value for (2)-(1)	(3)-(2)	*t*-value for (3)-(2)	(4)-(3)	*t*-value for (4)-(3)
Low DLI	0.002	0.006	0.026	0.016	0.004	0.39	0.021	3.07	-0.010	-2.19
2	-0.025	-0.022	-0.002	-0.006	0.002	0.17	0.020	1.39	-0.004	-0.34
3	-0.012	-0.014	-0.010	-0.008	-0.002	-0.14	0.004	0.42	0.001	0.14
4	-0.001	0.003	0.008	0.004	0.004	0.37	0.006	0.51	-0.004	-0.48
5	0.000	0.005	0.012	-0.008	0.005	0.47	0.007	0.53	-0.020	-2.18
6	0.000	-0.024	-0.009	-0.008	-0.024	-2.30	0.015	1.20	0.002	0.16
7	0.025	0.009	-0.002	-0.004	-0.016	-1.30	-0.011	-0.74	-0.003	-0.27
8	0.011	-0.001	-0.022	-0.016	-0.012	-0.84	-0.021	-1.45	0.006	0.58
9	0.002	0.004	-0.015	-0.006	0.002	0.11	-0.019	-1.15	0.009	0.72
High DLI	-0.080	-0.052	-0.110	-0.077	0.028	2.07	-0.058	-2.01	0.033	2.69
New High DLI	-0.078	0.036	-0.103	-0.069	0.114	2.39	-0.139	-2.11	0.034	0.61

Table 1.8 Cross-sectional regressions with stock characteristics

Each month from 1970/01 to 1999/12, we run a cross-sectional regression of the next month three-factor alphas on various current month stock characteristics. The alphas are estimated using rolling-window regressions. All variables are cross-sectionally demeaned so the intercept term is zero. In addition, the stock characteristics are also standardized so the regression slope coefficient can be interpreted as the impact on the return of a one standard deviation change in the variable. The slope coefficients are then averaged cross time and reported. The robust t value is computed using Newey-West autocorrelation adjusted standard error with 12 lags. Amihud is a liquidity measure; DLI is the Default Likelihood Indicator of Vassalou and Xing (2004); Size is the log of market capitalization; B/M is the book-to-market ratio and Pastret is the return one month prior to the portfolio formation. We exclude stocks with missing characteristics and negative B/M. The regressions are estimated for both the full sample (1589 stocks per month on average) and the top DLI Quintile (272 stocks per month on average). Panel A reports the correlations among the characteristics (Full sample in lower-triangular and the Top DLI-quintile in the upper triangular). Panel B reports the regression results. The robust t value is reported below the coefficient estimate in italic. The regression slopes are presented in the unit of percentage return. For Panel B, the factor loadings are computed using monthly data in a five-year rolling window.

Panel A: Correlations

		Top DLI Quintile				
		Amihud	DLI	Size	B/M	Pret
Full Sample	Amihud		0.140	-0.190	0.103	0.016
	DLI	0.183		-0.232	0.353	-0.123
	Size	-0.143	-0.318		-0.208	0.034
	B/M	0.128	0.426	-0.362		-0.089
	Pastret	-0.003	-0.117	0.043	-0.106	

Panel B: Regression results (pre-formation factor loadings estimated using 5-year monthly data)

	Full sample							Top DLI-quntile						
	Pastret	Amihud	DLI	Size	B/M	Pastret* Amihud	R-square	Pastret	Amihud	DLI	Size	B/M	Pastret* Amihud	R-square
Model 1			0.217				0.80%			0.660				1.00%
			3.16							*5.73*				
Model 2		0.233					0.74%		0.592					1.21%
		3.56							*4.67*					
Model 3	-0.995						1.17%	-2.263						2.38%
	-9.70							*-13.35*						
Model 4	-1.004	0.200	0.037				2.47%	-2.244	0.541	0.265				4.44%
	-9.97	*3.22*	*0.61*					*-13.66*	*4.25*	*2.42*				
Model 5	-1.023	0.178	-0.084	-0.019	0.284		3.24%	-2.215	0.345	-0.066	-0.649	0.611		6.10%
	-10.19	*3.26*	*-1.60*	*-0.31*	*4.97*			*-13.44*	*2.97*	*-0.70*	*-4.36*	*5.02*		
Model 6	-0.877	0.043	-0.101	-0.033	0.288	-0.597	3.68%	-1.954	0.276	-0.086	-0.625	0.629	-0.639	7.02%
	-9.36	*0.54*	*-1.90*	*-0.56*	*5.21*	*-7.80*		*-11.93*	*1.40*	*-0.89*	*-4.34*	*5.31*	*-4.16*	

Table 1.9: Evidence on Alternative Explanations

During our sample from 1970 to 1999, at the end of each month, we sort all stocks according to Default Likelihood Indicator (DLI) into deciles 1 to 10 (1: Low-default risk; 10: High-Default risk). We then consider two stock portfolios: (1) New High-DLI stocks: a portfolio of high-default-risk stocks which have become financially distressed only recently (They are in DLI-decile 10 during the current month but not in the previous month); (2) Characteristics-matched Low-DLI stocks: a portfolio of stocks with similar past returns and trading prices but relatively low default risk, constructed as follows: at the end of each month, we focus on Low-DLI stocks (stocks in DLI-decile 1 to 9) and further sort them on their past one-month returns and trading prices into 36 portfolios. Among the 36 portfolios, we choose the portfolio whose past one-month return and trading price are closest to those of the New High-DLI stocks on average. Panel A reports various characteristics of the two portfolios. Panel B reports mutual fund holding changes and institutional transaction information for the two portfolios. Finally, we sort all stocks according to institutional selling pressure using the actual transaction data from Plexus group into deciles and report the associated changes in DLI in Panel C.

Panel A: Portfolio Characteristics

Portfolio	# of stocks per month	Price ($) at formation (t = 0)	DLI one month before (t = -1)	DLI at formation (t = 0)	Return (%) during formation month (t = 0)	Return (%) one month after (t = 1)
New High-DLI	49.2	4.85	0.099	0.233	-12.75	2.93
Char-matched Low-DLI	63.3	4.14	0.020	0.033	-15.50	2.15

Panel B: Mutual Holding Changes and Institutional Transactions

	Plexus Group Data (Full sample)				Mutual Fund	
	Agg net buy/sell order (as % of shares outstanding) by Institutions (%)	*t*-value	Agg shares bought/sold (as % of shares outstanding) by Institutions (%)	*t*-value	Holdings Change (Q - Q-1) (%)	*t*-value
New High-DLI	-0.18	-2.26	-0.11	-2.32	-0.948	-16.52
Char-matched Low-DLI	-0.03	-0.34	-0.03	-0.37	-0.045	-9.23

Panel C: Institutional selling pressure and change in DLI

	High	2	3	4	5	6	7	8	9	Low
Agg shares bought/sold (as % of shares outstanding) by Institutions (%)	1.33	0.35	0.15	0.07	0.02	0.00	-0.03	-0.11	-0.30	-1.33
Change in DLI (%)	-0.07	-0.01	0.04	-0.03	0.05	-0.03	0.01	-0.06	0.07	0.07

Table 1.10: Effect of bid-ask bounce

At the end of each month from 1970/12 to 1999/12, we sort all stocks into 10 deciles according to their DLIs (decile 1: Low DLI and decile 10: High DLI). We report the equally-weighted return during the first month after portfolio formation. We also report the measure for return bias (in bp) due to bid-ask bounce computed as $\left(\frac{P_A-P_B}{P_A+P_B}\right)^2$ where P_A and P_B are the bid and ask price of the stock. We first compute the return bias for the full sample (1970-1999) by assuming a constant bid-ask spread of \$0.25. We also compute the return bias using the actual quoted spread (quoted ask – quoted bid) from quote data in TAQ (after 1993) and ISSM (before 1993). The sampling period for NYSE stocks is from 1983 to 1999 and the sampling period for NASDAQ stocks is from 1987 to 1999. Finally, we compute the monthly return using daily returns from the second positive trading-volume-day.

DLI Decile #	Return one month after formation	Return bias due to bid-ask bounce Assuming a spread of \$0.25, 1970-1999 (in bp)	Return bias due to bid-ask bounce Using actual quoted spread, 1983-1999 (in bp)	First-month return excluding the return on the first trading day
1	0.0113	1.75	0.56	0.0112
2	0.0107	2.35	0.91	0.0106
3	0.0138	3.14	1.29	0.0138
4	0.0133	4.33	1.84	0.0132
5	0.0138	5.83	2.58	0.0137
6	0.0140	7.92	3.71	0.0139
7	0.0123	11.38	5.58	0.0123
8	0.0126	17.61	9.74	0.0124
9	0.0118	28.28	17.01	0.0118
10	0.0210	53.81	42.09	0.0202

Table 1.11: Economic significance of the first-month high return on the High-DLI and New High-DLI stocks

We focus on the High-DLI stocks (stocks in the highest-DLI decile during the formation month) and New High-DLI stocks (stocks that enter the highest-DLI decile only during the formation month) and further sort them into quartiles according to their market capitalizations (in Panel A) or their trading prices (in Panel B). We then report various characteristics for each quartile. The percentage bid-ask spread and the return bias due to bid-ask bounce are both computed using the actual quoted spread (quoted ask – quoted bid) from quote data in TAQ (after 1993) and ISSM (before 1993). The sampling periods for these two characteristics are from 1983 to 1999 for NYSE stocks and from 1987 to 1999 for NASDAQ stocks. For other characteristics, the sampling periods are from 1971 to 1999.

Panel A: Size-sorted quartile

Quartile	# of stocks	Mktcap (million $)	Trading price	Return during formation month	Return one month after formation	Bid-ask spread (%)	Return bias due to bid-ask bounce (bp)
				High-DLI Stocks			
1	65	137.9	7.37	-0.0236	0.0037	5.18	11.08
2	65	13.6	3.51	-0.0257	0.0079	9.86	31.98
3	65	5.5	2.18	-0.0279	0.0152	14.27	63.65
4	65	2.0	1.27	-0.0585	0.0576	23.10	154.48
				New High DLI Stocks			
1	11	235.8	9.69	-0.0998	0.0199	3.69	5.50
2	12	21.8	4.93	-0.1170	0.0152	7.02	18.07
3	12	8.4	3.19	-0.1424	0.0234	10.70	39.93
4	11	3.0	1.79	-0.1801	0.0619	18.89	128.84

Panel B: Price-sorted quartile

Quartile	# of stocks	Mktcap (million $)	Trading price	Return during formation month	Return one month after formation	Bid-ask spread (%)	Return bias due to bid-ask bounce (bp)
				High-DLI Stocks			
1	65	123.3	8.56	-0.0177	0.0058	4.82	10.18
2	65	21.4	3.26	-0.0240	0.0074	9.21	29.78
3	66	9.6	1.77	-0.0324	0.0161	14.58	70.41
4	65	4.9	0.77	-0.0618	0.0553	23.80	147.50
				New High DLI Stocks			
1	11	209.3	11.14	-0.0852	0.0154	3.61	6.32
2	12	39.5	4.73	-0.1169	0.0176	6.74	17.49
3	12	14.3	2.63	-0.1449	0.0289	10.79	43.43
4	11	6.2	1.19	-0.1921	0.0582	19.71	127.92

Table 2.1: Descriptive statistics of analyst information weighting factor

This table reports the distributional characteristics of raw analyst information weighting factor, estimated stock by stock and quarter by quarter, using individual analyst's quarterly earning forecasts, from April 1984 to December 2005. The cutoff value for annual aggregation is based on the year of quarterly earning announcements, as reported in unadjusted I/B/E/S historical actual earning files. The stocks which are followed by less than two analysts, or have less than seven earning quarterly earning forecasts are dropped from the sample. The (raw) analyst information weighting factors are estimated based on the procedure documented in the text. N is the number of stock-quarters. ***, ** and * denote that the t-statistics (for mean) or the sign-rank test statistics (for median) indicate the mean and median of the raw information weighting factor values are different from 0.5 at 1%, 5% and 10% significance levels respectively.

Year	N	Mean		Std	Q1	Median		Q3
1984	493	0.54	***	0.25	0.33	0.50	***	0.71
1985	871	0.56	***	0.23	0.38	0.56	***	0.71
1986	1198	0.59	***	0.23	0.43	0.57	***	0.75
1987	1389	0.59	***	0.22	0.43	0.58	***	0.75
1988	1718	0.57	***	0.22	0.41	0.57	***	0.73
1989	2280	0.56	***	0.22	0.39	0.56	***	0.71
1990	2560	0.55	***	0.23	0.38	0.56	***	0.71
1991	2805	0.55	***	0.22	0.38	0.54	***	0.71
1992	3134	0.55	***	0.22	0.40	0.55	***	0.71
1993	2807	0.54	***	0.21	0.39	0.55	***	0.70
1994	3763	0.55	***	0.21	0.40	0.55	***	0.70
1995	3842	0.53	***	0.21	0.38	0.50	***	0.67
1996	3688	0.52	***	0.22	0.36	0.50	***	0.68
1997	3712	0.52	***	0.22	0.33	0.50	***	0.67
1998	4117	0.51	***	0.23	0.33	0.50	***	0.67
1999	4201	0.51	*	0.23	0.33	0.50	**	0.67
2000	3556	0.49		0.24	0.31	0.50	*	0.67
2001	4239	0.50		0.24	0.33	0.50		0.68
2002	3984	0.49	***	0.24	0.30	0.50	***	0.67
2003	3815	0.50		0.24	0.33	0.50		0.67
2004	4207	0.49	*	0.24	0.32	0.50	*	0.67
2005	4513	0.51	***	0.24	0.33	0.50	***	0.69
Pre-FD	46134	0.53	***	0.22	0.37	0.53	***	0.70
Post-FD	20758	0.50	*	0.24	0.33	0.50	*	0.67
All	66892	0.52	***	0.23	0.35	0.50	***	0.69

Table 2.2: Time-series regression of quarterly analyst information weighting factor

Panel A reports the cross-sectional average of full-sample (1985-2005) time-series regression coefficients of analyst information weighing factor (first column). The first stage is time-series regression of individual stock's (denoted as i) current quarter's (denoted as q) analyst information weighting factor, regressed on past for lags of information weighting factors. Then the time-series regressions' coefficients distributional characteristics (median, upper and lower quartiles) along with the t-statistics associated with mean are subsequently computed from the cross-section of stocks. Panel B and C report the estimates from two sub-samples periods, i.e., 1985 – 2001 (the pre-Regulation FD period) and 2001 – 2005 (the post-Regulation FD period). The cut-off dates are based on the I/B/E/S quarterly earning announcement dates. In the actual portfolio formation, only the forecasts and announcement information prior to formation month is used. *** denotes that the cross-sectional average coefficients are statistically different from zero and from 0.5 for the intercept term.

	$\bar{\alpha}_0$	$\bar{\beta}_1$	$\bar{\beta}_2$	$\bar{\beta}_3$	$\bar{\beta}_4$	R^2
Panel A: All Sample						
Mean of Regression Coefficients / *R*-squared	0.674***	-0.047***	-0.109***	-0.050***	-0.100***	25.94%
Standard Deviations of Regression Coefficients	0.555	0.404	0.512	0.435	0.452	
Q1 of Regression Coefficients	0.429	-0.201	-0.260	-0.198	-0.257	
Median of Regression Coefficients / *R*-squared	0.594	-0.006	-0.074	-0.035	-0.077	17.65%
Q3 of Regression Coefficients	0.852	0.152	0.077	0.121	0.073	
Panel B: Pre Regulation FD Sample						
Mean of Regression Coefficients / *R*-squared	0.714***	-0.059***	-0.116***	-0.068***	-0.106***	25.59%
Standard Deviations of Regression Coefficients	0.588	0.367	0.551	0.459	0.456	
Q1 of Regression Coefficients	0.447	-0.206	-0.272	-0.218	-0.275	
Median of Regression Coefficients / *R*-squared	0.625	-0.009	-0.087	-0.042	-0.088	17.34%
Q3 of Regression Coefficients	0.890	0.145	0.069	0.104	0.063	
Panel C: Post Regulation FD Sample						
Mean of Regression Coefficients / *R*-squared	0.762***	-0.103***	-0.181***	-0.076***	-0.139***	36.77%
Standard Deviations of Regression Coefficients	0.625	0.453	0.486	0.442	0.445	
Q1 of Regression Coefficients	0.447	-0.330	-0.439	-0.290	-0.366	
Median of Regression Coefficients / *R*-squared	0.707	-0.069	-0.164	-0.068	-0.141	31.40%
Q3 of Regression Coefficients	0.996	0.151	0.082	0.141	0.074	

Table 2.3: Correlation between Information Weighting Factors and Firm Characteristics

The pairwise correlation coefficients among various stock characteristics are computed from the quarterly cross-section each quarter, and then the time-series average and the t-statistics are reported. The sample period in Panel A is from Q2/1985 to Q4/2005; and the sample period in Panel B is from Q1/1991 to Q4/2005. Past Ret is the past 12-month cumulative return during months (t-11, t). Past LT Ret is the cumulative return during months (t-35, t-12). Past TO is the average turnover during months (t-11, t), and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. DISP is the analysts' quarterly earning forecast dispersions during the quarter up to month t. SUE is the standardized unexpected earnings estimated from the seasonal random walk model and normalized into the percentile ranking. SEO is a binary indicator variable taking the value of one if the firm conducted a seasoned equity offering during previous 36 months, and zero otherwise. Future TO denotes the stock's average monthly turnover ratio during months (t+2, t+7). ILLQ is the NASDAQ volume-adjusted Amihud illiquidity measure during months (t-11, t). The market capitalization of the firm (Size) is the logarithm of market capitalization of the firm (in 1000's dollars) at month t. IWF is the value of herding tendency measure at formation month (t). B/M is the book to market equity during portfolio formation month t. News coverage is the logarithm of one plus the 12-month average news coverage of individual stocks during months (t-11, t).

Panel A: Time-series average and *t*-statistics of pair-wise correlation coefficient among variables, April 1985 to December 2005.

	IWF	Past LT Ret	Past Ret	B/M	Past TO	Size	ILLQ	DISP	SUE
A1: Time-series Average of Pairwise Correlation Coefficients									
IWF	1.00	-0.04	-0.01	0.09	-0.09	0.11	-0.07	0.11	0.01
Past LT Ret	-0.04	1.00	-0.01	-0.28	0.23	0.09	-0.09	-0.07	-0.06
Past Ret	-0.01	-0.01	1.00	-0.32	0.07	0.16	-0.13	-0.06	0.25
B/M	0.09	-0.28	-0.32	1.00	-0.11	-0.22	0.15	0.26	-0.16
Past TO	-0.09	0.23	0.07	-0.11	1.00	-0.30	-0.02	0.06	-0.02
Size	0.11	0.09	0.16	-0.22	-0.30	1.00	-0.42	-0.03	0.08
ILLQ	-0.07	-0.09	-0.13	0.15	-0.02	-0.42	1.00	0.00	-0.07
DISP	0.11	-0.07	-0.06	0.26	0.06	-0.03	0.00	1.00	-0.07
SUE	0.01	-0.06	0.25	-0.16	-0.02	0.08	-0.07	-0.07	1.00

A2: Simple *t*-statistics of time-series Average of Pairwise Correlation Coefficients

IWF	-	-5.08	-0.86	14.56	-13.71	17.89	-7.92	16.01	1.03
Past LT Ret	-5.08	-	-0.56	-30.59	9.67	5.50	-8.97	-8.50	-7.47
Past Ret	-0.86	-0.56	-	-24.05	2.98	10.96	-16.65	-5.15	27.94
B/M	14.56	-30.59	-24.05	-	-12.38	-18.88	16.13	29.02	-19.32
Past TO	-13.71	9.67	2.98	-12.38	-	-32.25	-3.27	8.19	-2.82
Size	17.89	5.50	10.96	-18.88	-32.25	-	-36.10	-3.85	9.08
ILLQ	-7.92	-8.97	-16.65	16.13	-3.27	-36.10	-	0.30	-12.29
DISP	16.01	-8.50	-5.15	29.02	8.19	-3.85	0.30	-	-6.66
SUE	1.03	-7.47	27.94	-19.32	-2.82	9.08	-12.29	-6.66	-

Panel B: Time-series average and *t*-statistics of pair-wise correlation coefficient among variables, January-1991 to December 2005.

	IWF	Past LT Ret	Past Ret	B/M	Past TO	Size	ILLQ	DISP	SUE	News Coverage
B1: Time-series Average of Pairwise Correlation Coefficients										
IWF	1.00	-0.07	-0.03	0.10	-0.11	0.11	-0.05	0.13	0.00	0.04
Past LT Ret	-0.07	1.00	-0.02	-0.29	0.29	0.05	-0.07	-0.09	-0.07	0.04
Past Ret	-0.03	-0.02	1.00	-0.31	0.07	0.14	-0.12	-0.05	0.24	-0.02
B/M	0.10	-0.29	-0.31	1.00	-0.14	-0.26	0.18	0.25	-0.13	0.00
Past TO	-0.11	0.29	0.07	-0.14	1.00	-0.29	-0.02	0.04	-0.04	0.04
Size	0.11	0.05	0.14	-0.26	-0.29	1.00	-0.40	-0.04	0.07	0.43
ILLQ	-0.05	-0.07	-0.12	0.18	-0.02	-0.40	1.00	0.01	-0.06	-0.14
DISP	0.13	-0.09	-0.05	0.25	0.04	-0.04	0.01	1.00	-0.05	0.10
SUE	0.00	-0.07	0.24	-0.13	-0.04	0.07	-0.06	-0.05	1.00	-0.03
News Coverage	0.04	0.04	-0.02	0.00	0.04	0.43	-0.14	0.10	-0.03	1.00

B2: Simple *t*-statistics of time-series Average of Pairwise Correlation Coefficients

IWF		-11.30	-4.11	17.02	-17.01	16.11	-6.43	22.29	-0.27	2.87
Past LT Ret	-11.30		-0.69	-30.47	12.36	3.01	-9.83	-11.42	-8.01	1.88
Past Ret	-4.11	-0.69		-23.03	2.55	7.33	-15.70	-3.91	27.74	-1.43
B/M	17.02	-30.47	-23.03		-13.45	-20.96	18.45	23.54	-20.00	-0.36
Past TO	-17.01	12.36	2.55	-13.45		-23.79	-3.87	5.47	-5.16	2.11
Size	16.11	3.01	7.33	-20.96	-23.79		-34.31	-5.81	8.57	10.10
ILLQ	-6.43	-9.83	-15.70	18.45	-3.87	-34.31		1.62	-11.13	-8.28
DISP	22.29	-11.42	-3.91	23.54	5.47	-5.81	1.62		-5.24	8.31
SUE	-0.27	-8.01	27.74	-20.00	-5.16	8.57	-11.13	-5.24		-3.71
News Coverage	2.87	1.88	-1.43	-0.36	2.11	10.10	-8.28	8.31	-3.71	

Table 2.4: Determinants of Information Weighting Factor

This table explores the relationship of between information weighting factor (IWF) and other characteristics. In all regressions, the dependent variable is the information weighting factor (IWF) value at formation month (t). Let t be the portfolio formation month. Past Ret is the past 12-month cumulative return during months (t-11, t). Past LT Ret is the cumulative return during months (t-35, t-12). The sample period from models 1 and 2 is from Q2/1985 to Q4/2005; and the sample period for model 3 is Q1/1991 to Q4/2005. The estimated regression coefficients and the t-statistics from the pooled cross-sectional time-series regression are reported. The standard errors are computed by clustering at the stock level; and the estimation method is ordinary least square (OLS).

	Model 1		**Model 2**		**Model 3**	
	Estimates	*t*-values	Estimates	*t*-values	Estimates	*t*-values
Intercept	-0.0268	-2.63	-0.0298	-3.01	-0.0621	-5.10
Past LT Ret	-0.0004	-0.38	-0.0003	-0.31	0.0000	-0.04
Past Ret	-0.0001	-0.03	-0.0008	-0.38	-0.0027	-1.19
B/M	0.0273	7.93	0.0208	6.43	0.0263	6.49
Past TO	-0.0942	-8.42	-0.0639	-4.57	-0.0551	-3.49
Size	0.0065	6.21	0.0065	6.42	0.0099	7.83
ILLQ	-0.0013	-1.59	-0.0012	-1.56	-0.0008	-1.26
DISP	0.0991	5.48	0.2583	13.55	0.3414	10.21
SUE	0.0017	0.79	0.0029	1.31	-0.0007	-0.23
IPO	0.0013	0.17	0.0036	0.49	-0.0037	-0.44
SEO	0.0065	2.45	0.0056	2.16	0.0083	2.54
Future TO			-0.0317	-3.12	-0.0278	-2.53
News Coverage					-0.0030	-2.25
Adj. R^2	4.09%		5.32%		5.59%	
N	43,358		42,658		28,329	

Table 2.5: Summary Statistics of Basic Portfolio Characteristics

This table reports the time-series average of stock characteristic at portfolio level. All values are computed at the portfolio formation month, unless otherwise stated. The sample period in this table is from Q2/1985 to Q5/2005. "R1" is the portfolio of 20 percent of the stocks with the lowest returns over the previous twelve months; R5 is the portfolio of the 20 percent of the stocks with the highest past twelve month returns, and so on. "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) predicted information weighting factor values. Panel A reports the information weighting factor values, past 12-month excess returns, average prices, percentages of NYSE, AMEX and NASDAQ capitalization, and number of stocks in each portfolio. In the calculation of percentages of capital of the NYSE, AMEX and NASDAQ, only the common shares are used. Panel B reports the time-series average of dollar value and percentile rankings of each portfolio. Each month, I compute the market capitalization percentile breakpoints for all common shares traded on NYSE, AMEX and NASDAQ. Each stock in the momentum portfolios interacted with the herding tendency measure is compared to the breakpoints to obtain the associated market capitalization percentile rankings. The average capitalization percentile rankings reported in the Panel B are the time-series average of such percentile rankings.

Panel A: Time-series average of past 12 month returns and IWF values double sorted portfolios: IWF values, past 12 month cumulative returns, prices, percentage of capital and number of stocks

		R1 (losers)	R2	R3	R4	R5 (winners)
	IWF1	-0.0862	-0.0591	-0.0498	-0.0550	-0.0768
IWF Values	IWF2	0.0360	0.0486	0.0517	0.0471	0.0354
	IWF3	0.1426	0.1426	0.1438	0.1400	0.1357
	IWF1	-26.59%	-2.44%	12.48%	29.04%	66.90%
Past 12 Month Returns	IWF2	-23.80%	-2.17%	12.38%	28.89%	63.02%
	IWF3	-23.67%	-2.36%	12.42%	28.47%	62.80%
	IWF1	20.88	32.62	39.29	42.55	43.27
Prices	IWF2	25.14	38.21	44.82	49.04	48.80
	IWF3	27.09	38.64	43.99	49.03	51.02
Percentage of NYSE, AMEX and NASDAQ Capitalization	IWF1	1.65%	3.06%	3.79%	4.15%	3.05%
	IWF2	3.23%	5.97%	6.79%	7.19%	5.97%
	IWF3	2.25%	3.72%	4.69%	4.64%	4.35%
	IWF1	33	33	33	33	33
Number of Stocks	IWF2	42	42	42	42	43
	IWF3	33	33	33	33	33

Panel B: Time-series average of past 12 month returns and IWF values double sorted portfolios: market capitalization in dollar values and percentile ranks

		R1 (losers)	R2	R3	R4	R5 (winners)
Market Capitalization (in percentile rank)	IWF1	56.02	69.55	74.60	75.96	70.95
	IWF2	65.15	77.05	80.46	81.27	77.49
	IWF3	63.48	74.66	77.89	78.92	76.27
Market Capitalization (in 1000's dollars)	IWF1	150,319	268,845	311,226	328,334	248,568
	IWF2	284,343	511,064	578,596	577,301	493,819
	IWF3	179,739	302,343	392,047	361,690	333,863

Table 2.6: Monthly Returns from Simple Momentum Strategy and Momentum Strategy Interacted with Herding Tendency Measures

The simple price momentum portfolios are formed based on 12-month lagged returns and held for 3, 6, 9, 12, 24 and 36 months, with skipping one month between the portfolio formation and return accumulation months. Only common shares traded on NYSE/AMEX/NASD with the end of month price greater than $5 and are selected into the portfolios. The portfolio returns are equally weighted returns. In the construction of price momentum portfolios interacting with the herding tendency measure (IWF), stocks in the intersection of I/B/E/S and CRSP are sorted into equally-spaced quintiles based on past twelve month returns; in each quintile, stocks are sequentially sorted into terciles based on the predicted information weighting factors, setting top thirty and bottom thirty percentiles as the breakpoints for the terciles. "R1" is the equal-weighted portfolio of 20 percent of the stocks with the lowest returns over the previous twelve months; "R5" is the equal-weighted portfolio of the 20 percent of the stocks with the highest past twelve month returns, and so on. "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) predicted information weighting factor values. The full sample period is April 1985 to December 2005. Panel A reports the average return and the associated t-statistics of each of the momentum portfolio constructed based on past 12 month returns with holding horizon ranging from 3 months to 36 months; Panel B reports the average return and the associated t-statistics of momentum strategy returns (past winner – past loser) for all months, January only and February to December; Panel C reports the monthly average returns of momentum portfolios interacted with the herding measure at different holding; Panel D reports the average return and the associated t-statistics of momentum strategy interacted with herding tendency measure for all months, January only and February to December.

Panel A: Monthly average returns of momentum portfolios with different holding horizons, April-1985 to December-2005.

Holding Horizon	R1 (Losers)	R2	R3	R4	R5 (Winners)
3 Months (W-L)	1.10%	1.17%	1.26%	1.37%	1.74%
6 Months (W-L)	1.22%	1.20%	1.31%	1.36%	1.67%
9 Months (W-L)	1.34%	1.23%	1.32%	1.33%	1.61%
12 Months (W-L)	1.43%	1.26%	1.32%	1.32%	1.54%
24 Months (W-L)	1.55%	1.32%	1.35%	1.32%	1.40%
36 Months (W-L)	1.50%	1.34%	1.37%	1.33%	1.41%

Panel B: Simple momentum strategy returns, April-1985 to December-2005.

	All Months	January Only	February - December	All Months	January Only	February - December
	Holding Horizon = 3 Months			Holding Horizon = 6 Months		
Mean (W-L)	0.63%	-1.21%	0.79%	0.45%	-1.34%	0.60%
t-statistics	1.75	-0.86	2.13	1.30	-1.23	1.67
	Holding Horizon = 9 Months			Holding Horizon = 12 Months		
Mean (W-L)	0.27%	-1.45%	0.42%	0.12%	-1.60%	0.27%
t-statistics	0.83	-1.63	1.23	0.39	-2.06	0.84
	Holding Horizon = 24 Months			Holding Horizon = 36 Months		
Mean (W-L)	-0.14%	-1.70%	-0.01%	-0.09%	-1.45%	0.03%
t-statistics	-0.59	-2.34	-0.02	-0.50	-2.22	0.17

Panel C: Monthly average returns of momentum portfolios interacted with the herding measure at different holding horizons, April 1985 to December 2005.

lding Horizon	IWF Rank	R1 (Losers)	R2	R3	R4	R5 (Winners)
	IWF1	0.80%	1.13%	1.28%	1.43%	1.94%
Ionths (W-L)	IWF2	1.15%	1.12%	1.27%	1.35%	1.65%
	IWF3	1.34%	1.28%	1.23%	1.34%	1.63%
	IWF1	0.99%	1.21%	1.42%	1.42%	1.83%
Ionths (W-L)	IWF2	1.26%	1.17%	1.25%	1.32%	1.63%
	IWF3	1.40%	1.23%	1.31%	1.36%	1.53%
	IWF1	1.19%	1.25%	1.38%	1.41%	1.77%
Ionths (W-L)	IWF2	1.37%	1.19%	1.30%	1.29%	1.58%
	IWF3	1.46%	1.25%	1.30%	1.31%	1.46%
	IWF1	1.32%	1.28%	1.37%	1.39%	1.70%
Months (W-L)	IWF2	1.43%	1.25%	1.31%	1.28%	1.53%
	IWF3	1.53%	1.25%	1.30%	1.29%	1.40%
	IWF1	1.41%	1.36%	1.40%	1.38%	1.55%
Months (W-L)	IWF2	1.57%	1.29%	1.35%	1.29%	1.39%
	IWF3	1.65%	1.32%	1.31%	1.29%	1.26%
	IWF1	1.38%	1.38%	1.42%	1.38%	1.54%
Months (W-L)	IWF2	1.53%	1.31%	1.37%	1.30%	1.40%
	IWF3	1.57%	1.34%	1.33%	1.31%	1.28%

Panel D: Monthly average returns of the strategy of momentum portfolios interacted with the herding measure at different holding horizons, April 1985 to December 2005.

	IWF Rank	All Month	January Only	February - December	All Month	January Only	February - December
		Holding Horizon = 3 Months			Holding Horizon = 6 Months		
W-L)	IWF1	1.14%	-0.33%	1.26%	0.85%	-0.71%	0.98%
ics		2.86	-0.22	3.07	2.31	-0.62	2.55
W-L)	IWF 2	0.50%	-1.76%	0.70%	0.37%	-1.48%	0.54%
ics		1.30	-1.28	1.74	1.02	-1.40	1.39
W-L)	IWF 3	0.29%	-1.42%	0.44%	0.13%	-1.80%	0.30%
ics		0.79	-0.89	1.18	0.38	-1.37	0.84
		Holding Horizon = 9 Months			Holding Horizon = 12 Months		
W-L)	IWF 1	0.59%	-1.37%	0.76%	0.38%	-1.67%	0.56%
ics		1.70	-1.34	2.09	1.17	-1.86	1.62
W-L)	IWF 2	0.22%	-1.34%	0.35%	0.10%	-1.39%	0.23%
ics		0.63	-1.56	0.96	0.31	-1.80	0.68
W-L)	IWF 3	0.00%	-1.69%	0.15%	-0.13%	-1.81%	0.02%
ics		0.01	-1.59	0.46	-0.44	-1.84	0.05
		Holding Horizon = 24 Months			Holding Horizon = 36 Months		
W-L)	IWF 1	0.15%	-1.25%	0.27%	0.16%	-0.95%	0.26%
ics		0.55	-1.52	0.96	0.78	-1.17	1.21
W-L)	IWF 2	-0.18%	-1.77%	-0.04%	-0.13%	-1.65%	0.00%
ics		-0.73	-2.06	-0.17	-0.71	-2.11	0.00
W-L)	IWF 3	-0.39%	-2.08%	-0.24%	-0.29%	-1.73%	-0.17%
ics		-1.65	-2.89	-0.98	-1.67	-3.04	-0.92

Table 2.7: Factor Regression Adjusted Monthly Returns from Simple Momentum Strategy and Momentum Strategy Interacted with Herding Tendency Measures

This table reports the factor-regression adjusted monthly returns from the simple momentum strategy (Panel A), and the factor-regression adjusted monthly returns from the momentum strategy interacted with the herding tendency measure (Panel B). The simple price momentum portfolios are formed based on 12-month lagged returns and held for 3, 6, 9, 12, 24 and 36 months, with skipping one month between the portfolio formation and return accumulation months. Only common shares traded on NYSE/AMEX/NASD with the end of month price greater than \$5 and are selected into the portfolios. The portfolio returns are equally weighted returns. In the construction of price momentum portfolios interacting with the herding tendency measure (IWF), stocks in the intersection of I/B/E/S and CRSP are sorted into equally-spaced quintiles based on past twelve month returns; in each quintile, stocks are sequentially sorted into terciles based on the predicted information weighting factors, setting top thirty and bottom thirty percentiles as the breakpoints for the terciles. In Panel B, "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) herding tendency measures (IWF). The first model is the Fama-French three factor model, and the intercept from the regression is denoted as "FF Alpha". Model 2 is the Fama-French three-factor model with the liquidity risk factor l, and the intercept from the regression is denoted as "FF + LIQ Alpha". Average factor adjusted monthly returns and associated t-statistics from all month, January-only and February-to-December portfolios are reported. The sample period is from April 1985 to December 2005.

Panel A: Factor adjusted returns from the simple momentum strategy; sample period is from April 1985 to December 2005.

All Months		Non January Months		January Only	
FF Alpha	FF + LIQ Alpha	FF Alpha	FF + LIQ Alpha	FF Alpha	FF + LIQ Alpha
Holding Horizon = 3 Months					
0.95%	0.63%	1.11%	0.77%	-0.16%	-0.66%
2.58	1.88	2.98	2.26	-0.10	-0.41
Holding Horizon = 6 Months					
0.83%	0.54%	0.98%	0.65%	-0.33%	-0.78%
2.45	1.73	2.78	2.03	-0.26	-0.63
Holding Horizon = 9 Months					
0.67%	0.41%	0.80%	0.52%	-0.56%	-0.96%
2.12	1.41	2.44	1.70	-0.53	-0.97
Holding Horizon = 12 Months					
0.51%	0.29%	0.64%	0.40%	-0.72%	-1.13%
1.77	1.06	2.11	1.39	-0.81	-1.44
Holding Horizon = 24 Months					
0.20%	0.07%	0.29%	0.16%	-0.64%	-0.93%
0.85	0.30	1.22	0.68	-0.89	-1.40
Holding Horizon = 36 Months					
0.18%	0.10%	0.26%	0.18%	-0.28%	-0.41%
1.13	0.64	1.57	1.11	-0.48	-0.71

Panel B: Factor adjusted returns of the momentum portfolio interacting with the herding tendency measure (IWF); sample period is from April 1985 to December 2005

IWF Ranking	All Months		Non January Months		January Only	
	FF Alpha	FF + LIQ Alpha	FF Alpha	FF + LIQ Alpha	FF Alpha	FF + LIQ Alpha
Holding Horizon = 3 Months						
IWF1	1.42%	1.08%	1.58%	1.21%	-0.15%	-0.69%
	3.49	2.90	3.81	3.18	-0.08	-0.39
IWF2	0.89%	0.55%	1.07%	0.71%	-0.23%	-0.77%
	2.29	1.56	2.70	1.95	-0.15	-0.56
IWF3	0.54%	0.26%	0.68%	0.39%	-0.09%	-0.50%
	1.44	0.75	1.81	1.10	-0.05	-0.26
Holding Horizon = 6 Months						
IWF1	1.21%	0.90%	1.36%	1.02%	-0.27%	-0.62%
	3.28	2.66	3.56	2.90	-0.20	-0.45
IWF2	0.83%	0.51%	0.97%	0.62%	-0.21%	-0.63%
	2.30	1.56	2.58	1.80	-0.17	-0.55
IWF3	0.46%	0.20%	0.60%	0.34%	-0.54%	-1.11%
	1.32	0.62	1.71	1.01	-0.34	-0.74
Holding Horizon = 9 Months						
IWF1	0.96%	0.69%	1.13%	0.83%	-0.78%	-1.07%
	2.79	2.17	3.15	2.50	-0.64	-0.87
IWF2	0.66%	0.38%	0.78%	0.46%	-0.36%	-0.73%
	1.97	1.23	2.21	1.43	-0.36	-0.77
IWF3	0.37%	0.15%	0.49%	0.26%	-0.59%	-1.13%
	1.20	0.53	1.53	0.86	-0.46	-0.98
Holding Horizon = 12 Months						
IWF1	0.77%	0.53%	0.93%	0.68%	-0.96%	-1.29%
	2.38	1.76	2.77	2.13	-0.93	-1.28
IWF2	0.53%	0.28%	0.64%	0.36%	-0.55%	-0.90%
	1.72	0.99	1.98	1.22	-0.59	-1.06
IWF3	0.23%	0.04%	0.35%	0.15%	-0.70%	-1.26%
	0.81	0.15	1.16	0.52	-0.60	-1.27
Holding Horizon = 24 Months						
IWF1	0.48%	0.35%	0.56%	0.44%	-0.30%	-0.56%
	1.83	1.37	2.06	1.62	-0.38	-0.72
IWF2	0.18%	0.04%	0.27%	0.13%	-0.61%	-0.93%
	0.74	0.16	1.10	0.52	-0.65	-1.04
IWF3	-0.07%	-0.19%	0.04%	-0.08%	-1.03%	-1.32%
	-0.31	-0.84	0.16	-0.34	-1.36	-1.84
Holding Horizon = 36 Months						
IWF1	0.43%	0.35%	0.49%	0.42%	0.17%	0.02%
	2.15	1.77	2.41	2.06	0.23	0.03
IWF2	0.15%	0.06%	0.24%	0.15%	-0.28%	-0.46%
	0.87	0.38	1.34	0.88	-0.37	-0.61
IWF3	-0.03%	-0.10%	0.06%	-0.02%	-0.75%	-0.80%
	-0.16	-0.65	0.37	-0.12	-1.34	-1.35

Table 2.8: Monthly Returns from the "Within-News" Strategy

This table reports the return spreads between high and low herding tendency measure values (IWF) portfolios conditional on past returns (i.e., "within-news" strategy) with holding horizons ranging from 3 months to 36 months. "R1" is the equal-weighted portfolio of 20 percent of the stocks with the lowest returns over the previous twelve months; "R5" is the equal-weighted portfolio of the 20 percent of the stocks with the highest past twelve month returns, and so on. "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) predicted information weighting factor values. Average monthly returns and associated t-statistics from all month, January-only and February-to-December portfolios are reported. The sample period is from April 1985 to December 2005.

	Return Rank	All Months	January Only	February - December	All Months	January Only	February - December
		Holding Horizon = 3 Months			Holding Horizon = 6 Months		
Mean (IWF3-IWF1)	R1	0.54%	0.31%	0.56%	0.42%	0.24%	0.43%
t-statistics	(losers)	2.74	0.50	2.69	2.33	0.44	2.29
Mean (IWF3-IWF1)	R2	0.15%	-0.73%	0.23%	0.02%	-0.42%	0.06%
t-statistics		1.21	-1.36	1.80	0.23	-0.84	0.59
Mean (IWF3-IWF1)	R3	-0.05%	-0.71%	0.01%	-0.10%	-0.69%	-0.05%
t-statistics		-0.37	-2.75	0.08	-0.98	-2.25	-0.47
Mean (IWF3-IWF1)	R4	-0.09%	-0.34%	-0.07%	-0.06%	0.02%	-0.07%
t-statistics		-0.71	-0.76	-0.51	-0.52	0.06	-0.56
Mean (IWF3-IWF1)	R5	-0.31%	-0.78%	-0.26%	-0.30%	-0.84%	-0.25%
t-statistics	(winners)	-1.73	-1.05	-1.46	-1.97	-1.27	-1.63
		Holding Horizon = 9 Months			Holding Horizon = 12 Months		
Mean (IWF3-IWF1)	R1	0.25%	-0.53%	0.32%	0.20%	-0.67%	0.27%
t-statistics	(losers)	1.52	-1.10	1.85	1.21	-1.36	1.61
Mean (IWF3-IWF1)	R2	0.02%	-0.60%	0.08%	0.00%	-0.60%	0.06%
t-statistics		0.25	-1.32	0.82	0.03	-1.33	0.59
Mean (IWF3-IWF1)	R3	-0.08%	-0.39%	-0.05%	-0.07%	-0.37%	-0.04%
t-statistics		-0.85	-1.32	-0.51	-0.83	-1.28	-0.48
Mean (IWF3-IWF1)	R4	-0.08%	-0.45%	-0.05%	-0.09%	-0.44%	-0.06%
t-statistics		-0.85	-1.19	-0.51	-0.93	-1.08	-0.59
Mean (IWF3-IWF1)	R5	-0.24%	-0.62%	-0.21%	-0.23%	-0.59%	-0.19%
t-statistics	(winners)	-1.74	-1.15	-1.44	-1.72	-1.15	-1.42
		Holding Horizon = 24 Months			Holding Horizon = 36 Months		
Mean (IWF3-IWF1)	R1	0.21%	-0.12%	0.24%	0.16%	-0.24%	0.20%
t-statistics	(losers)	1.52	-0.29	1.64	1.18	-0.48	1.40
Mean (IWF3-IWF1)	R2	-0.02%	-0.48%	0.02%	-0.03%	-0.45%	0.00%
t-statistics		-0.29	-1.41	0.20	-0.41	-1.30	0.05
Mean (IWF3-IWF1)	R3	-0.09%	-0.50%	-0.05%	-0.08%	-0.48%	-0.05%
t-statistics		-1.05	-1.24	-0.62	-1.04	-1.47	-0.59
Mean (IWF3-IWF1)	R4	-0.07%	-0.69%	-0.01%	-0.04%	-0.73%	0.02%
t-statistics		-0.76	-1.64	-0.14	-0.46	-1.90	0.25
Mean (IWF3-IWF1)	R5	-0.24%	-0.73%	-0.20%	-0.22%	-0.88%	-0.16%
t-statistics	(winners)	-2.10	-1.73	-1.66	-1.99	-2.38	-1.38

Table 2.9: Monthly Returns and Factor Regression Adjusted Monthly Returns from the "Within-News" Strategy

This table reports the factor-model adjusted monthly return spreads between high and low herding tendency measure values (IWF) portfolios conditional on past returns with holding horizons ranging from 3 months to 36 months. "R1" is the equal-weighted portfolio of 20 percent of the stocks with the lowest returns over the previous twelve months; "R5" is the equal-weighted portfolio of the 20 percent of the stocks with the highest past twelve month returns, and so on. "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) herding tendency measure values. Average monthly returns and associated t-statistics from all month, January-only and February-to-December portfolios are reported. Model 1 is the Fama-French three factor model. Model 2 is the Fama-French three-factor model with short-term reversal factor, STREV, constructed as in Jegadeesh (1990) and Lehmann (1990). Model 3 is the Fama-French three-factor model with the liquidity risk factor LIQ. The sample period is from April 1985 to December 2005.

Panel A: Factor regression adjusted monthly return, all months; sample period is from April 1985 to December 2005

	Return Rank	FF Alpha	FF + LIQ Alpha	FF + TREV Alpha	FF Alpha	FF + PS Alpha	FF + TREV Alpha
		Holding Horizon = 3 Months			Holding Horizon = 6 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.60%	0.52%	0.62%	0.47%	0.41%	0.48%
t-statistics		3.10	2.71	3.18	2.64	2.30	2.72
Mean (IWF3-IWF1)	R2	0.16%	0.16%	0.17%	0.03%	0.02%	0.04%
t-statistics		1.27	1.22	1.31	0.32	0.20	0.35
Mean (IWF3-IWF1)	R3	0.02%	-0.02%	0.06%	-0.06%	-0.08%	-0.03%
t-statistics		0.13	-0.18	0.49	-0.56	-0.80	-0.29
Mean (IWF3-IWF1)	R4	-0.14%	-0.16%	-0.11%	-0.08%	-0.10%	-0.06%
t-statistics		-1.10	-1.26	-0.91	-0.73	-0.89	-0.57
Mean (IWF3-IWF1)	R5 (Winners)	-0.28%	-0.29%	-0.28%	-0.28%	-0.29%	-0.29%
t-statistics		-1.60	-1.68	-1.61	-1.91	-1.96	-1.96
		Holding Horizon = 9 Months			Holding Horizon = 12 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.26%	0.28%	0.26%	0.23%	0.23%	0.23%
t-statistics		1.67	1.73	1.61	1.43	1.47	1.42
Mean (IWF3-IWF1)	R2	0.03%	0.04%	0.02%	0.00%	0.01%	-0.02%
t-statistics		0.35	0.40	0.19	-0.01	0.07	-0.16
Mean (IWF3-IWF1)	R3	-0.07%	-0.05%	-0.06%	-0.08%	-0.07%	-0.08%
t-statistics		-0.80	-0.60	-0.74	-1.03	-0.87	-0.99
Mean (IWF3-IWF1)	R4	-0.12%	-0.10%	-0.11%	-0.12%	-0.11%	-0.12%
t-statistics		-1.26	-1.07	-1.14	-1.34	-1.16	-1.23
Mean (IWF3-IWF1)	R5 (Winners)	-0.24%	-0.24%	-0.25%	-0.23%	-0.23%	-0.24%
t-statistics		-1.79	-1.81	-1.82	-1.80	-1.78	-1.84
		Holding Horizon = 24 Months			Holding Horizon = 36 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.24%	0.24%	0.24%	0.18%	0.19%	0.19%
t-statistics		1.74	1.80	1.73	1.41	1.45	1.43
Mean (IWF3-IWF1)	R2	-0.05%	-0.04%	-0.06%	-0.06%	-0.06%	-0.07%
t-statistics		-0.58	-0.50	-0.69	-0.85	-0.75	-0.87
Mean (IWF3-IWF1)	R3	-0.13%	-0.12%	-0.14%	-0.14%	-0.14%	-0.15%
t-statistics		-1.70	-1.55	-1.70	-1.97	-1.85	-1.96
Mean (IWF3-IWF1)	R4	-0.11%	-0.09%	-0.10%	-0.08%	-0.07%	-0.08%
t-statistics		-1.28	-1.09	-1.18	-1.04	-0.88	-0.97
Mean (IWF3-IWF1)	R5 (Winners)	-0.24%	-0.24%	-0.24%	-0.21%	-0.20%	-0.21%
t-statistics		-2.16	-2.13	-2.19	-2.02	-1.98	-2.10

Panel B: Factor regression adjusted monthly return, non-January months only; sample period is from April 1985 to December 2005

	Return Rank	FF Alpha	FF + PS Alpha	FF + STREV Alpha	FF Alpha	FF + PS Alpha	FF + STREV Alpha
		Holding Horizon = 3 Months			Holding Horizon = 6 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.61%	0.61%	0.60%	0.48%	0.47%	0.46%
t-statistics		3.01	3.00	2.96	2.56	2.54	2.49
Mean (IWF3-IWF1)	R2	0.23%	0.23%	0.23%	0.06%	0.06%	0.06%
t-statistics		1.77	1.76	1.75	0.61	0.61	0.56
Mean (IWF3-IWF1)	R3	0.08%	0.06%	0.06%	-0.01%	-0.02%	-0.02%
t-statistics		0.63	0.48	0.43	-0.06	-0.15	-0.18
Mean (IWF3-IWF1)	R4	-0.13%	-0.14%	-0.14%	-0.11%	-0.11%	-0.11%
t-statistics		-1.02	-1.06	-1.06	-0.95	-0.99	-0.97
Mean (IWF3-IWF1)	R5 (Winners)	-0.29%	-0.29%	-0.29%	-0.29%	-0.28%	-0.28%
t-statistics		-1.62	-1.62	-1.62	-1.91	-1.90	-1.88
		Holding Horizon = 9 Months			Holding Horizon = 12 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.33%	0.33%	0.33%	0.29%	0.29%	0.30%
t-statistics		1.97	1.97	1.95	1.77	1.77	1.81
Mean (IWF3-IWF1)	R2	0.08%	0.07%	0.06%	0.04%	0.04%	0.02%
t-statistics		0.77	0.76	0.58	0.42	0.40	0.24
Mean (IWF3-IWF1)	R3	-0.04%	-0.05%	-0.04%	-0.06%	-0.06%	-0.06%
t-statistics		-0.49	-0.54	-0.47	-0.72	-0.76	-0.72
Mean (IWF3-IWF1)	R4	-0.10%	-0.11%	-0.10%	-0.11%	-0.11%	-0.11%
t-statistics		-1.06	-1.11	-1.03	-1.13	-1.17	-1.10
Mean (IWF3-IWF1)	R5 (Winners)	-0.24%	-0.24%	-0.24%	-0.22%	-0.22%	-0.23%
t-statistics		-1.72	-1.72	-1.75	-1.69	-1.69	-1.71
		Holding Horizon = 24 Months			Holding Horizon = 36 Months		
Mean (IWF3-IWF1)	R1 (Losers)	0.25%	0.25%	0.26%	0.20%	0.20%	0.21%
t-statistics		1.77	1.76	1.81	1.50	1.51	1.60
Mean (IWF3-IWF1)	R2	-0.02%	-0.02%	-0.02%	-0.04%	-0.04%	-0.04%
t-statistics		-0.19	-0.21	-0.30	-0.50	-0.52	-0.55
Mean (IWF3-IWF1)	R3	-0.10%	-0.10%	-0.10%	-0.11%	-0.11%	-0.11%
t-statistics		-1.23	-1.26	-1.25	-1.45	-1.47	-1.47
Mean (IWF3-IWF1)	R4	-0.06%	-0.06%	-0.05%	-0.03%	-0.03%	-0.03%
t-statistics		-0.70	-0.74	-0.65	-0.40	-0.42	-0.36
Mean (IWF3-IWF1)	R5 (Winners)	-0.21%	-0.21%	-0.21%	-0.17%	-0.17%	-0.17%
t-statistics		-1.85	-1.86	-1.88	-1.58	-1.60	-1.65

Table 2.10: Additional Portfolio Characteristics

This table reports additional time-series average of portfolio characteristics including illiquidity, analyst forecasts biases, news media coverage, information uncertainty and limit-to-arbitrage. The sample period in this table is from Q2/1985 to Q5/2005 (except news media coverage, which is from Q1/1991 to Q4/2005). "R1" is the equal-weighted portfolio of 20 percent of the stocks with the lowest returns over the previous twelve months; "R5" is the equal-weighted portfolio of the 20 percent of the stocks with the highest past twelve month returns, and so on. "IWF1" ("IWF3") subsamples comprise of stocks with lowest 30 percent (and highest 30 percent) herding tendency measure values. Panel A reports the NASDAQ volume-adjusted Amihud measure at portfolio formation month, and 12 month average turnover ratio including portfolio formation month. Panel B reports the median forecast errors (in dollar terms) during portfolio formation quarter, and the analyst forecast dispersion (in dollar terms) during the quarter up to the month of actual quarterly earning announcement. The median forecasts are computed 1-month prior to the actual quarterly earning announcements up to the portfolio formation month. The analysts' quarterly earning forecast dispersions is computed as the standard deviation of analyst forecasts during the quarter up to month t. Panel C reports the time-series average of 12-month size adjusted average including formation month. Panel D reports the time-series average of book to market equity ratio during formation month. Panel E reports the two arbitrage risks measures, in terms of the residual return variance (multiplied by 1000, ARBRISK) and the total return variance (multiplied by 1000, VOLA).

Panel A: Time-series average of past 12 month returns and IWF values double sorted portfolios – liquidity characteristics

		R1 (losers)	R2	R3	R4	R5 (winners)
Average Amihud Illiquidity Measure (12 month average)	IWF1	0.1005	0.0567	0.0449	0.0398	0.0547
	IWF2	0.0536	0.0335	0.0276	0.0270	0.0340
	IWF3	0.0603	0.0384	0.0350	0.0336	0.0397
Average Turnover Ratio (12 month average)	IWF1	0.1349	0.0868	0.0778	0.0845	0.1487
	IWF2	0.1166	0.0766	0.0695	0.0757	0.1305
	IWF3	0.1177	0.0789	0.0727	0.0787	0.1197

Panel B: Time-series average of past 12 month returns and IWF values double sorted portfolios – analysts forecast errors and forecast dispersions

		R1 (losers)	R2	R3	R4	R5 (winners)
Median Forecast Errors (quarterly average up to formation month)	IWF1	-0.0163	-0.0031	0.0022	0.0053	0.0104
	IWF2	-0.0097	-0.0023	0.0040	0.0087	0.0129
	IWF3	-0.0061	-0.0002	0.0057	0.0090	0.0126
Forecast Dispersion (up to formation month)	IWF1	0.0367	0.0341	0.0271	0.0254	0.0264
	IWF2	0.0458	0.0432	0.0391	0.0388	0.0368
	IWF3	0.0581	0.0498	0.0469	0.0432	0.0469

Panel C: Time-series average of past 12 month returns and IWF values double sorted portfolios – size adjusted news media coverage

		R1 (losers)	R2	R3	R4	R5 (winners)
Size Adjusted News Coverage (Median, Past 12 Month)	IWF1	0.12	-0.54	-0.76	-0.72	-0.10
	IWF2	0.09	-0.73	-0.88	-0.83	-0.42
	IWF3	0.10	-0.61	-0.85	-0.85	-0.41

Panel D: Time-series average of past 12 month returns and IWF values double sorted portfolios: book equity and book to market equity

		R1 (losers)	R2	R3	R4	R5 (winners)
Book to market equity ratio (at formation month)	IWF1	0.70	0.54	0.47	0.42	0.30
	IWF2	0.76	0.61	0.55	0.47	0.36
	IWF3	0.77	0.66	0.59	0.51	0.40

Panel E: Time-series average of past 12 month returns and IWF values double sorted portfolios – arbitrage risk and total return volatilities

		R1 (losers)	R2	R3	R4	R5 (winners)
ARBRISK (at formation month)	IWF1	0.8417	0.5324	0.4610	0.4827	0.7510
	IWF2	0.6725	0.4210	0.3718	0.3895	0.6353
	IWF3	0.6847	0.4272	0.3789	0.4087	0.6125
VOLA (at formation month)	IWF1	0.9954	0.6416	0.5621	0.5883	0.8831
	IWF2	0.8130	0.5259	0.4687	0.4883	0.7630
	IWF3	0.8141	0.5267	0.4693	0.5000	0.7236

Table 2.11: Fama-MacBeth Regressions of Returns on Past Returns, Interaction of Past Returns and Herding Tendency Measure and Stock Characteristics Controls

Fama and MacBeth (1973) cross-sectional regressions are run each month from April 1985 to December 2005. The average slope coefficients are reported as the Fama-MacBeth regression coefficients. Let t denote the portfolio formation month. In all regressions, the dependent variable is the individual stock 6-month cumulative return during months (t+2, t+8). Past Ret is the cumulative past return of stocks during months (t-12, t-1), and Past LT Ret is the cumulative return during (t-36, t-13). IWF (i) is the tercile ranking indicator function taking the value of one if the ranking of IWF belongs to the i-th tercile (i = 1, 2 and 3) and zero otherwise. B/M is the book to market equity of the firm at the end of month t. Size is the logarithm of market capitalization of the firm at the portfolio formation month (t). Past TO is the average turnover ratio of the stock during months (t-12, t-1), and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. DISP is the analysts' quarterly earning forecast dispersions during the quarter up to month t. SUE is the standardized unexpected earnings estimated from seasonal random walk model and normalized into the percentile ranking following Chan, Jegadeesh and Lakonishok (1996). SEO is a binary indicator variable taking the value of one if the firm conducted a seasoned equity offering during previous 36 months, and zero otherwise. Future TO is the stock's average monthly turnover ratio during months (t+2, t+8), and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. In all the regressions, both dependent and independent variables are first demeaned then divided by the cross-sectional standard deviations each month. Panel A reports the Fama-MacBeth regressions of individual stock momentum return and firm characteristic control variables; Panel B reports Fama-MacBeth regressions of individual stock momentum return conditional on information weighting factor; Panel C reports Fama-MacBeth regressions of individual stock momentum return conditional on information weighting factor and firm characteristic control variables. Newey-West HAC (with five lags) standard errors are computed to obtain the t-statistic. The adjusted R2 is the time-series average of each cross-sectional regression's adjusted R2.

Panel A: Fama-MacBeth regressions of individual stock momentum return and stock characteristics

Past Ret	Past LT Ret	BM	Past TO	Size	ILLQ	DISP	SUE	SEO	Future TO	*Adj. R^2*
Model 1: six month return										
0.0405										*2.79%*
1.94										
Model 2: six month return with set of firm characteristic control variables										
0.0484	-0.0224	0.0180	0.0234	-0.0059	0.0176	-0.0078	0.0160	-0.0255		*11.17%*
2.87	-1.47	1.09	1.01	-0.32	1.66	-0.86	1.98	-3.15		
Model 3: six month return with set of firm characteristic control variables and future turnover as additional control										
0.0471	-0.0155	0.0197	-0.0131	-0.0029	0.0175	-0.0053	0.0146	-0.0251	0.0429	*12.17%*
2.84	-1.12	1.20	-0.47	-0.16	1.71	-0.60	1.79	-3.11	1.52	

Panel B: Fama-MacBeth regressions of individual stock momentum returns conditional on information weighting factor

Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	*Adj. R^2*
0.0422	0.0245	0.0104	0.0018	-0.0040	0.0058	*3.58%*
2.97	1.52	0.81	0.26	-0.73	0.69	

Panel C: Fama-MacBeth regressions of individual stock momentum returns conditional on information weighting factor and stock characteristics

Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	Past LT Ret	BM	Past TO	Size	ILLQ	DISP	SUE	SEO	Future TO	*Adj. R^2*
Model 1: six month return, without future turnover as additional control															
0.0487	0.0303	0.0155	0.0073	-0.0056	0.0021	-0.0223	0.0159	0.0231	-0.0056	0.0211	-0.0070	0.0139	-0.0246		*11.69%*
4.14	2.21	1.47	1.07	-1.04	0.26	-1.46	0.99	0.98	-0.30	2.16	-0.79	1.73	-3.04		
Model 2: six month return, with future turnover as additional control															
0.0473	0.0286	0.0163	0.0074	-0.0055	0.0013	-0.0153	0.0176	-0.0135	-0.0024	0.0204	-0.0042	0.0130	-0.0243	0.0428	*12.69%*
4.15	2.07	1.56	1.11	-1.02	0.17	-1.10	1.09	-0.48	-0.13	2.16	-0.49	1.59	-3.01	1.50	

Table 2.12: Time-Series and Cross-Sectional Regressions of Returns on Past Returns, Interaction of Past Returns and Herding Tendency Measure and Stock Characteristic Controls

This table reports the time-series and cross-sectional regressions of the sample from April 1985 to December 2005. Let t denote the portfolio formation month. In all regressions, the dependent variable is the individual stock 6-month cumulative return during months (t+2, t+8). Past Ret is the cumulative past return of stocks during months (t-12, t-1), and Past LT Ret is the cumulative return during (t-36, t-13). IWF (i) is the tercile ranking indicator function taking the value of one if the ranking of IWF belongs to the i-th tercile ($i = 1$, 2 and 3) and zero otherwise. B/M is the book to market equity of the firm at the end of month t. Size is the logarithm of market capitalization of the firm at the portfolio formation month (t). Past TO is the average turnover ratio of the stock during months (t-12, t-1), and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. DISP is the analysts' quarterly earning forecast dispersions during the quarter up to month t. SUE is the standardized unexpected earnings estimated from seasonal random walk model and normalized into the percentile ranking following Chan, Jegadeesh and Lakonishok (1996). SEO is a binary indicator variable taking the value of one if the firm conducted a seasoned equity offering during previous 36 months, and zero otherwise. Future TO is the stock's average monthly turnover ratio during months (t+2, t+8), and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. In all the regressions, both dependent and independent variables are first demeaned then divided by the cross-sectional standard deviations each month. Panel A reports the pooled time-series and cross-sectional regressions of individual stock momentum return and firm characteristic control variables; Panel B reports the pooled time-series and cross-sectional regressions of individual stock momentum return conditional on information weighting factor; Panel C reports the pooled time-series and cross-sectional regressions of individual stock momentum return conditional on information weighting factor and firm characteristic control variables. The spatial HAC standard errors are computed to obtain the t-statistics, where the "economic distance" in the cross-sections is the Euclidian norm of Z-scores of the past 12 month return, size and book-to-market equity, and the time-dimension includes five lags (see appendix B for details about calculations of spatial HAC standard errors). The adjusted R2 is the pooled time-series cross-sectional regression R2 adjusted by degree of freedoms.

Panel A: Time-series cross-sectional regressions of individual stock momentum returns and characteristics

ıst Ret	Past LT Ret	BM	Past TO	Size	ILLQ	DISP	SUE	SEO	Future TO	*Adj. R^2*
odel 1: six month return (N = 133,540)										
.0394										*0.15%*
2.13										
odel 2: six month return with set of firm characteristic control variables (N=120,913)										
.0450	-0.0224	0.0187	0.0200	-0.0192	0.0153	0.0069	0.0126	-0.0272		*0.46%*
2.22	-1.44	1.50	0.75	-1.02	1.75	0.77	1.58	-2.75		
odel 3: six month return with set of firm characteristic control variables and future turnover as additional control (N=120,913)										
.0438	-0.0222	0.0189	0.0086	-0.0186	0.0155	0.0066	0.0123	-0.0273	0.0135	*0.47%*
2.15	-1.43	1.51	0.32	-0.97	1.75	0.75	1.53	-2.78	0.49	

Panel B: Time-series cross-sectional regressions of individual stock momentum returns conditional on information weighting factor

Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	*Adj. R^2*
Model 1: six month return (N= 133,540)						
0.0324	0.0265	0.0143	0.0063	-0.0025	0.0012	*0.17%*
2.67	1.91	1.23	0.94	-0.50	0.20	

Panel C: Time-series cross-sectional regressions of individual stock momentum returns conditional on information weighting factor and characteristics

Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	Past LT Ret	BM	Past TO	Size	ILLQ	DISP	SUE	SEO	Future TO	*Adj. R^2*
Model 1: six month return, without future turnover as additional control (N=120,913)															
0.0345	0.0286	0.0216	0.0115	-0.0025	-0.0032	-0.0220	0.0183	0.0203	-0.0211	0.0160	0.0064	0.0124	-0.0271		*0.48%*
2.60	1.87	1.69	1.83	-0.49	-0.53	-1.41	1.46	0.76	-1.10	1.83	0.73	1.55	-2.75		
Model 2: six month return, with future turnover as additional control (N=120,913)															
0.0336	0.0279	0.0210	0.0115	-0.0025	-0.0031	-0.0219	0.0184	0.0095	-0.0205	0.0162	0.0061	0.0122	-0.0272	0.0128	*0.48%*
2.59	1.88	1.68	1.87	-0.49	-0.51	-1.43	1.50	0.35	-1.10	1.82	0.71	1.54	-2.88	0.47	

Table 2.13: Fama-MacBeth Regressions of Returns on Interaction of Past Returns Ranking and Herding Tendency Measure, and Stock Characteristics Controls

Fama and MacBeth (1973) cross-sectional regressions are run each month from April 1985 to December 2005. The average slope coefficients are reported as the Fama-MacBeth regression coefficients. Let t denote the portfolio formation month. In all regressions, the dependent variable is the individual stock cumulative return during months (t+2, t+2+J), where J = 1, 3, and 6. RET (i) is the quintile ranking indicator variable taking the value of one if the ranking of past 12 month returns belongs to the i-th quintile (i = 1, 2, 3, 4 and 5) and zero otherwise. Past Ret is the cumulative past return of stocks during months (t-12, t-1). IWF is the value of information weighting factor value at portfolio formation month. Only the regression coefficients for the interaction between past return ranking and past return, and the interaction between past return ranking and information weighting factor values are reported. In all regressions, the set of control variables include Past LT Ret, B/M, Size, Past TO, DISP, SUE, SEO and Future TO. Except the definition of Future TO, other variables are similarly defined as in previous table. Future TO is the stock's average monthly turnover ratio during months (t+2, t+2+J), where J = 1, 3, and 6, and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. In all the regressions, both dependent and independent variables are first demeaned then divided by the cross-sectional standard deviations each month. Newey-West HAC (with zero, three and five lags respectively) standard errors are computed to obtain the t-statistic. The adjusted R2 is the time-series average of each cross-sectional regression's adjusted R^2.

IWF x RET(*i*=2)	IWF x RET(*i*=3)	IWF x RET(*i*=4)	IWF x RET(*i*=5)	Past Ret x RET(*i*=1)	Past Ret x RET(*i*=2)	Past Ret x RET(*i*=3)	Past Ret x RET(*i*=4)	Past Ret x RET(*i*=5)	*Adj.* R^2
0.0078	0.0031	0.0018	-0.0114	0.0173	0.0044	0.0133	0.0195	0.0494	12.39%
2.08	0.83	0.47	-2.14	2.07	0.87	2.33	3.21	5.34	
0.0123	0.0003	-0.0038	-0.0132	0.0170	0.0080	0.0272	0.0385	0.0779	13.53%
2.49	0.06	-1.05	-1.73	1.43	1.34	3.48	4.86	5.78	
0.0115	0.0003	-0.0068	-0.0109	-0.0107	-0.0008	0.0253	0.0416	0.0844	13.20%
2.32	0.06	-1.51	-1.49	-0.70	-0.10	2.87	3.95	4.91	

Table 2.14: Time-Series Cross-Sectional Regressions of Returns on Interaction of Past Returns Ranking and Herding Tendency Measure, and Stock Characteristics Controls

This table reports the pooled time-series and cross-sectional regressions of the sample from April 1985 to December 2005. Let t denote the portfolio formation month. In all regressions, the dependent variable is the individual stock cumulative return during months (t+2, t+2+J), where J = 1, 3, and 6. RET (i) is the quintile ranking indicator variable taking the value of one if the ranking of past 12 month returns belongs to the i-th quintile (i = 1, 2, 3, 4 and 5) and zero otherwise. Past Ret is the cumulative past return of stocks during months (t-12, t-1). IWF is the value of information weighting factor value at portfolio formation month. Only the regression coefficients for the interaction between past return ranking and past return, and the interaction between past return ranking and information weighting factor values are reported. In all regressions, the set of control variables include Past LT Ret, B/M, Size, Past TO, DISP, SUE, SEO and Future TO. Except the definition of Future TO, other variables are similarly defined as in previous table. Future TO is the stock's average monthly turnover ratio during months (t+2, t+2+J), where J = 1, 3, and 6, and the monthly stock turnover ratios are adjusted by stock exchange where the stock is listed. In all the regressions, both dependent and independent variables are first demeaned then divided by the cross-sectional standard deviations each month. The spatial HAC standard errors are computed to obtain the t-statistics, where the "economic distance" in the cross-sections is the Euclidian norm of Z-scores of the past 12 month return, size and book-to-market equity, and the time-dimension includes one, two and five lags (see appendix B for details about calculations of spatial HAC standard errors). The adjusted R2 is the pooled time-series cross-sectional regression R^2 adjusted by degree of freedoms. The number of observations in these regressions is 120,913 stock-months.

IWF x RET(*i=2*)	IWF x RET(*i=3*)	IWF x RET(*i=4*)	IWF x RET(*i=5*)	Past Ret x RET(*i=1*)	Past Ret x RET(*i=2*)	Past Ret x RET(*i=3*)	Past Ret x RET(*i=4*)	Past Ret x RET(*i=5*)	*Adj.* R^2
0.0076	0.0034	-0.0059	-0.0023	0.0189	-0.0007	0.0100	0.0268	0.0398	*0.33%*
2.27	1.11	-2.18	-0.55	2.58	-0.16	2.60	5.89	4.83	
0.0075	0.0032	-0.0040	-0.0010	0.0257	0.0014	0.0136	0.0323	0.0526	*0.76%*
1.87	0.87	-1.27	-0.19	2.34	0.27	2.70	5.18	4.31	
0.0027	-0.0018	-0.0048	0.0032	0.0027	-0.0037	0.0176	0.0343	0.0605	*0.61%*
0.64	-0.44	-1.36	0.54	0.19	-0.49	2.84	4.08	3.57	

Table 2.15: Additional Robustness Checks

This table reports additional robustness checks. Panel A reports the Fama-MacBeth regressions of individual stock momentum returns conditional on information weighting factor and stock characteristics control variables including Past LT Ret, B/M, Size, Past TO, DISP, SUE, SEO and Future TO. Panel B reports the Fama-MacBeth regressions of individual stock information weighting factor conditional last return rankings and stock characteristics control variables including Past LT Ret, B/M, Size, Past TO, DISP, SUE, SEO and Future TO. Panel D is similar to Panel B, and it reports the sub-sample period regressions from January 1990 to December 2005. Regression coefficients and associated t-statistics for control variables are not reported for brevity. In Panels A and B, all stocks with the standardized unexpected earnings (SUE) percentile ranking greater than 90th percentile or lower than 10th percentile during the portfolio formation quarters are excluded in the monthly regressions. Similar to Panel A, Panel C reports the subsample period regressions from January 1990 to December 2005. In all the regressions, both dependent and independent variables are first demeaned then divided by the cross-sectional standard deviations each month. Newey-West HAC (with zero, three and five lags respectively) standard errors are computed to obtain the t-statistic. The adjusted R^2 is the time-series average of each cross-sectional regression's adjusted R^2.

Panel A: Fama-MacBeth regressions of individual stock momentum returns conditional on information weighting factor and stock characteristics, excluding top and bottom 10% SUE stocks.

	Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	*Adj. R^2*
6 month	0.0348	0.0264	0.0176	0.0129	-0.0113	-0.0038	*13.24%*
	3.41	1.78	1.75	1.63	-1.33	-0.52	

Panel B: Fama-MacBeth regressions of individual stock information weighting factor conditional last return rankings and stock characteristics, excluding top and bottom 10% SUE stocks.

	IWF x RET(*i=1*)	IWF x RET(*i=2*)	IWF x RET(*i=3*)	IWF x RET(*i=4*)	IWF x RET(*i=5*)	Past Ret x RET(*i=1*)	Past Ret x RET(*i=2*)	Past Ret x RET(*i=3*)	Past Ret x RET(*i=4*)	Past Ret x RET(*i=5*)	*Adj. R^2*
1 month	0.0125	0.0080	0.0038	-0.0007	-0.0169	0.0239	0.0037	0.0131	0.0240	0.0563	*12.57%*
	2.26	1.91	0.81	-0.18	-1.57	2.67	0.64	1.85	3.27	3.73	
3 month	0.0225	0.0133	-0.0027	-0.0085	-0.0191	0.0225	0.0133	-0.0027	-0.0085	-0.0191	*13.58%*
	3.27	2.51	-0.58	-1.77	-1.42	3.27	2.51	-0.58	-1.77	-1.42	
6 month	0.0161	0.0110	-0.0033	-0.0077	-0.0222	-0.0042	-0.0026	0.0273	0.0474	0.0969	*13.12%*
	1.72	1.84	-0.63	-1.49	-1.21	-0.24	-0.31	2.67	3.82	3.54	

Panel C: Fama-MacBeth regressions of individual stock momentum returns conditional on information weighting factor and stock characteristics, January 1990 to December 2005

	Past Ret x IWF(*i=1*)	Past Ret x IWF(*i=2*)	Past Ret x IWF(*i=3*)	IWF x IWF(*i=1*)	IWF x IWF(*i=2*)	IWF x IWF(*i=3*)	*Adj. R^2*
6 month	0.0474	0.0285	0.0163	0.0074	-0.0054	0.0014	*12.71%*
	4.27	2.13	1.62	1.14	-1.02	0.19	

Panel D: Fama-MacBeth regressions of individual stock information weighting factor conditional last return rankings and stock characteristics, January 1990 to December 2005

	IWF x RET(*i=1*)	IWF x RET(*i=2*)	IWF x RET(*i=3*)	IWF x RET(*i=4*)	IWF x RET(*i=5*)	Past Ret x RET(*i=1*)	Past Ret x RET(*i=2*)	Past Ret x RET(*i=3*)	Past Ret x RET(*i=4*)	Past Ret x RET(*i=5*)	*Adj. R^2*
1 month	0.0084	0.0107	-0.0015	-0.0055	-0.0093	0.0170	0.0028	0.0107	0.0174	0.0413	11.50%
	2.06	3.00	-0.47	-1.92	-2.14	2.07	0.65	2.28	3.48	4.63	
3 month	0.0128	0.0133	-0.0031	-0.0036	-0.0086	0.0176	0.0023	0.0156	0.0242	0.0579	12.66%
	2.47	3.01	-0.90	-1.04	-1.53	1.40	0.40	2.76	3.28	4.68	
6 month	0.0082	0.0070	-0.0063	-0.0066	-0.0061	-0.0113	-0.0048	0.0165	0.0275	0.0609	12.32%
	1.31	1.53	-1.69	-1.81	-1.03	-0.68	-0.62	2.53	2.82	3.47	

Figure 2.1: Past Returns and Information Weighting Factor Double Sorted Portfolio Formation

This figure illustrates the time line of estimating the analyst herding tendency measure of individual stocks, and forming portfolios based on the interaction of past 12-month returns and analyst herding tendency measure. The herding tendency measures are estimated based on past four quarter's raw information weighting factor values. The holding horizon of momentum portfolios ranges from J = 3 months to J = 36 months. There is one month lag between portfolio formation month and the beginning of return accumulation month.

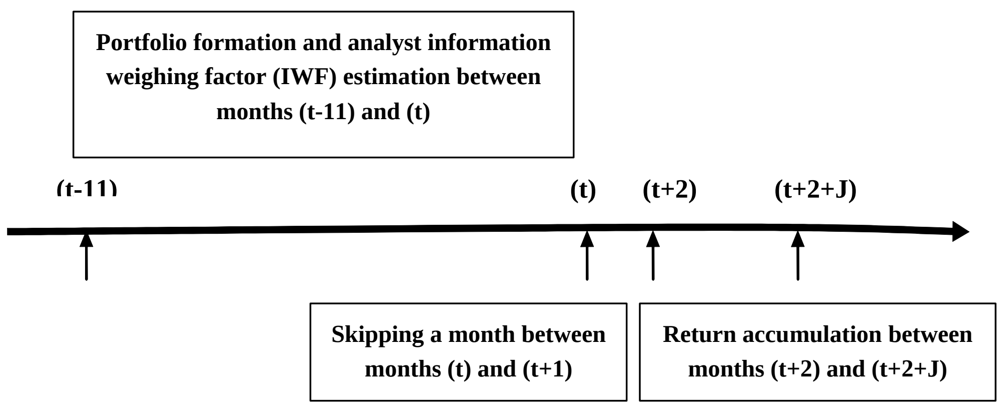


Figure 2.2: Yearly Sample Coverage by CRSP and IBES

This figure plots the percentage of total market capitalization of the stocks covered by the momentum portfolios interacting with analyst herding tendency measure relative to all NYSE, AMEX and NASDAQ common stocks. The annual percentage is averaged across the monthly percentage within the calendar year.

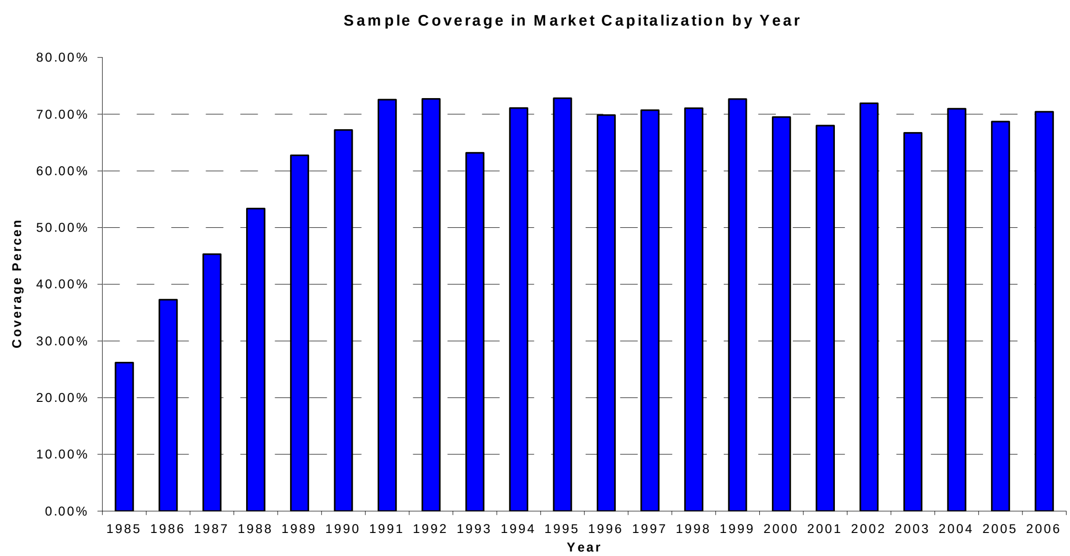


Figure 2.3: Spatial Covariance Estimate and Economics Distance

This figure plots the economic distance constructed as the Euclidian normal of Z-scores of past return, logarithm of market capitalization and book to market equity ratios against the spatial covariance estimate of residuals. The left vertical bar denotes the breakpoints of the economic distance associated with lower 2.5% of total population of pairwise residual correlation, and the right vertical bar denotes the breakpoint of the economic distance associated with upper 2.5% of the total population of pairwise residual correlation.

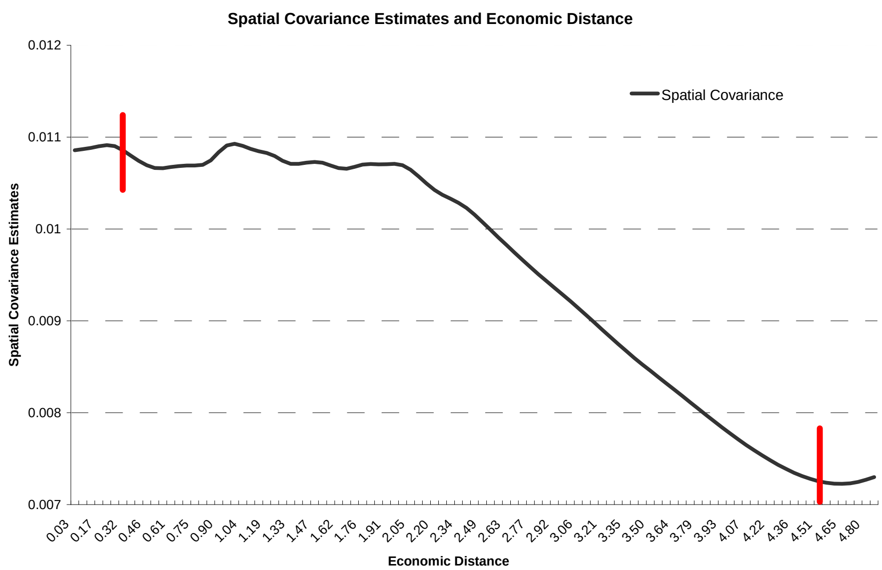


Figure 2.4: Empirical Distribution of Pairwise Correlations

This figure plots the empirical distribution of pairwise correlations of residuals along the economic distance constructed as in figure C1. The left vertical bar denotes the breakpoints of the economic distance associated with lower 2.5% of total population of pairwise residual correlation, and the right vertical bar denotes the breakpoint of the economic distance associated with upper 2.5% of the total population of pairwise residual correlation.

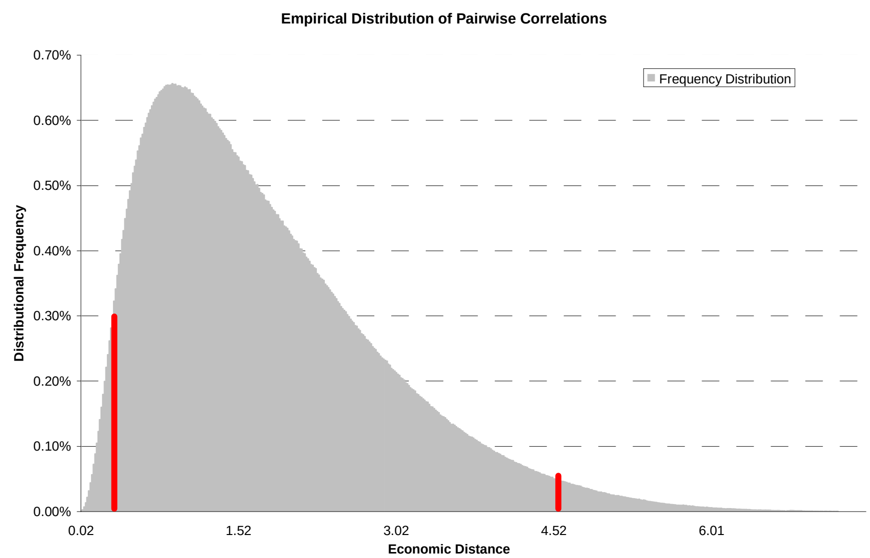
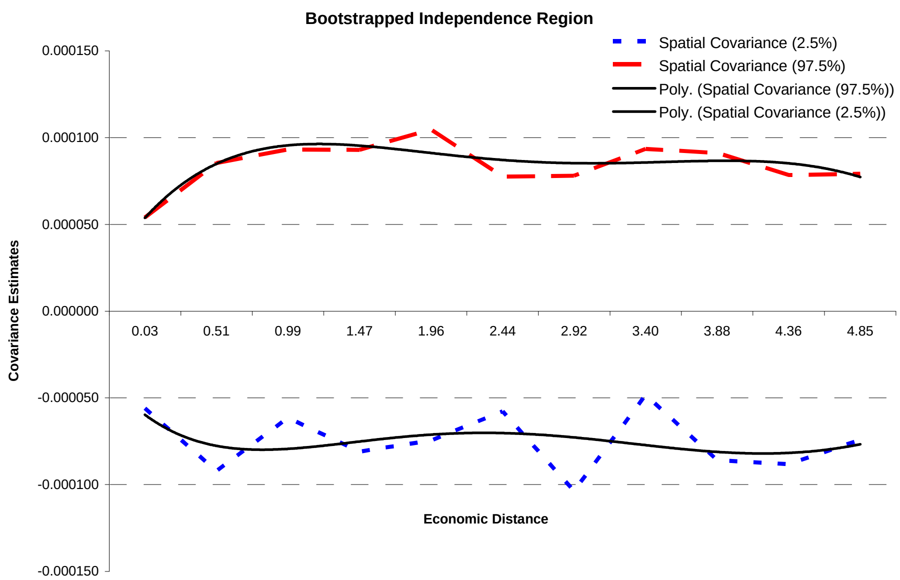


Figure 2.5: Bootstrapped Independence Region

This figure plots the Economic Distance constructed as the Euclidian normal of Z-scores of past return, logarithm of market capitalization and book to market equity ratios against the acceptance region of serial and spatial independence among residuals. The dashed lines denote the 97.5% and 2.5% bounds of the acceptance region, and the solid lines denote the smoothed 97.5% and 2.5% bounds of the acceptance region using polynomials of degree five.



References

[1] Acharya, V. V. and L. H. Pedersen (2005): "Asset pricing with liquidity risk," *Journal of Financial Economics*, Vol 72, 375-410.

[2] Andrade, G. and S. N. Kaplan (1998): "How costly is financial (not economic) distress? Evidence from highly leveraged transactions that became distressed," *Journal of Finance*, Vol. 53, 1443-1498.

[3] Almazan, A. K. C. Brown, M. Carlsonb and D. A. Chapman (2004): "Why constrain your mutual fund manager?" *Journal of Financial Economics*, Vol 73, 289–321.

[4] Altman, E. (1968): "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *Journal of Finance*, Vol 23, 589-609.

[5] Asparouhova, Elena, Michael Hertzel, and Michael Lemmon, 2005,Behavioral Biases and Investor Behavior: Predicting the Next Step of a Random Walk (Revisited and Extended), *Working Paper*, University of Utah.

[6] Atkins, Allen B. and Edward A. Dyl, 2005, Market structure and reported trading volume, *Journal of Financial Research* 28, 115 - 131.

[7] Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

[8] Amihud, Y., and H. Mendelson (1986): "Asset pricing and the bid-ask spread," *Journal of Financial Economics*, Vol 17, 223-250.

[9] Antweiler, Werner, and Murray Z. Frank, 2004, The market impact of corporate news stories, *Working Paper*, Carlson School of Management, University of Minnesota.

[10] Asness, Clifford S., 1997, The interaction of value and momentum strategies, *Financial Analysts Journal* (March/April), 29-36.

[11] Avramov D., T. Chordia and A. Goyal (2005): "Liquidity and autocorrelations in individual stock returns," *Journal of Finance*, forthcoming.

[12] Avramov D., T. Chordia, G. Jostova and A. Philipov (2006): "Credit ratings and the cross-section of stock returns," *Working Paper.*

[13] Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159-178.

[14] Banerjee, A.V., 1992, A Simple-Model of Herd Behavior, *Quarterly Journal of Economics* 107, 797-817.

[15] Barberis, N., A. Shleifer, and R. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.

[16] Barber, Brad, Reuven Lehavy, Maureen McNichols, and Brett Trueman, 2001, Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns, *Journal of Finance* 56, 531-563.

[17] Barberis, Nicholas and Richard Thaler, 2003, A Survey of Behavior Finance, in George Constantinides, Milton Harris and Rene Stulz (eds.), *Handbook of the Economics of Finance*, North Holland.

[18] Barberis, Nicholas and Wei Xiong, 2006, What Drives the Disposition Effect? An Analysis of a Long-standing Preference-based Explanation, Working paper, Yale School of Management, Yale University.

[19] Berndt, Antje, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranzk, 2005, Measuring default risk premia from default swap rates and EDFs, *Working Paper*, Standford University.

[20] Berk, Jonnathan B., Richard C. Green, and V. Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553-1607.

[21] Bernhardt, Dan, Murillo Campello, and Edward Kutsoati, 2006, Who herds?, *Journal of Financial Economics* 80, 657-675.

[22] Bessembinder, H., K. Chan and P. J. Seguin (1996): "An empirical examination of information, differences of opinion, and trading activity," *Journal of Financial Economics*, Vol 40, 105-134.

[23] Bhandari, L.C. (1988): "Debt/equity ratio and expected common stock returns: Empirical evidence," *Journal of Finance*, Vol 43, 507-528.

[24] Biais, Bruno, Denis Hilton, Karine Mazurier, and Sebastien Pouget, 2005, Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market, *Review of Economic Studies* 72, 287-312.

[25] Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A Theory of Fads, Fashion, Custom, and Cultural-Change as Informational Cascades, *Journal of Political Economy* 100, 992-1026.

[26] Black, F. and M. Scholes (1973): "The pricing of options and corporate liabilities," *Journal of Political Economy*, Vol 81, 637-659.

[27] Blacke, Robert R, Harry Helson, and Jane Srybley Mouton, 1957, The Generality of Conformity Behavior as a Function of Factual Anchorage. Difficulty of Task, and Amount of Social Pressure, *Journal of Personality* 25, 294-305.

[28] Bloomfield, Robert, and Jeffrey Hales, 2002, Predicting the next step of a random walk: experimental evidence of regime-shifting beliefs, *Journal of Financial Economics* 65, 397-414.

[29] Blume. M.E. and R. F. Stambaugh (1983): "Biases in computed returns : an application to the size effect," *Journal of Financial Economics*, Vol 12, 387-404.

[30] Boni, Leslie, and Kent L. Womack, 2006, Analysts, Industries, and Price Momentum, *Journal of Financial & Quantitative Analysis* 41, 85-109.

[31] Branch, B. (1977): "A tax loss trading rule," *Journal of Business*, Vol 50, 198-207.

[32] Brennan, Michael J., Tarun Chordia and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345-373.

[33] Bradley, Daniel J., Bradford D. Jordan and Jay R. Ritter, 2006, Analyst Behavior Following IPOs: The "Bubble Period" Evidence, *Review of Financial Studies*, forthcoming.

[34] Brav, Alon, Christopher Geczy, and Paul A. Gompers, 2000, Is the abnormal return following equity issuances anomalous?, *Journal of Financial Economics* 56, 209-249.

[35] Brunnermeier, M.K., Nagel, S., 2004, Hedge funds and the technology bubble. *Journal of Finance* 59, 2013-2040.

[36] Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2006, Robust Inference with Multiway Clustering, working paper, University of California-Davis.

[37] Campbell, J. Y., S. J. Grossman, and J. Wang (1993): "Trading volume and serial correlation in stock returns," *Quarterly Journal of Economics*, 108(4), 905–939.

[38] Campbell J. Y., J. Hilscher and J. Szilagyi (2004): "In search of distress risk," *Working Paper*, Harvard University.

[39] Çelen, Boğaçhan, and Shachar Kariv, 2004, Distinguishing Informational Cascades from Herd Behavior in the Laboratory, *American Economic Review* 94, 484-498.

[40] Chan, K.C. and N. Chen (1991): "Structural and return characteristics of small and large firms," *Journal of Finance*, Vol 46, 1467-1484.

[41] Chan, Wesley S., 2003, Stock price reaction to news and no-news: drift and reversal after headlines, *Journal of Financial Economics* 70, 223-260.

[42] Chan, Louis, Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum Strategies, Journal of Finance 51, 1681-1713.

[43] Chari, V. V., and Patric J. Kehoe, 2004, Financial crises as herds: overturning the critiques, *Journal of Economic Theory* 119, 128-150.

[44] Chen, Qi, and Wei Jiang, 2006, Analysts' Weighting of Private and Public Information, *Review of Financial Studies*, Forthcoming.

[45] Cheng, Yingmei, Mark H. Liu, and Jun Qian, 2006, Buy-Side Analysts, Sell-Side Analysts, and Investment Decisions of Money Managers, *Journal of Financial and Quantitative Analysis* 41, 51-83.

[46] Chen, Zhiwu, Werner Stanzl and Masahiro Watanabe, 2001, Price Impact Costs and the Limit of Arbitrage, *Working Paper*, Yale School of Management.

[47] Chevalier, Judith A., and Glenn Ellison, 1999, Career Concerns of Mutual Fund Managers, *Quarterly Journal of Economics* 114, 389-432.

[48] Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle and time-varying expected returns, *Journal of Finance* 57, 985-1019.

[49] Chordia, T., R. Roll, and A. Subrahmanyam (2000): "Commonality in liquidity," *Journal of Financial Economics*, Vol 56, 3–28.

[50] Chordia, T., R. Roll, and A. Subrahmanyam (2002): "Order imbalance, liquidity, and market returns," *Journal of Financial Economics*, Vol 65, 111–130.

[51] Clement, Michael B., and Senyo Tse, 2003, Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters?, *Accounting Review* 78, 227–249.

[52] Clement, Michael B., and Senyo Y.Tse, 2005, Financial Analyst Characteristics and Herding Behavior in Forecasting, *Journal of Finance* 60, 307-341.

[53] Conley, Timothy G, 1999, GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics* 92, 1-45.

[54] Conley, Timothy G, and Bill Dupor, 2003, A Spatial Analysis of Sector Complementarity, *Journal of Political Economy* 111, 311-352.

[55] Conley, Timothy G, and Xiaohong Chen, 2001, A New Semiparametric Spatial Model for Panel Time Series, *Journal of Econometrics* 105, 59-83.

[56] Cooper Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market States and Momentum, *Journal of Finance* 59, 1345-1365.

[57] Cooper, Rick A., Theodore E. Day, and Craig M. Lewis, 2001, Following the leader: A study of individual analysts' earnings forecasts, *Journal of Financial Economics* 61, 383–416.

[58] Conrad, J. S., A. Hameed and C. Niden (1994): "Volume and autocovariances in short-horizon individual security returns", *Journal of Finance*, Vol 49, 1305-1329.

[59] Conrad, J. S., K. M. Johnson and S. Wahal (2003): "Institutional trading and alternative trading systems", *Journal of Financial Economics*, Vol 70, 99–134.

[60] Coval, J. and E. Stafford (2005): "Asset fire sales (and purchases) in equity markets," *Journal of Finance*, forthcoming.

[61] Connor, Gregory, and Robert A. Korajczyk, 1988, Risk and Return in an Equilibrium APT: Application of a New Test Methodology, Journal of Financial Economics 21, 255-289.

[62] Da, Zhi and Ernst Schaumburg, 2005, Target prices, relative valuations and the compensation for liquidity provision, *Working Paper*, Kellogg School of Management.

[63] D'Avolio, Gene, 2002, The market for borrowing stock, Journal of Financial Economics 66, 271-306.

[64] Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839-1885.

[65] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic Based Benchmarks, *Journal of Finance* 52, 1035-58.

[66] Daniel, Kent, and Sheridan Titman, 2000, Market Efficiency in an Irrational World, *NBER Working Paper 7489*.

[67] De Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-805.

[68] De Bondt, Werner F. M., and Richard Thaler, 1987, Further Evidence on Investor Overreaction and Stock Market Seasonality, *Journal of Finance* 42, 557-581.

[69] De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-738.

[70] De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann, 1991, The survival of noise traders in financial markets, *Journal of Business* 64, 1 - 20.

[71] DellaVigna, Stefano, and Joshua Pollet, 2004, Strategic Release of Information on Fridays: Evidence from Earnings Announcements, *Working Paper*, Department of Economics, University of California, Berkeley.

[72] Dichev, I. (1998): "Is the risk of bankruptcy a systematic risk?," *Journal of Finance*, Vol 53, 1131-47.

[73] Diether, K. B., C. J. Malloy, and A. Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

[74] Duarte, J., F. A. Longstaff and F. Yu (2005): "Risk and return in fixed income arbitrage: nickels in front of a steamroller," *Review of Financial Studies*, forthcoming.

[75] Durham, Gregory R., Michael G. Hertzel, and J. Spencer Martin, 2005, The Market Impact of Trends and Sequences in Performance: New Evidence, *Journal of Finance* 60, 2551-2569.

[76] Durnev, A., R. Morck, B. Yeung, and P. Zarowin (2003): "Does greater firm-specific return variation mean more or less informed stock pricing?" *Journal of Accounting Research*, Vol 41, 797-836.

[77] Ehrbeck, Tilman, and Robert Waldmann, 1996, Why are Professional Forecasters Biased? Agency Versus Behavioral Explanations, T*he Quarterly Journal of Economics* 111, 21-40.

[78] Ellis, K., R. Michaely and M. O'Hara (2000): "When the underwriter is the market maker: An examination of trading in the IPO aftermarket," *Journal of Finance*, Vol 55, 1039-1074.

[79] Engleberg, Joey and Pengjie Gao, 2006, Soure of Profitability from Pair-trades, Working Paper, Kellogg School of Management, Northwestern University.

[80] Fama, Eugene F., Kenneth R. French, 1993, The cross-section of expected stock returns," *Journal of Finance*, Vol 47, 427-465.

[81] Fama, Eugene F., Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

[82] Fama, Eugene F., Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55-84.

[83] Fama, Eugene F., Kenneth R. French, 2001, Disappearing dividends: changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics* 60, 3-43.

[84] Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *The Journal of Political Economy* 81, 607-636.

[85] Fan, Jianqiang, and Irene I. Gijbels, 1996, Local Polynomial Modelling and Applications, Chapman and Hall, London.

[86] Frankel, Richard, and Sugata Roychowdhury, 2004, Testing the clientele effect: an explaination for non-GAAP earnings adjustments used to compute I/B/E/S earnings, *Working Paper*, Sloan School of Management, Massachusetts Institute of Technology.

[87] Frazzini, Andrea, 2006, The Disposition Effect and Under-Reaction to News, *Journal of Finance* 61, 2017-2046.

[88] Froot, Kenneth A., Scharfstein, David S., Stein, Jeremy C., 1992, Herd on the street: informational inefficiencies in a market with short-term speculation, *Journal of Finance* 47, 1461-1484.

[89] Gabaix, Xavier, Arvind Krishnamurthy and Olivier Vigneron, 2006, Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market, *Journal of Finance*, Forthcoming.

[90] Garlappi, L., T. Shu and H. Yan (2005): "Default risk, shareholder advantage, and stock returns," *Review of Financial Studies*, forthcoming.

[91] Gao, George, 2006, A Spatial Analysis of Stock Comovements: Theory and Evidence, working paper, Graduate School of Business, University of Chicago.

[92] E. Gatev, W. N. Goetzmann and K. G. Rouwenhorst (2006): "Pairs trading: performance of a relative-value arbitrage rule," *Review of Financial Studies*, Vol 19, 797-827.

[93] George, Thomas J., and Chan-Yang Hwang, 2004, The 52-Week High and Momentum Investing, *Journal of Finance* 59, 2145-2176.

[94] Gilson, S.C. (1995): "Investing in distressed situations: a market survey," *Financial Analyst Journal*, Nov/Dec, 8-27.

[95] Gleason, Cristi A., and Charles M. C. Lee, 2003, Analyst Forecast Revisions and Market Price Discovery, *Accounting Review* 78, 193-225.

[96] Gomes, Armando, Gary Gorton, and Leonardo Madureira, 2004, SEC Regulation Fair Disclosure, Information, and the Cost of Capital, *Working Paper 10567*, NBER.

[97] Graham, John R., 1999, Herding among investment newsletters: theory and evidence, *Journal of Finance* 54, 237-268.

[98] Grundy, B. D., and J. S. Martin, 2001, Understanding the nature of the risks and the source of the rewards to momentum investing, *Review of Financial Studies* 14, 29-78.

[99] Green, T. Clifton, 2006, The Value of Client Access to Analyst Recommendations, *Journal of Financial and Quantitative Analysis* 41, 1-24.

[100] Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311-339.

[101] Grinblatt, Mark, and Tobias J. Moskowitz, 2004, Predicting stock price movements from past returns: the role of consistency and tax-loss selling, *Journal of Financial Economics* 71, 541-579.

[102] Griffin, John M., Susan Ji, and J. Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515-2547.

[103] Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior, *American Economic Review* 85, 1088 - 1105.

[104] Grossman S. and M. H. Miller (1988): "Liquidity and market structure," *Journal of Finance*, Vol 43, 617-633.

[105] Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy*, Vol. 99, 225-262.

[106] Harris, J. H., V. Panchapagesan and I. Werner (2004): "From pink slips to pink sheets: Liquidity and shareholder wealth consequences of Nasdaq delistings," *Working Paper*, Dice Center.

[107] Hasbrouck, J. and D. J. Seppi (2001): "Common factors in prices, order flows, and liquidity," *Journal of Financial Economics*, Vol 59, 383–411.

[108] Haugen, R. and J. Lakonishok (1988): "The incredible January effect: The stock market's unsolved mystery," Dow Jones-Irwin, Homewood, Illinois.

[109] Hendershott, T. and M. S. Seasholes (2006): "Market maker inventories and stock prices," *Working Paper*.

[110] Hillegeist, S.A., Keating, E.K., Cram D.P. and K.G. Lundstedt (2003): "Assessing the probability of bankruptcy," *Review of Accounting Studies*, Vol 9, 5-34.

[111] Hirshleifer, David, and Siew Hong Teoh, Herding and Cascading in Capital Markets: A Review and Synthesis, *European Financial Management* 9, 25-66.

[112] Hong, Harrison, and Jeffrey D. Kubik, 2003, Analyzing the analysts: Career concerns and biased forecasts, *Journal of Finance* 58, 313–351.

[113] Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, 2000, Security analysts' career concerns and herding of earnings forecasts, *RAND Journal of Economics* 31, 121–144.

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

[115] Hong, Harrison, Terrance Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265-295.

[116] Hong, Harrison, Jose Scheinkman and Wei Xiong, 2005, Advisors and Asset Prices: A Model of the Origins of Bubbles, Working Paper, Princeton University.

[117] Ivkovic, Zoran, and Narasimhan Jegadeesh, 2004, The timing and value of forecast and recommendation revisions, *Journal of Financial Economics* 73, 433-463.

[118] Jegadeesh, N. (1990): "Evidence of predictable behavior of security returns," *Journal of Finance*, Vol 45, 881-898.

[119] Jagannathan, Ravi, Tongshu Ma, and Antonio Baldaque da Silva, 2005, Analyst earnings forecast and equity return anomalies, *Working Paper*, Kellogg School of Management, Northwestern University.

[120] Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.

[121] Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D. Krische, Charles M. C. Lee, Analyzing the Analysts: When Do Recommendations Add Value?, *Journal of Finance* 59, 1083 - 1124.

[122] Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

[123] Jegadeesh, Narasimhan, and Sheridan Titman, 2002, Cross-sectional and time-series determinants of momentum returns, *Review of Financial Studies* 15, 143-157.

[124] Jiang, Guohua, Charles M.C. Lee, and Yi Zhang, 2005, Information Uncertainty and Expected Returns, *Review of Accounting Studies* 10, 185-221.

[125] Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585-608.

[126] Kahneman, Daniel, Paul Slovic, and Amos Tversky, 1982, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge University Press, New York).

[127] Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263-291.

[128] Keim, D.B. and A. Madhavan (1995): "Anatomy of the trading process: empirical evidence on the behavior of institutional traders," *Journal of Financial Economics*, Vol 37, 371-398.

[129] Keim, D.B. and A. Madhavan (1996): "The upstairs market for large-block transactions: Analysis and measurement of price effects," *Review of Financial Studies*, Vol 9, 1-36.

[130] Kirchler, Erich, and Boris Maciejovsky, 2002, Simultaneous Over- and Underconfidence: Evidence from Experimental Asset Markets, Journal of Risk and Uncertainty 25, 65-85.

[131] Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039-1082.

[132] Kyle, A. S. (1985): "Continuous auctions and insider trading," *Econometrica*, Vol 53, 1315-1336.

[133] Lamont, Owen A., 2002, Macroeconomic forecasts and microeconomic forecasters, J*ournal of Economic Behavior & Organization* 48, 265-280.

[134] Lee, C. M. C., and B. Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017-2069.

[135] Lehman, B. (1990): "Fads, martingales and market efficiency," *Quarterly Journal of Economics*, Vol 105, 1–28.

[136] Lettau, Martin and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, *Journal of Political Economy* 109, 1238-1287.

[137] Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289-314.

[138] Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence Phillips, 1982, Calibration of Probabilities: The State of the Art to 1980, in *Judgment under Uncertainty: Heuristics and Biases*, edited by Daniel Kahneman, Paul Slovic, and Amos Tversky, Cambridge, U.K., and New York: Cambridge University Press, 306-334.

[139] Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The Illusory nature of momentum profits, *Journal of Financial Economics* 71, 349-380.

[140] Lim, Seongyeon, 2003, Essays in Financial Economics: Mental Accounting and Selling Decisions of Individual Investor; Analysts' Reputational Concerns and Underreaction to Public News, *Unpublished Ph.D. Dissertation*, Ohio State University.

[141] Lim, Terrence, 2001, Rationality and Analysts' Forecast Bias, *Journal of Finance* 56, 369-385.

[142] Lin, Hsiou-wei, and Maureen F. McNichols, 1998, Underwriting relationships, analysts' earnings forecasts and investment recommendations, *Journal of Accounting and Economics* 25, 101-127.

[143] Liu, Laura, Jerry Warner, and Lu Zhang, 2004, Momentum profits and macroeconomic risk, *Working Paper*, Simon School of Business, University of Rochester.

[144] Loughran, Tim, and Jay R. Ritter, 1995, The New Issues Puzzle, *Journal of Finance* 50, 23-51.

[145] MacKinlay, A.C., 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

[146] Macey, J., M. O'Hara and D. Pompilio (2004): "Down and out in the stock market: The law and finance of the delisting process," *Working Paper.*.

[147] Malmendier, Ulrike M. and Devin M. Shanthikumar, 2005, Are Investors Naive About Incentives?, *Working Paper*, Harvard Business School.

[148] Merton, R.C. (1974): "On the pricing of corporate debt: The risk structure of interest rates," *Journal of Finance*, Vol 29, 449-470.

[149] Merton, R.C. (1987): "A simple model of capital market equilibrium with incomplete information," *Journal of Finance*, Vol 42, 483-510

[150] Michaely, Roni, and Kent L. Womack, 1999, Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12, 653-686.

[151] Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum?, *Journal of Finance* 54, 1249-1290.

[152] Newey, Whitney K, and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

[153] Nofsinger, John, Sias, Richard, 1999, Herding and feedback trading by institutional and individual investors, *Journal of Finance* 54, 2263-2295.

[154] Nieuwerburgh, Stijn Van, and Laura Veldkamp, 2006, Information Acquisition and Portfolio Under-Diversification, Working Paper, Stern Business School, New York University.

[155] Odders-White E.R. and M.J. Ready (2005): "Credit ratings and stock liquidity," *Review of Financial Studies*, Vol 19, 119-157.

[156] Odean, Terrance, 1998, Volume, volatility, price and profit when all trades are above average, *Journal of Finance*, 53, 1887-1934.

[157] Ohlson, J. A. (1980): "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, Vol 18, 109-131.

[158] Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

[159] Payne, Jeff L. and Thomas, Wayne B., The Implications of Using Stock-Split Adjusted I/B/E/S Data in Empirical Research, The Accounting Review 78, 1049-1067.

[160] Prendergast, Canice, and Lars Stole, 1996, Impetuous youngsters and jaded oldtimers: Acquiring a reputation for learning, *Journal of Political Economy* 104, 1105–1134.

[161] Petersen, Mitchell A., 2006, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, Working Paper, Kellogg School of Management, Northwestern University.

[162] Polk, Christopher, and Paola Sapienza, 2006, The stock market and corporate investment: a test of catering theory, *Review of Financial Studies*, Forthcoming.

[163] Pulvino, Todd, Do Asset Fire Sales Exist? An Empirical Investigation of Commercial Aircraft Transactions, J*ournal of Finance* 53, 939-978.

[164] Rogers, William, 1993, Regression Standard Errors in Clustered Samples, *Stata Technical Bulletin* 13, 19-23.

[165] Rouwenhorst, K. Geert, 1998, International momentum strategies, *Journal of Finance* 53, 267– 284.

[166] Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80, 309-349.

[167] Sadka, R. and A. Scherbina (2006): "Analyst disagreement, mispricing and liquidity," *Journal of Finance*, forthcoming.

[168] Santos, Tano, Veronesi, Pietro, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

[169] Scharfstein, David, and Jeremy Stein, 1990, Herd behavior and investment, *American Economic Review* 80, 465–479.

[170] Shefrin, Hersh, and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance* 40, 777-791.

[171] Seguin P. J. and M. M. Smoller (1997): "Share price and mortality: An empirical evaluation of newly listed Nasdaq stocks," *Journal of Financial Economics*, volume 45, 333-363.

[172] Shiller, Robert J., 1984, Stock prices and social dynamics, *Brookings Papers on Economic Activity* 2, 457-498.

[173] Shiller, Robert J., 1995, Conversation, Information, and Herd Behavior, *American Economic Review* 85, 181–85.

[174] Shiller, Robert J., and John Pound, 1989, Survey Evidence on Diffusion of Interest and Information Among Investors, *Journal of Economic Behavior and Organization* 12, 47-66.

[175] Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

[176] Shumway, T. (2001): "Forecasting bankruptcy more accurately: A simple hazard model," *Journal of Business*, Vol 74, 101-124.

[177] Shumway, T. and V. A. Warther (1999): "The delisting bias in CRSP's Nasdaq data and its implications for the size effect," *Journal of Finance*, Vol 54, 2361-2379.

[178] Silva, Antonio Baldaque da, 2003, Forecasting Earnings on Industry Stock Index Portfolios, Unpublished Ph.D. Dissertation, Northwestern University.

[179] Skoulakis, Georgios, 2006, Panel Data inference in Finance: Least-Squares vs Fama-MacBeth, working paper, University of Maryland.

[180] Spiegel, M. and A. Subrahmanyam (1995): "On intraday risk premia," *Journal of Finance*, Vol 50, 319-339.

[181] Spiegel, M. and X. Wang (2005): "Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk," *Working Paper*, Yale School of Management.

[182] Temin, Peter and Voth, Hans-Joachim, 2004, Riding the South Sea Bubble, *American Economic Review* 94, 1654 - 1668.

[183] Thaler, Richard, 1985, Mental accounting and consumer choice, *Marketing Science* 4, 199 - 214.

[184] Thompson, Samuel, 2006, Simple Formulas for Standard Errors that Cluster by Both Firm and Time, Working Paper, Harvard University.

[185] Trueman, Brett, 1994, Analyst Forecasts and Herding Behavior, *Review of Financial Studies* 7, 97-124.

[186] Vega, Clara, 2006, Stock price reaction to public and private information, *Journal of Financial Economics* 82, 103-133.

[187] Vassalou, M. and Y. Xing (2004): "Default risk in equity returns," *Journal of Finance*, Vol 59, 831-868.

[188] Vuolteenaho, T. (2002): "What drives firm-level stock returns," *Journal of Finance*, Vol 57, 233-264.

[189] Welch, Ivo, 2000, Herding among security analysts, *Journal of Financial Economics* 58, 369-396

[190] Wermers, Russ, 1999, Mutual Fund Herding and the Impact on Stock Prices, *Journal of Finance* 53, 581 - 622.

[191] White, Hal, 1984, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity, *Econometrica* 48, 817-838.

[192] Wilder, David A., 1977, Perception of groups, size of opposition and social influence, *Journal of Experimental and Social Psychology* 13, 253-268.

[193] Wurgler, J. and K. Zhuravskaya, 2002, Does arbitrage flatten demand curves for stocks?, *Journal of Business* 75, 583-608.

[194] Zhang, Frank X., 2006, Information Uncertaity and Stock Returns, *Journal of Finance* 61, 105-137.

[195] Zitzewitz, E., 2001, Measuring herding and exaggeration by equity analysts and other opinion sellers, *Working Paper*, Graduate School of Business, Standford University.

APPENDIX A

Variance Decomposition of the Default Likelihood Indicator (DLI)

Empirically, the Default Likelihood Indicator is computed as a function of three variables: leverage (lev), past stock return (ret) and asset volatility (σ_A), i.e., $DLI = N(-DD) = f(lev, ret, \sigma_A)$. We want to examine the relative importance of these three variables in a variance decomposition framework. Theoretically, as normal CDF is a monotone transformation of its argument ($-DD$), we can either work with the transformed variable (DLI) or the original variable($-DD$). Unfortunately, directly working with DLI is challenging because DLI is highly skewed due to the nonlinear transformation of normal CDF. Therefore, we decide to study the variance decomposition of the equivalent variable: $-DD$, which is better-behaved statistically.

Applying the first-order Taylor series expansion of $-DD$ around the cross-sectional mean of lev, ret and σ_A, we have[1]:

(A.1)
$$-\overline{DD} = \frac{\partial f}{\partial lev}\overline{lev} + \frac{\partial f}{\partial ret}\overline{ret} + \frac{\partial f}{\partial \sigma}\overline{\sigma_A} + \kappa,$$

[1]For simiplicity of notation, we omit the time subscript t and firm superscript i.

where κ captures the approximation error, and variables with upper bar are cross-sectionally demeaned. Therefore, we have

$$\text{(A.2)}\quad var(DD) = \frac{\partial f}{\partial lev} cov(DD, lev) + \frac{\partial f}{\partial ret} cov(DD, ret) + \frac{\partial f}{\partial \sigma} cov(DD, \sigma_A) + cov(DD, \kappa),$$

where $var(\cdot)$ and $cov(\cdot)$ are the cross-sectional variance and covariance, respectively. Dividing both sides of the above equation by $var(DD)$, we then have

$$\text{(A.3)}\qquad 1 = \frac{\partial f}{\partial lev}\beta_{lev} + \frac{\partial f}{\partial ret}\beta_{ret} + \frac{\partial f}{\partial \sigma}\beta_{\sigma_A} + \beta_{\kappa}.$$

The term $\frac{\partial f}{\partial(\cdot)}\beta_{(\cdot)}$ then measures the contribution of each input to the cross-sectional variations of DLI. The sum of the contribution from the three factors is less than one, and the difference, as captured by β_{κ}, is due to the approximation error in the Taylor series expansion. The partial derivatives, or sensitivity, $\frac{\partial f}{\partial(\cdot)}$ are computed numerically by the finite difference method. β can be measured by regression. For instance, β_{lev} is estimated by regressing $\overline{lev}$ on $-\overline{DD}$ cross-sectionally (so the intercept of the regression is zero by construction). Empirically, we have a panel data of $-DD$, lev, ret and σ_A. To estimate β, we follow Vuolteenaho (2002) and run a Weighted Least Squares (WLS) regression. In practice, this means deflating the data for each firm-date by the number of firms in the corresponding cross-section. The results are reported in Table I. We report only simple WLS standard error. The simple WLS standard errors translate to t-values above fifty for all estimates; therefore, we are confident that all the estimates will still be significant, even if we adjust for auto-correlation and cross-sectional correlation of the error terms. Of course, this is hardly surprising, as (A.3) is merely a statement of an identity.

APPENDIX B

Additional Discussions of DLI Sorted Portfolio Returns

B.1. Are results driven by bid-ask spreads?

One particular problem associated with illiquid stocks traded at low prices is that the bid-ask bounce could lead to a non-negligible upward bias in average return computation, as discussed in Blume and Stambaugh (1983) and most recently in Canina, Michaely, Thaler and Womack (1998). A natural question is whether the first-month high return on the highest-DLI stock portfolio is entirely driven by the bias due to the bid-ask bounce. We believe that the answer is no.

We approach this question by estimating the impact of the bid-ask bouce on return. Blume and Stambaugh (1983) show that the bias on return per period due to the bid-ask bounce can be measured by $\left(\frac{P_A - P_B}{P_A + P_B}\right)^2$, where P_A and P_B are bid and ask price, respectively. First, assuming a bid-ask spread of \$0.25 and given an average price of \$3.58 for stocks in the highest-DLI stock portfolio, a rough estimate for the bias is 12 bps $= \left(\frac{0.25}{3.58 \times 2}\right)^2$, which is much smaller than the 90 bp return premium the highest-DLI stocks earn over the stocks in the next highest-DLI decile. Second, we also compute the bias measure for individual stock and average the bias measures first within each DLI decile and then across time. We report the average monthly return bias due to bid-ask bounce for each DLI decile in Table B1. For this calculation, we again assume a bid-ask spread of \$0.25, which is typically higher than the actual bid-ask spread especially for penny stocks.[1] Therefore, the return bias measures we compute are most likely overestimates and can thus serve as an upper bound for the true return bias. For example, Blume and Stambaugh (1993) choose a single day at random - Dec 13, 1973 - and select all NYSE common stocks with bid prices less than \$8. The average bias measure for these 332 stocks is only 5 bps. As expected, the bias increases with DLI. The bias is 54 bps for the highest-DLI decile, which is higher than the rough estimate of 12 bps we calculated earlier using the average price. This is because our assumption on the bid-ask spread generates extremely large bias measure on penny stocks which overstates the average. However, the difference in the average bias measure between stocks in DLI-decile 10 and stocks in DLI-decile 9 is only 26 bps, again much smaller than their return difference of 90 bps. As a robustness check, we also compute an alternative return bias measure in a subsample from 1983 to 1999, using the actual quoted spread (quoted ask – quoted bid) from quote data in TAQ and ISSM.[2] As trade could happen between the quoted bid and quoted ask, the alternative return bias measure is again likely to be overstated.[3] The alternative return bias measure is uniformly smaller than the first return bias measure. Again the difference in the average bias measure between stocks in DLI-decile 10 and stocks in DLI-decile 9 (25 bps) is much smaller than their return difference of 90 bps.

[1] For a stock with a price less than or equal to \$0.25, the assumption of a \$0.25 bid-ask spread does not make much sense. We therefore assume a bid-ask spread equal to 50% of the trading price for such a stock.

[2] The sampling period for NYSE stocks is from 1983 to 1999 and the sampling period for NASDAQ stocks is from 1987 to 1999. Because the ISSM data were constructed in early years through data collection from various sources, not all transaction records are in the database. In particular, six months worth of data for NASD stocks from 1987 through 1989 in ISSM are missing.

[3] For example, floor traders at NYSE can cross the trades by taking the opposite side of the incoming order and execute at the better of bid or ask quotes. It is also possible that large blocks can be executed on the up-stair market.

As a more direct way of accoutning for the bid-ask bouce, we compute the monthly return using daily returns from the second positive trading-volume-day. This resulting return measure is therefore largely free from the bid-ask-bounce bias.[4] After excluding the return on the first trading day of the calendar month, the return drops only slightly. For example, the first-month return of the highest-DLI stocks drops to 2.01% from 2.10%, indicating that the impact of bid-ask-bounce is small. To conclude, all the above evidence seems to suggest that random bounce between bid and ask does not fully explain the first-month high return on the highest-DLI stock portfolio.[5]

B.2. Are results driven by increased uncertainty?

A sharp increase in a stock's DLI measure is usually associated with higher uncertainty regarding the firm's "fundamentals" at least temporarily. The increased uncertainty could lead to a higher expected stock return in the near future as in Merton (1987). Later on, as uncertainty resolves, the expected return goes back to its normal level. If such uncertainty-based explanation is true, we would expect stocks with higher level of uncertainty to have higher returns during the first month after portfolio formation. Empirically, we focus on the group of New High DLI stocks since they drive most of the results in the paper. We use a cash-flow based uncertainty measure developed by Zhang (2006). At the end of each month, we further sort New High DLI stocks into two portfolios according to the uncertainty measures and compute the equally-weighted portfolio return during the first

[4]We thank Nai-fu Chen for suggesting this return measure.

[5]It is possible that for stocks in the highest-DLI decile, their prices bounce systematically from bid at the end of portfolio formation month to ask at the end of the first month after. This systematic bid-ask bounce will lead to a much larger first-month return on these stocks. However, such systematic bid-ask bounce is entirely consistent with our explanation. The fact that trade occurs at the bid during portfolio formation indicates large selling pressure after the stock becomes financially distressed. As more buyers come to the market in the next month, trade occurs at the ask.

month after portfolio formation for each portfolio separately. The first-month returns on these two portfolios turn out to be similar: 3.17% for New High DLI stocks with high level of uncertainty measures and 3.35% for New High DLI stocks with low level of uncertainty measures. The difference of 18 bps is not significant (t-value = 0.5). We therefore believe that increased uncertainty is unlikely to be the main explanation of the first-month high return on the highest-DLI stock portfolio.

B.3. Economic significance of the first-month high returns

In this paper, we focus on stocks with high DLIs. These stocks earn about 90 basis points more than otherwise similar stocks during the first month after portfolio formation. These stocks with large exposure to default risk, are more likely to have smaller market capitalizations, lower trading prices and higher percentage trading costs, as shown in Panel A of Table VII.[6] Naturally, a question arises, is the first-month high return on these stocks economically significant? In other words, can such high return be captured by portfolio trading strategies after accounting for transaction costs? This subsection answers this question in detail.

We further sort these stocks into quartiles according to their market capitalizations. We then compute the average monthly returns for each quartile. We also compute the average percentage bid-ask spread and the average return bias due to bid-ask bounce for each quartile. Again, both measures are computed using the actual quoted spread (quoted ask – quoted bid) from quote data in TAQ and ISSM. The sampling periods for these two measures are from 1983 to 1999 for NYSE stocks and from 1987 to 1999 for NASDAQ

[6]During the sampling period from 1971 to 1999, there are on average 260 stocks in the highest DLI-decile per month, with a total market capitalization slightly above 10 billion dollars (from 3 billion dollars at in 1971 to 30 billion dollars in 1999).

stocks. This quoted spread is likely to over-estimate the true “effective” bid-ask spread. The results are presented in Panel A of Table B2. First, for all four quartiles, the first-month returns after portfolio formation are much higher than the return bias measures. Therefore, random bid-ask bounce do not completely explain the high first-month returns. It is more likely that trading price, on average, systematically bounces from bid at portfolio formation to ask a month later, which is consistent with our liquidity-based explanation. Second, the first-month high returns are primarily driven by penny stocks in the lowest-size quartile. These stocks have an average market cap of 2 million dollars, an average trading price of only $1.27 and an average first-month return of 5.76%. This relatively high return is expected. Given its low price, the same bounce from bid to ask will result in a higher return. Finally, For all four quartiles, the average transaction costs as measured by the percentage bid-ask spreads are much higher than the first-month returns, which means that the first-month high return on high-DLI stocks is, on average, economically insignificant. Our liquidity-based explanation would predict a more pronounced price reversal for the subsample of high-DLI stocks that have recently experienced increases in DLIs. When we examine the New DLI stocks, which enter the highest-DLI decile only during the portfolio formation month, we observe larger (in absolute term) negative returns during the portfolio formation month and higher positive returns in the month after. However, these high returns are still not economically significant since they are on average smaller than the transaction costs. Similar results are obtained when we sort high-DLI stocks into quartiles according to their trading prices at portfolio formation as in Panel B of Table B2. In conclusion, outside investors (other than the market makers) cannot consistently capture the first-month high returns on high-DLI stocks by trading at monthly frequency. This is consistent with the findings in Avramov, Chordia and Goyal (2005) in which they show the profits to contrarian trading strategy are smaller than the likely transaction costs and therefore short-term return reversal does not constitute a violation of efficient market hypothesis. Finally, this is also consistent with the view that market makers, generically defined, are compensated by providing liquidity when it is most needed.

B.4. Compensation for liquidity provision during the later sub-sample

Liquidity of the stock market improves significantly since July, 1997 due to various institutional changes on the exchanges. The new Order Handling Rules (OHR) was phrased-in during early 1997 for all NASDAQ stocks, which allows the general public to compete more effectively with NASDAQ market makers in liquidity provision via limit orders. In addition, tick size was cut down from \$1/8 to \$1/16 for both NYSE and NASDAQ stocks on June of 1997. If part of the higher return on High-DLI stocks is indeed a compensation for liquidity provision, we would expect it to decrease after June 1997.[7] The sub-sample result during the later period from July 1997 to the end of 1999 confirms this observation. The Fama-French three-factor risk-adjusted return on the highest-DLI stock portfolios is only 6 bps on average during this period and is not statistically significant.

B.5. Market Maker's Inventory

Our liquidity provision explanation would predict a temporary increase in market maker's inventory when a stock recently becomes finally distressed. Due to data limitation, we cannot directly test the market maker's inventory changes. An indirect (albeit

[7]We thank Joel Hasbrock and Larry Glosten for pointing this out.

imperfect) measure is the stock-level aggregate order imbalances, which capture market making and inventory by traders other than the specialists as well as the specialists, and should be related to the specialists' end-of-day inventory position (at least for NYSE traded stocks). Results from the order imbalance diagnostics (see Panel C of Table VII) suggest that market maker on average take large long positions in the new high DLI stocks, and provide liquidity to the markets. This interpretation is reinforced by findings in Hendershott and Seasholes (2006). Using actual NYSE specialist data, they show the NYSE specialists' inventory positions are negatively correlated with past returns, and large increase in inventory is negatively related to future returns.

APPENDIX C

Data Construction and Definitions

This appendix provides details on the construction of firm characteristic variables.

ME – The market equity of the stock at the month of portfolio formation. The market equity is computed from the end of portfolio formation month share price and number of shares outstanding.

ME Rank – The market equity percentile ranking at the portfolio formation month. As the ME in CRSP sample grows over time, and exhibits large time series variations, I use the NYSE ME percentile breakpoints as a benchmark, and compute the decile ranking of each stock.

DISP – the average analyst quarterly forecast dispersions at the most recent quarter as of month t. Following Deither, Malloy, and Scherbina (2002), I use the unadjusted I/B/E/S detailed summary file and actual earnings file to compute this value.

Size – the logarithm of market capitalization of the firm (in 1000's dollars).

IWF – It is the (normalized predicted) information weighting factor values.

Past Ret – It is the logarithm of the past 12-month cumulative return including portfolio formation month (t).

Past LT Ret – It is the logarithm of the cumulative return during (t-35) and (t-12) months, where t is the portfolio formation month.

Size Adjusted News Coverage – It is the 12-month average news coverage of individual stocks during (t) and (t-11). The news media coverage for stock j at month t is defined as log (1+ number of news of the stock j) / market capitalization (in millions of dollars) of the stock j.

SUE – the latest standardized unexpected earnings in percentile ranking up to the portfolio formation quarter. The standardized unexpected earnings are computed using the seasonal adjusted random walk model, or $SUE_{i,q} = (E_{i,q} - E_{i,q-4}) / \sigma_{i,q}$, where $E_{i,q}$ is the most recent quarterly earning announced as of month t for stock i, $E_{i,q-4}$ is earnings for the prior four quarters, and $\sigma_{i,q}$ is the standard deviation of $(E_{i,q} - E_{i,q-4})$ over the preceding eight quarters.

Turnover - the share turnover ratio. It is computed as the monthly share trading volume (as reported by CRSP MSF) divided by the share outstanding and then average across month (t-11) and month (t). In some cases, I use the rule-of-thumb adjustment procedure suggested by (Atkins and Dyl, 1997; Atkins and Dyl, 2005) and divide NASDAQ stocks' trading volumes reported in CRSP by two.

SEO - an indicator variable taking value of one if there is a seasoned equity offering (SEO) by the firm during any month between $(t - 35)$ and (t).

IPO - an indicator variable taking value of one if there is an initial public offering (IPO) by the firm during any month between $(t - 35)$ and (t).

VOLA - the total volatility of stock returns during the past 12 months. The VOLA measure is scaled by 1000 to improve expositions.

ARBRISK – the arbitrage risk measure (**ARBRISK**) depicts the nondiversifiable risk of individual stock return, as measured by the return volatility not attributable to the market index. Wurgler and Zhuravskaya (2002), and Mendenhall (2004) argue the idiosyncratic volatilities mean higher nondiversifiable risk held by arbitragers. Each month, for every stocks with more than thirty valid observations of past 12 months, we estimate the following regression using daily returns and factors of past 12 months,

$$\text{(C.1)} \quad \begin{aligned} R_{i,t} - R_{RF,t} &= \alpha + \beta_{i,0} MKTRF_t + \beta_{i,1} MKTRF_{t-1} \\ &\quad + \beta_{i,2} \left[\frac{MKTRF_{t-2} + MKTRF_{t-3} + MKTRF_{t-4}}{3} \right] + \varepsilon_{i,t} \end{aligned}$$

where $MKTRF$ is the market excess return factor, $R_{RF,t}$ is the risk free rate, and $R_{i,t}$ is the daily stock return. To control for the nonsynchronous trading, we use the sum-beta method in Dimson (1979). The ARBRISK measure is scaled by 1000 to improve expositions.

ILIQ – the illiquidity measure suggested by Amihud (2002):

$$\text{(C.2)} \qquad Amihud_t = \frac{1}{T} \sum\nolimits_{d=1}^{T} \frac{|R_{i,t-d}|}{Vol_{i,t-d}} \ .$$

I average the daily absolute value of the ratio between return and dollar trading volume of individual stocks during the portfolio formation month t to get the Amihud measure for month t – $Amihud_t$. I use the filtering rules suggested by Amihud (2002), but I do not exclude NASDAQ stocks. Because the NYSE/AMEX and NASDAQ volumes are not directly comparable due to inter-dealer volume double count, I use the rule-of-thumb adjustment procedure suggested by Atkins and Dyl (2005) and divide NASDAQ stocks' trading volumes reported in CRSP by two. To assess the robustness of our empirical measures, I also consider a few variants of the above construction: (1) I experiment exclusion of the top and bottom 1 percent of the annual observations to mitigate the influence of outliers. (2) I replace the missing value of the daily liquidity measure with concurrent year minimum, mean, median and maximum illiquidity measures. All of the results are quantitatively and qualitatively similar. To improve exposition, the original Amihud measure is scaled by $10,000$.

APPENDIX D

Calculation of Standard Errors

This appendix provides necessary background information on how to calculate the spatial heteroskedasticity and autocorrelation consistent (SHAC) standard errors in the characteristic regressions, built on the theoretical work of Conley (1999). First, I briefly review the necessary concepts about covariance estimator in standard time series context to motivate the further discussions. Second, I explain the spatial covariance estimator in both cross section and panel data. Third, I compare the SHAC covariance estimator with some other covariance estimators in the literature. Fourth, I point out some implementation issues on spatial covariance estimator. Finally, I give details on constructing the economic distances.

D.1. Basic Concepts in Time Series

Let $\{V_t\}_{t=1}^{T}$ be a sequence of an $N \times 1$ random vector where N is fixed and t goes to infinity ("large T" case). Under independent and identically distributed (*iid*) distributional assumption of $\{V_t\}_{t=1}^{T}$, or some suitable mixing conditions of the distribution of $\{V_t\}_{t=1}^{T}$, the law of large number (LLN) implies

$$\frac{1}{T}\sum_{t=1}^{T} V_t \longrightarrow E\left[V_t\right] \tag{D.1}$$

Without loss of generality, the expected value of $\{V_t\}_{t=1}^{T}$ is assumed to be zero, $E\left(V_t\right) = 0$.

The central limit theorem (CLT) under *iid* assumption implies

$$\text{(D.2)} \qquad \frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_t \longrightarrow N(0\ ,\ \Omega)$$

where

$$\text{(D.3)} \qquad \Omega \triangleq Var\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_t\right) = \frac{Var\,(V_t)}{T} = \frac{1}{T} E\,(V_t V_t') \equiv \frac{1}{T} C(0)$$

Define $C(0) \triangleq E\,(V_t V_t')$, and its sample estimate is

$$\widehat{C}(0) = \frac{1}{T} \sum_{t=1}^{T} V_t V_t'$$

By LLN in (D.1), $\widehat{C}(0)$ is a consistent estimator for $C(0)$. Under mixing condition, (D.3) implies

$$\text{(D.4)} \qquad \Omega \triangleq \lim_{T \longrightarrow \infty} Var\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} V_t\right) = \sum_{k=-\infty}^{\infty} C(k)$$

where the autocovariance function $C(k)$ is defined to as $C(k) \triangleq E\left(V_t V_{t+k}'\right)$. For a given k, the sample autocovariance estimator is

$$\widehat{C}(k) = \frac{1}{T} \sum_{t=1}^{T-k} V_t V_{t+k}', \text{ and } \widehat{C}(-k) = \widehat{C}(k)'$$

By LLN in (D.1), $\widehat{C}(k)$ is also a consistent estimator for $C(k)$, for each k.

D.2. Time-Series HAC Estimator for Ω

Newey and West (1987) suggest the following heteroscedasticity autocorrelation consistent (HAC) estimator for Ω,

$$\widehat{\Omega}_T = \sum_{k=-(T-1)}^{T-1} \lambda_T(k) \cdot \widehat{C}_T(k) = \sum_k \lambda_T(k) \cdot \left(\frac{1}{T} \sum_t V_t V'_{t+k} \right) \tag{D.5}$$

where the weighting function $\lambda_T(k) \to 1$ for all k as $T \to \infty$, but slowly enough so that $Var(\widehat{\Omega}_T) \to 0$ as $T \to \infty$ to ensure the consistent estimate of Ω. To guarantee the positive definiteness of covariance estimate $\widehat{\Omega}_T$, Newey and West (1987) also propose to choose the following weighting function (i.e., Bartlett kernel),

$$\lambda_T(k) = \begin{cases} 1 - \frac{|k|}{L_T}, \ |k| \leq L_T \\ 0, \text{ otherwise} \end{cases} \tag{D.6}$$

where L_T is the bandwidth.[1]

D.3. Spatial HAC Estimator for Ω

The calculations of standard errors in the characteristic regressions follows Conley (1999) who studies cross sectional generalized methods of moments estimate (GMM). He generalizes the Newey-West standard covariance estimator so that the weighting schemes depend on the "economic distance" (rather than time) between observations. In the time-series context, time is a natural index to describe the correlation structure. But in the cross-sectional context, usually it is difficult to find such straightforward index. Spatial

[1]The weights are constructed by resorting to the spectral density representation of the time series. In particular, the Fourier transform of Bartlett window is non-negative, so the estimated covariance matrix will always be positive semidefinite. See Priestley (1982) for detailed discussions of spectral analysis.

model uses the economic distance as an index and provides a parsimonious and sensible solution to model the correlation structure. In the literature, this type of variance-covariance estimator underlying the standard error calculation is called spatial HAC (SHAC) estimator. To motivate the spatial HAC estimator, and illustrate the connection between SHAC estimator and Newey-West time series HAC estimator, we first rewrite (D.5) in a slightly different form,

$$\widehat{\Omega}_T = \sum_{k=-(T-1)}^{T-1} \lambda_T(k) \cdot \widehat{C}_T(k) = \frac{1}{T} \sum_{r=1}^{T} \sum_{c=1}^{T} \lambda_T(|c-r|) \cdot V_r V_c' \tag{D.7}$$

where the weights depend on the time lag (the distance in time), namely, $\lambda_T(\cdot)$ is near 1 for small $|c-r|$ and near 0 for large $|c-r|$.

Specifically, consider an example of linear cross-sectional regression model,

$$y_{s_i} = X_{s_i}' \beta + u_{s_i} \text{ for } i = 1, ..., N_\tau$$

where there are N_τ cross-sectional units, and each unit i has a "location share" s_i and all units are in the k dimensional space Z^k, $\{s_i \in Z^k\}_{i=1}^{N_\tau}$. The dependent variable y_{s_i} is a scalar and the independent variables X_{s_i} is $k \times 1$ vector. In the rather general case, one may assume that there exist an instrument Z_{s_i} such that the moment condition holds, $E(Z_{s_i} u_{s_i}) = 0$, then the least square estimate is

$$\widehat{\beta}_\tau = \left(\frac{1}{N_\tau} \sum_{i=1}^{N_\tau} Z_{s_i} X_{s_i}' \right)^{-1} \left(\frac{1}{N_\tau} \sum_{i=1}^{N_\tau} Z_{s_i} y_{s_i} \right). \tag{D.8}$$

The asymptotic distribution of β follows,

$$\text{(D.9)} \qquad \sqrt{N_\tau}(\widehat{\beta}_\tau - \beta) = \left(\frac{1}{N_\tau}\sum_{i=1}^{N_\tau} Z_{s_i}X'_{s_i}\right)^{-1}\left(\frac{1}{\sqrt{N_\tau}}\sum_{i=1}^{N_\tau} Z_{s_i}u_{s_i}\right).$$

In order to use spatial law of large number (LLN) and the central limit theorem (CLT), covariance stationarity and mixing conditions are imposed. Briefly speaking, stationarity means that the covariance between X_{s_i} and X_{s_j} only depends on the distance between unit i and j in the space Z^k, not on the location s_i and s_j; mixing means that the covariance between X_{s_i} and X_{s_i+h} approaches zero as distance $h \to \infty$.[2] For formal definitions and other regularity conditions in spatial asymptotic theory, see Conley (1999) for details. Assume the processes $\{X_{s_i}, Z_{s_i}, u_{s_i} : s_i \in Z^k\}_{i=1}^{N_\tau}$ are covariance stationary, mixing and well-behaved, by spatial LLN, as $N_\tau \to \infty$ the first part of (D.9) yields,

$$\text{(D.10)} \qquad \frac{1}{N_\tau}\sum_{i=1}^{N_\tau} Z_{s_i}X'_{s_i} \longrightarrow^p E\left(Z_sX'_s\right)$$

and by spatial CLT, as $N_\tau \to \infty$ the second part of (D.9) yields,

$$\text{(D.11)} \qquad \frac{1}{\sqrt{N_\tau}}\sum_{i=1}^{N_\tau} Z_{s_i}u_{s_i} \longrightarrow^D N(0,\, \Omega) \ \text{ where } \Omega = \sum_{s\in Z^k} E\left(Z_0u_0\right)\left(Z_su_s\right)'.$$

In (D.11), the subscripts 0 and s stand for the origin point 0 and point s in the k-dimensional space Z^k, respectively.[3] For example, in the 2-dimensional space Z^2, the

[2] In fact, when the distance is defined as time period, these are exactly the covariance stationary and mixing definitions in time series.

[3] Note if the spatial process is covariance stationary, we can start at any origin point.

coordinate of point $s \triangleq (i,j)$ and Ω in the equation (D.11) becomes

$$\Omega = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} E\left(Z_{i,j}u_{i,j}\right)\left(Z_{i+m,j+n}u_{i+m,j+n}\right)'.$$

where the subscripts m and n are used to denote the distances between any two points in the space Z^2.

Define the residuals $\widehat{u}_{s_i} = y_{s_i} - X'_{s_i}\widehat{\beta}_\tau$, and we estimate Ω_τ as

$$\widehat{\Omega}_\tau = \frac{1}{N_\tau}\sum_{i=1}^{N_\tau}\sum_{j=1}^{N_\tau}\lambda_\tau(||s_i - s_j||)\cdot(Z_{s_i}\widehat{u}_{s_i})(Z_{s_j}\widehat{u}_{s_j})' \tag{D.12}$$

where $\lambda_\tau(\cdot)$ is a spatial weighting function of economic distance between location s_i and s_j. Under the condition that $\lambda_\tau(d) \to 1$ for any distance d as $N_\tau \to \infty$, but slowly enough so that $Var(\widehat{\Omega}_\tau) \to 0$ as $N_\tau \to \infty$, $\widehat{\Omega}_\tau$ is a consistent estimate of Ω, and we call $\widehat{\Omega}_\tau$ the spatial type HAC estimator. Note the asymptotic results used from equation (D.9) to equation (D.12) are based on $N_\tau \to \infty$, i.e., the number of units in the space Z^k increases to infinity. One kernel function analogous to the Bartlett kernel in the Newey-West time series HAC estimator is specified as

$$\lambda_\tau(d) = \begin{cases} 1 - \frac{|d|}{L_\tau}, \ |d| \le L_\tau \\ 0, \text{otherwise} \end{cases} \tag{D.13}$$

where d is the input of "economic distance" between unit i and j, and L_τ is the bandwidth (see Conely, 1999; Chen and Conley, 2001).[4] Similar to HAC estimator in time series

[4]This estimator is always positive semi definite because the spectral window corresponding to the Bartlett function space domain window is always non-negative. According to Bochner's theorem (see Priestley, 1982), a necessary and sufficient condition for a valid covariance function is that its Fourier transform is non-negative.

context, the choice of bandwidth L_τ also reflects a tradeoff between bias and variance of the estimate $\widehat{\Omega}_\tau$. The nonparametric spatial estimator in (D.12) is quite robust; no particular data generating process (DGP) is assumed for the error dependence structure as long as some regularity conditions are satisfied, the economic distances can be endogenous, and they can be measured with errors as long as the errors are bounded.

D.4. Spatial HAC Estimator for Ω in Time-Series Cross-Sectional Regression

The previous sections outlined the relationship between the familiar Newey-West HAC estimator and generic SHAC estimator. Now we consider the representation of SHAC estimator in the context of time-series cross-sectional regressions. The typical balanced panel data includes N firms $(i = 1, ..., N)$ and T time periods $(t = 1, ..., T)$.[5] The more general spatial model nests the time into a separate dimension. The economic distance configuration are represented by a set of points in the k-dimensional space Z^k (one dimension is time and $k \geq 2$), each firm i at each time t is modelled to reside in Z^k, with location $s_{i,t}$. The linear time-series cross-sectional regression model under this setting is given by

$$y_{it} = X_{it}'\beta + \epsilon_{it} \text{ for } i = 1, ..., N \text{ and } t = 1, ..., T \tag{D.14}$$

where the dependent variable y_{it} is a scalar and the regressors X_{it} is $k \times 1$ vector. In the rather general case, one may assume that there exist a $k \times 1$ instrument variable Z_{it}

[5]For notation brevity, I focus on balanced panel data. All the derivation of asymptotic theory and estimation procedures are also applicable for the unbalanced panel. In practice, the estimator given in equation (D.19) below will automatically handle the unbalanced panel case since it is implemented on pairwise distances and cross products.

such that $E\left(Z_{it}\epsilon_{it}\right)=0$, then the least square estimate is

(D.15) $$\widehat{\beta}_{OLS}=\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}X_{it}'\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}y_{it}\right)$$

The asymptotic distribution of $\widehat{\beta}_{OLS}$ follows,

(D.16) $$\sqrt{NT}(\widehat{\beta}_{OLS}-\beta)=\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}X_{it}'\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}\epsilon_{it}\right)$$

The balanced panel data consists of the realizations of stochastic process $\{X_{it},Z_{it}\}$ at a collection of locations $s_{i,t}$ for each firm i at each time period t, $\{X_{it},Z_{it}:i=1,...,N \text{ and } t=1,...,T\}$.[6] Under the assumption that this process is covariance stationary, mixing and well-behaved, by spatial LLN as $NT\to\infty$, the first part of (D.16) yields,

(D.17) $$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}X_{it}'\longrightarrow^{p}E\left(Z_{it}X_{it}'\right)$$

Furthermore, assume the process of error $\{\epsilon_{it}\}$ at locations $s_{i,t}$ is also covariance stationary and mixing, by spatial CLT as $NT\to\infty$ the second part of (D.16) yields,

(D.18) $$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}\epsilon_{it}\longrightarrow^{D}N(0,\,\Omega)\quad\text{where }\Omega=\sum_{s\in Z^{k}}E\left(Z_{0}\epsilon_{0}\right)\left(Z_{s}\epsilon_{s}\right)'$$

where the subscripts 0 and s denote the origin point 0 and point s in the k-dimensional space Z^k respectively. Recall time is always a dimension in this setting, and hence implicitly there is an infinite sum over time lags like the equation (D.4) in time series context. For example, in the 3-dimensional space Z^3 (two cross section dimension plus one time dimension), the coordinate of point $s=(i,j,t)$ and Ω in the equation (D.18)

[6]This is the typical definition of spatial model in the context of random field in the geostatistics literature.

simply becomes

$$\Omega = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} E\left(Z_{i,j,t} u_{i,j,t}\right)\left(Z_{i+m,j+n,t+k} u_{i+m,j+n,t+k}\right)'.$$

where the subscripts m and n are used to denote the distances in the cross-sectional dimensions, and the subscript k is used to denote the time lag in time-series dimension, all of which are used to describe the distance between two points in Z^3.

Define the residuals $u_{it} = y_{it} - X_{it}' \widehat{\beta}_{OLS}$, and similar to the equation in (D.12), we obtain the consistent estimate of Ω

(D.19)

$$\begin{aligned}
&\widehat{\Omega}_{N,T} \\
= \quad & \frac{1}{N} \sum_{k=-(T-1)}^{T-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{T} \left(\sum_{t} \lambda_{N,T}(||s_{i,t} - s_{j,t+k}||) \cdot (Z_{i,t} u_{i,t})(Z_{j,t+k} u_{j,t+k})' \right) \\
= \quad & \frac{1}{NT} \left\{ \begin{array}{l}
\sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{N,T}(||s_{i,t} - s_{j,t}||) \cdot (Z_{i,t} u_{i,t})(Z_{j,t} u_{j,t})' \\
+ \sum_{k=1}^{T-1} \sum_{t=1}^{T-k} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{N,T}(||s_{i,t} - s_{j,t+k}||) \cdot (Z_{i,t} u_{i,t})(Z_{j,t+k} u_{j,t+k})' \\
+ \sum_{k=1}^{T-1} \sum_{t=k+1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{N,T}(||s_{i,t} - s_{j,t-k}||) \cdot (Z_{i,t} u_{i,t})(Z_{j,t-k} u_{j,t-k})'
\end{array} \right\}
\end{aligned}$$

Denote the distance as $(d, k) \triangleq ||s_{i,t} - s_{j,t+k}||$, the combined kernel is specified as

$$\text{(D.20)} \qquad \lambda_{N,T}(d,k) = \lambda_N(d) \cdot \lambda_T(k) = \begin{cases} (1 - \frac{|d|}{L_N})(1 - \frac{|k|}{L_T}), \ |d| \leq L_N \text{ and } |k| \leq L_T \\ 0, \text{ otherwise} \end{cases}$$

Under the condition that $\lambda_{N,T}(\cdot) \to 1$ for any distance as $NT \to \infty$, but slowly enough so that $Var(\widehat{\Omega}_{N,T}) \to 0$ as $NT \to \infty, \widehat{\Omega}_{N,T}$ is a consistent estimate of Ω.

Finally, we observe that the asymptotic results from equation (D.17) to equation (D.20) are based on $NT \to \infty$, either large N fixed T case, or large T fixed N case. Also this kernel is specified as a product of the Bartlett kernel in (D.6) and (D.13), and the choice of bandwidth L_N and L_T reflects a tradeoff between bias and variance of the estimate $\widehat{\Omega}_{N,T}$. This general spatial HAC estimator takes care of the time-series serial-correlations, cross-sectional correlations, time-series cross-autocorrelations and heteroscedasticity.[7]

D.5. Special Cases of Spatial HAC Estimator for Ω in Time-Series Cross-Sectional Regressions

The spatial HAC estimator nests several commonly used variance-covariance estimator as special cases, which illustrate the generality of spatial HAC estimator. The easiest way to see these connections is to look at the weighting functions in various covariance estimators.

- Newey-West Time-Series HAC Estimator: The first special case is the *Newey-West time series HAC estimator*. When we use the time as the natural index of distance in (D.12), the spatial HAC estimator is the Newey-West time-series HAC estimator. In the time-series cross-sectional regression set up in (D.14), if we ignore the cross-sectional dependence structure at any point in time, it is equivalent to set $d = 0$ in (D.20). Clearly, the resulting weighting function

[7]However, unlike the Newey-West time series HAC estimator, where there is symmetry of the form $\widehat{C}(k) = \widehat{C}(-k)$, generally there is no such symmetry in the spatial HAC estimator. This is because the kernel function of Newey-West estimator is $1 - \frac{|k|}{L_T}$, and so the k-th lag and lead get the same weight. In spatial context, the distance between firm i at time t and firm j at time $t+k$ is in general not equal to that between firm i at time $t-k$ and firm j at time t. Hence the kth lag and lead may get the different weights.

expressed in Kernel form, is the same as the weighting function of Newey-West time-series HAC estimator in (D.6).

- Cluster Type Covariance Estimator of Ω: The second special case is the *cluster type covariance estimator* in Rogers (1993). For cluster type covariance estimator, the weighting function is a function of the locations s_i and s_j with respect to their homogeneity,

$$\widehat{\Omega}_\tau = \frac{1}{N_\tau} \sum_{i=1}^{N_\tau} \sum_{j=1}^{N_\tau} \lambda_\tau(s_i, s_j) \cdot (Z_{s_i} \widehat{u}_{s_i})(Z_{s_j} \widehat{u}_{s_j})' \tag{D.21}$$

where

$$\lambda_\tau \triangleq \begin{cases} 1, \text{ if unit } i \text{ and } j \text{ in same group} \\ 0, \text{ otherwise} \end{cases}$$

Note this estimator in equation (D.21) assumes no correlation structure between clusters.

- White heteroscedasticity estimator of Ω: The third special case happens when every unit is a cluster (i.e., each cluster has only one observation), the estimator $\widehat{\Omega}_\tau$ is in fact the *White heteroscedasticity estimator* of Ω (White, 1984).
- Two-way cluster type covariance estimator of Ω: The fourth special case is so-called *two-way cluster type covariance estimator* in Cameron, Gelbach and Miller (2006), and Thompson (2006). Under the SHAC estimator, we may define one

cluster as

$$\lambda_\tau \triangleq \begin{cases} 1, \text{ if } i \text{ and } j \text{ share the same firm} \\ 1, \text{ if } i \text{ and } j \text{ share the same time} \\ 1, \text{ if } i = j \\ 0, \text{ otherwise} \end{cases} . \tag{D.22}$$

We may also define three clusters as in Cameron, Gelbach and Miller (2006), and Thompson (2006),

$$\begin{aligned} \lambda_\tau^{[1]} &\triangleq \begin{cases} 1, \text{ if } i \text{ and } j \text{ share the same firm} \\ 1, \text{ if } i = j \\ 0, \text{ otherwise} \end{cases} \\ \lambda_\tau^{[2]} &\triangleq \begin{cases} 1, \text{ if } i \text{ and } j \text{ share the same time} \\ 1, \text{ if } i = j \\ 0, \text{ otherwise} \end{cases} \\ \lambda_\tau^{[3]} &\triangleq \begin{cases} 1, \text{ if } i = j \\ 0, \text{ otherwise} \end{cases} \end{aligned} \tag{D.23}$$

where $\lambda_\tau^{[1]}$ denotes the cluster on firm, $\lambda_\tau^{[2]}$ denotes the cluster on time, $\lambda_\tau^{[3]}$ denotes an "adjustment cluster". It is clear that

$$\lambda_\tau = \lambda_\tau^{[1]} + \lambda_\tau^{[2]} - \lambda_\tau^{[3]},$$

hence the two-way cluster type covariance estimator in (D.23) is a simple rewriting of (D.22), and they are mathematically equivalent. Compared to the regular cluster covariance estimator characterized by (D.22), the only merit of the two-way cluster covariance estimator characterized by (D.23) is ease of obtaining $\lambda_{\tau}^{[1]}$, $\lambda_{\tau}^{[1]}$ and $\lambda_{\tau}^{[1]}$ from some "canned" statistical packages. As a note to the two-way cluster type of covariance estimator, one may adjusting the weighting function in (D.22) accordingly to obtain multi-way cluster type of covariance estimator.

For the cluster type of covariance estimator, as in spatial HAC estimator, the choices of clusters and the choice of number of clusters also reflect a tradeoff between bias and variance of the covariance estimates. However, it is clear that even sometimes the choices of clusters seem to be intuitive, the discrete clusters could be a rough measure to model dependence structure because the way of assigning the weight in clusters is rather extreme.[8]

D.6. Implementation Issues

There are two key issues related to the implementation of the spatial HAC estimator. First, the spatial model assumes that the correlation structure is a function of economic distance, and it is imperative to examine whether this type of spatial covariance pattern described by the economic distance indeed shows up in the data. The fundamental assumption of spatial model is that as the distance between two points in a space getting far away their pairwise correlation diminishes. If the spatial model is valid empirically, then we expect to see that the estimated covariance is a decreasing function of distance. Second, in order to use spatial HAC estimator in (D.20), we need a judgement call on the bandwidth

[8]From a theoretical point of view, the cluster types of covariance estimators will not always gurantee positive definitness of the estimator of the estimated covariance. This is because the uniform weighting function implies that the spectral window corresponding to its Fourier transform can be negative in some regions. Fortunately, this rarely happens in practice.

L_N and L_T. The threshold allowing for spatial correlation in the cross-sectional dimensions is much lower than its counterpart of serial correlation in the time-series dimension. This is rather intuitive. Let time be the only dimension in the space Z^1 and all the firms reside like points on this straight line. For each point in Z^1 the correlation is only through its neighbors in this 1-dimensional straight line. Now suppose we add another dimension and transform the original space Z^1 into the space Z^2. For each point in Z^2 space, the correlation is characterized through its neighbors on 2-dimensional sphere, which has much stronger impact than before. Therefore, to some extent choosing bandwidth L_N is more delicate than L_T in the Newey-West type of time-series HAC estimator.

To resolve these important issues, in the first step I perform the local average estimation on the covariance of residuals u_{it}.[9] This type of non-parametric estimator is suggested in Conley and Dupor (2003). Let $d_{ij,t} \triangleq ||s_{i,t} - s_{j,t}||$, for a given input of economic distance h, the spatial covariance estimate for residuals u_{it} is

$$\widehat{C}(h) = \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} W_{N,T}(h - d_{ij,t}) \cdot u_{it}u_{jt} \text{ for } h > 0 \tag{D.24}$$

where the weighting function $W_{N,T}(\cdot)$ is normalized to sum to one and concentrates its mass at zero as the sample size goes to infinity.[10] The plot of spatial covariance estimate $\widehat{C}$ against the distance h is very helpful to validate the spatial model assumption. The estimator in equation (D.24) assumes no parametric model structure in residuals correlation and is essentially based on the replicates of residual covariance patterns across firms. This nonparametric approach is robust to model misspecification and outliers.

[9]The covariance of the elements in $Z_{it}u_{it}$ is also examined since this is directly linked to standard errors.
[10]I use the Gaussian kernel with adaptive bandwidth selection as in Fan and Gijbels (1996).

To further examine whether the estimated spatial covariances are statistically different from zero, in the second step I use a bootstrap method to construct an acceptance region for the null hypothesis of serial and cross-sectional independence. This acceptance region facilitates the choice of appropriate bandwidth L_N in spatial HAC estimator. Specifically, conditional on each firm's location $\{s_{i,t}\}$, I simulate the draws $\{u_{it}^{BS}\}$ from a distribution with the same stationary and marginal distribution of the residuals $\{u_{it}\}$, i.e., IID samples with replacement from the residuals empirical distribution. The bootstrapped samples are generated by preserving heterogeneity across firms and by imposing homogeneity across time for each firm. After obtaining the IID samples, the spatial covariance estimate for bootstrapped residuals u_{it}^{BS} follows exactly the same from the equation (D.24)

$$\widehat{C}_{BS}(h) = \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} W_{N,T}(h - d_{ij,t}) \cdot u_{it}^{BS} u_{jt}^{BS} \text{ for } h > 0 \tag{D.25}$$

where the given distance input h, pairwise distance $d_{ij,t}$ and the weight $W_{N,T}(h - d_{ij,t})$ are kept the same as in (D.24). By design, the estimated bootstrap covariance $\widehat{C}_{BS}$ is equal to zero. To construct the acceptance region of serial and spatial independence, the bootstrap experiment above is repeated many times. For each distance h, the lower 2.5 and upper 97.5 percentile of the estimated bootstrap covariance $\widehat{C}_{BS}$ yields the lower and upper bound of independence region respectively. This region is symmetrically around zero by design. When the estimated $\widehat{C}\,(h)$ in (D.24) is contained in the region at certain distance, we cannot reject the null hypothesis that the spatial correlation at this distance h is not statistically different from zero. In summary, the bootstrap exercise helps to examine the precision of spatial covariance estimate and yields an objective choice of the bandwidth length in spatial HAC estimator.

D.7. Construction of "Economic Distance"

There is a great amont of flexibility in specifying both the determinant of economic distance and how the economic distance determines the weighting scheme. To avoid the danger of being completely *ad hoc*, I construct several sets of economic distances motivated by empirical regularities from the prior literatures.

The first economic distance is built on two dimensions. The first dimension is natural time. The natural time dimension is easy to understand as the data structure is a time-series. The second dimension is the firm's pairwise Euclidian distance in the cross-sectionally demeaning and standardized book-to-market equity ratios (B/M), logarithm of market capitalization and past 12-month returns. The empirical evidence suggests that characteristics such as B/M, size and past returns provide better *ex-ante* forecasts of cross-sectional patterns of future returns, and also does a better job in matching future realized returns (see Daniel and Titman, 1997; and Daniel, Grinblatt, Titman and Wermers, 1997). Essentially, the empirical evidence suggests the B/M, size and past returns determine the economic distance among firms, and in turn such distance provide good characterization of return structure. Therefore, if one cares about the correlation structures of returns, both *ex-ante* and *ex-post*, then B/M, size and past returns provide one way to characterize such return correlation structure. In particular, we define the economic distance as the following Euclidian distance characterization,

$$\lambda_\tau(||s_i - s_j||) \triangleq \sqrt{(bm_{i,t} - bm_{j,t})^2 + (size_{i,t} - size_{j,t})^2 + (pastret_{i,t} - pastret_{j,t})^2}$$

where bm_i is the cross-sectionally demeaned and normalized B/M ratio for observation i for $i = 1,...,N_\tau$

$$bm_i \triangleq \frac{\left(B/M_{i,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} B/M_{n,t}\right)}{\sqrt{\frac{1}{N_\tau - 1}\sum_{n=1}^{n=N_\tau}\left(B/M_{n,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} B/M_{n,t}\right)^2}};$$

$size_{i,t}$ is the cross-sectionally demeaned and normalized logarithm of market capitalization for observation i for $i = 1,...,N_\tau$

$$size_i \triangleq \frac{\left(size_{i,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} size_{n,t}\right)}{\sqrt{\frac{1}{N_\tau - 1}\sum_{n=1}^{n=N_\tau}\left(size_{n,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} size_{n,t}\right)^2}};$$

$pastret_{i,t}$ is the cross-sectionally demeaned and normalized past 12-month return for observation i for $i = 1,...,N_\tau$

$$pastret_i \triangleq \frac{\left(ret_{i,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} ret_{n,t}\right)}{\sqrt{\frac{1}{N_\tau - 1}\sum_{n=1}^{n=N_\tau}\left(ret_{n,t} - \frac{1}{N_\tau}\sum_{n=1}^{n=N_\tau} ret_{n,t}\right)^2}}.$$

Standard errors computed from this estimator is denoted as "characteristic distance standard errors". This is the standard errors I reported in this paper.

The second economic distance is the aggregated industry classification based on Fama and French (1998). Using industry classification as the clustering mode is motivated by Moskwitz and Grinblatt (1999) who show that "... *industry momentum drives much of individual stock momentum, and stocks within an industry tend to be much more highly correlated than stocks across industries...*" (p. 1251). Using the industry classification as the economic distance essentially boils down to using the cluster type standard errors in (D.21) where firms belonging to the same industry are assumed to have arbitrary correlation structure in the dimensions of time and cross-section but there is no correlation between industries. Standard errors computed from this cluster type estimator is denoted as "industry cluster standard errors". I have verified that the cluster type of standard errors are similar to the standard errors obtained from the Newey-West HAC estimator.[11]

The last economic distance measure is not based on any economic meaning. Rather it is based on residual cluster analysis. Essentially this estimation procedure is to let the error distributional structure to determine the formation of clusters. As the purpose of cluster analysis is to maximize the within-cluster correlations and minimize the between-cluster correlations, it is suitable to assign cluster identification to each firm-year observations. In practice, we first estimate the time-series and cross-sectional residuals from the firm by firm time-series regressions. Then we carry out the cluster analysis of these residuals. Finally we map the residuals cluster position from the cluster analysis back to the original cross-sectional time-series regressions, and form clusters to compute the regular cluster standard errors. Standard errors computed from this estimator are denoted as "generalized cluster standard errors". I have verified that this type of standard errors are close to the "industry cluster standard errors".

[11] Admittedly, assumption of no correlation between industries is extreme. For example, in an empirical test of APT model, Connor and Korajczyk (1988) reject the block-diagonality of the idiosynractic covariance metric in which they define blocks by three-digit SIC code industries. One may modify this approach by allowing non-zero economic distance among industries. This is done by estimating the correlation coefficients from the daily industry portfolio returns. The correlation coefficients are used as the distance between industries.

D.8. Implementation of Spatial HAC Estimator for Ω in Time-Series Cross-Sectional Regressions

Figure 2.3 plots the economic distance constructed as the Euclidian normal of Z-scores of past return, logarithm of market capitalization and book to market equity ratios against the spatial covariance estimate of residuals for the following regression model in (2.11).

Except for a small amount of local variations, the global pattern implies that the constructed economic distance characterize the covariance structure well. For the distance over the range 0.3 to 4.8, the estimated residual covariances are clearly persistent. In other words, as the economic distance moves further away, the estimated covariances decay monotonically. Figure 2.4 plots the empirical distribution of pairwise correlations of residuals along the economic distance constructed as in figure 2.3. When the distance is too small or too large, there is only small amount of pairwise distances and residual cross products. Since the estimator in equation (D.24) is based on the replicates of residual covariance patterns across firms, the estimated spatial covariances are less precise at these extreme distances. However, because these correlations only accounts for at most 5% of total correlations, the overall impact is small. Figure 2.5 plots the economic distance constructed as the Euclidian normal of Z-scores of past return, logarithm of market capitalization and book to market equity ratios against the acceptance region of spatial independence among residuals. Because of computational burden, in the construction of acceptance region, I only choose 10 equally space points between zero and the maximal economic distance, and conduct the bootstrap replicates 50 times. The bootstrap evidence suggests that the residual covariance estimates are far from being zero, even at 75% of the maximal economic distance. Therefore, being conservative, I retain the maximal economic distance as the bandwidth in my estimation of spatial HAC estimator. However, using the 75% of the maximal economic distance as the bandwidth does not change the estimate of t-statistics in any noticeable way.