May 2008
McCombs Research Paper Series No. FIN-02-08

THE UNIVERSITY OF TEXAS AT AUSTIN
McCOMBS SCHOOL OF BUSINESS

Retail Clienteles and the Idiosyncratic Volatility Puzzle

Bing Han
McCombs School of Business
The University of Texas at Austin
bhan@mail.utexas.edu

Alok Kumar
McCombs School of Business
The University of Texas at Austin
akumar@mail.utexas.edu

Retail Clienteles and the Idiosyncratic Volatility Puzzle*

Bing Han

Alok Kumar

Current Version: May 17, 2008

*Both authors are at the Department of Finance, McCombs School of Business, University of Texas at Austin, 1 University Station, B6600, Austin, TX 78712. Bing Han can be reached at 512-232-6822; *email:* `bhan@mail.utexas.edu`. Alok Kumar can be reached at 512-232-6824; *email:* `akumar@mail.utexas.edu`. We thank Aydogan Alti, Robert Battalio, Keith Brown, Fangjian Fu, Lorenzo Garlappi, Rick Green, John Griffin, Jay Hartzell, Jennifer Huang, Narasimhan Jegadeesh, Danling Jiang, Paul Koch, George Korniotis, Francisco Perez-Gonzalez, Stefan Ruenzi, Clemens Sialm, Sheridan Titman, Masahiro Watanabe, Mark Weinstein, Yuhang Xing, Joe Zhang and seminar participants at UT-Austin, Texas Tech University, Southwind Finance Conference at the University of Kansas, the Tenth Texas Finance Festival, and UT-Dallas for helpful comments and valuable suggestions. We would like to thank Jeremy Page for excellent research assistance. We also thank Brad Barber, Sudheer Chava, Paul Koch, Terrance Odean, and Amiyatosh Purnanandam for providing some of the data used in the paper. Of course, we are responsible for all remaining errors and omissions.

Retail Clienteles and the Idiosyncratic Volatility Puzzle

ABSTRACT

This study provides a simple economic explanation for the puzzling negative relation between idiosyncratic volatility and average stock returns identified in Ang, Hodrick, Xing, and Zhang (2006, 2008). We show that retail investors prefer to hold and actively trade high idiosyncratic volatility stocks due to their propensity to speculate and because those stocks offer greater opportunities for experiencing utility over realizing gains. We find that stocks with high proportion of retail trading tend to earn lower future returns, especially if they are speculative stocks and are more difficult to arbitrage. Furthermore, the negative volatility-return relation is concentrated among stocks that are dominated by retail investors, exhibit strong speculative characteristics, and have high arbitrage costs. Among stocks with low levels of retail trading, average returns increase with idiosyncratic volatility. Collectively, our evidence indicates that the level of retail trading in a stock is a critical determinant of the volatility-return relation.

I. Introduction

The trade-off between risk and return is a central theme in asset pricing and investment. On the one hand, traditional asset pricing models based on frictionless markets and complete information argue that only systematic risk should command a risk premium (e.g., the Sharpe (1964) and Lintner (1965) capital asset pricing model and the Ross (1976) arbitrage pricing theory). On the other hand, Merton (1987), Jones and Rhodes-Kropf (2004), Malkiel and Xu (2006), and Barberis and Huang (2008) develop asset pricing models in which returns are a positive function of idiosyncratic risk.[1] The arguments of these models center on the inability of a certain class of investors to hold diversified portfolios. Those under-diversified investors would require extra compensation for holding idiosyncratic risk, which could generate a positive relation between idiosyncratic volatility and returns. If one group of investors is unable or unwilling to hold the market portfolio for exogenous reasons, the remaining investors will also be unable to hold the market portfolio. Therefore,

[1] An exception is Johnson (2004), who develops a model in which stocks with high leverage and high idiosyncratic returns earn lower returns. However, Ang, Hodrick, Xing, and Zhang (2008) provide empirical evidence inconsistent with this prediction. Also, see the evidence in Section V.E.

idiosyncratic risk could be priced to compensate rational investors for their inability to hold the market portfolio.

In contrast to the theoretical predictions, recent empirical studies document a negative idiosyncratic risk-return relation. This evidence contradicts extant theories that predict either a zero or positive premium for idiosyncratic risk. For example, Ang, Hodrick, Xing, and Zhang (2006) find that a stock's daily return idiosyncratic volatility over the previous month negatively predicts its average return next month. Ang, Hodrick, Xing, and Zhang (2008) show that this negative volatility-return relation is robust. It holds even after controlling for other well-known predictors of cross-sectional returns and the puzzling negative relation is observed in several international stock markets. We refer to the robust, negative idiosyncratic risk-return relation identified in AHXZ studies as the idiosyncratic volatility "puzzle."

In this study, we attempt to resolve this idiosyncratic volatility puzzle. Specifically, we examine whether the heterogeneity in investors' idiosyncratic volatility preferences affects the pricing of idiosyncratic risk. Investors are assumed to dislike volatility in traditional theories that focus on the risk-return trade-off. However, several recent studies suggest the possibility that certain types of investors might be attracted toward high volatility stocks and volatility induced investor clienteles might exist. The preferences of those volatility-seeking investors could influence the volatility-return relation.

In particular, Barberis and Xiong (2008a) study investors who derive utility by realizing gains from the stocks they own.[2] They show that realization utility investors would prefer high volatility stocks because those stocks offer a greater chance of delivering a sizable gain. Further, in the preferred risk habitat hypothesis of Dorn and Huberman (2007), investors specialize in the volatilities of the stocks they hold, where the variation in this specialization corresponds to the variation in risk aversion. They demonstrate that less risk-averse investors prefer to hold and trade high volatility stocks. Investors' affinity for high idiosyncratic volatility stocks could also reflect their strong desire to speculate or gamble (e.g., Shefrin and Statman (2000), Barberis and Huang (2008), Kumar (2008)). Finally, high idiosyncratic volatility stocks could be attractive to sensation-seeking investors and to investors who trade for entertainment (Dorn and Sengmueller (2006), Grinblatt and Keloharju (2008)).

Motivated by these recent studies, we conjecture that volatility-induced investor clienteles influence the relation between idiosyncratic volatility (IVOL) and expected returns. Specifically,

[2]Shefrin and Statman (1985) propose realization utility as one of the key building blocks of the disposition effect, along with prospect theory. Also, see Barberis and Xiong (2008b).

investors who are attracted toward high idiosyncratic volatility stocks would derive additional non-wealth utility from the act of holding and trading those stocks.[3] Therefore, all else equal, those investors would be willing to pay a premium for high idiosyncratic volatility stocks. In addition, stocks that attract a large clientele of volatility-seeking investors would be more highly valued and have lower subsequent returns. In contrast, among stocks that are dominated by volatility-averse investors, the idiosyncratic volatility premium would be positive. For stocks in the middle, where the influences of investors with heterogeneous volatility preferences balance out, the idiosyncratic volatility premium would be insignificant.

To test this conjecture, we need to measure the extent to which each stock is held and traded by volatility-seeking investors. For this purpose, we compute the "retail trading proportion" (RTP) of each stock using data from the Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ). Specifically, the retail trading proportion of a stock as the monthly dollar value of the buy- and sell-initiated small-trades (trade size below $5,000) divided by the dollar value of the total market volume in the same month.

Our trading-based measure of idiosyncratic volatility-based retail clientele is motivated by two observations. First, although previous studies suggest that multiple mechanisms could generate idiosyncratic volatility clienteles, they all imply that investors who prefer high volatility stocks would trade them more often. Therefore, the magnitude of investor-level trading could be useful for identifying idiosyncratic volatility-based clienteles.

Second, we focus on the trades of retail investors because institutional investors may be reluctant to hold high idiosyncratic volatility stocks (we confirm this empirically). Many institutions are governed by prudent man rules, which require institutional investors with fiduciary obligations to invest in "high quality" stocks (e.g., Badrinath, Gay, and Kale (1989), Del Guercio (1996)). It might be difficult to justify high idiosyncratic volatility stocks as prudent investments. Institutional investors typically hold diversified portfolios and are reluctant to deviate from the market (e.g., Cohen, Gompers, and Vuolteenaho (2002), Almazan, Brown, Carlson, and Chapman (2004)). Institutions are also less likely to respond to fads or factors that generate "noise" trading. For example, if the attraction for high volatility stocks reflects a desire to experience greater realization utility, then the volatility-seeking clientele would display greater disposition effect.[4] However, previous studies have already demonstrated that institutional investors are less subject to the disposition

[3] In addition to the channels mentioned above, investors could derive anticipatory utility such as dream utility or hope utility (Clotfelter and Cook (1989)) from holding stocks with high idiosyncratic volatility.

[4] Barberis and Xiong (2008a) show that realization utility investors exhibit the disposition effect.

effect (e.g., Shapira and Venezia (2001), Feng and Seasholes (2005)).

Using actual retail holdings and trading data from a large U.S. discount brokerage house, we show that brokerage investors as a group overweight and more actively trade stocks that have high idiosyncratic volatility and high small-trades volume. Furthermore, using cross-sectional regressions, we show that a stock's retail trading proportion is significantly positively associated with its idiosyncratic volatility. In contrast, using the 13F data, we find that institutional investors as a group overweight low idiosyncratic volatility stocks. These results support our conjecture that retail investors are more likely to exhibit a strong preference for high idiosyncratic volatility stocks.

Examining the reasons for retail investors' strong attraction to high volatility stocks, we find that investors are more likely to hold and actively trade high idiosyncratic volatility stocks for speculative reasons. The levels of retail trading are higher among stocks with strong speculative characteristics. We also find that the characteristics of retail clienteles of high idiosyncratic volatility stocks are remarkably similar to the characteristics of investors who exhibit greater propensity to speculate and gamble as documented in Kumar (2008). In addition, we find support for the predictions of the Barberis and Xiong (2008a) model, which posits that investors prefer high idiosyncratic volatility stocks because those stocks offer a greater opportunity for experiencing higher levels of realization utility.

Due to retail investors' speculative preferences, stocks with high proportion of retail trading tend to be relatively overpriced and have significantly lower future returns, especially if those stocks are more difficult to arbitrage. More importantly, we show that the relation between idiosyncratic volatility and expected returns depends crucially on the extent of speculative retail trading. The negative idiosyncratic volatility premium is concentrated in a segment of the market that is dominated by retail investors (i.e., RTP level is high). Furthermore, this relation is stronger among stocks with speculative characteristics such as high idiosyncratic skewness and for stocks that are more costly to arbitrage. In contrast, among stocks with low RTP, we find that the future returns are positively related to idiosyncratic volatility. These results cannot be successfully explained by alternative hypotheses based on institutional preferences, short-term return reversals, or leverage.

Collectively, our results suggest that the "noise" generated by speculative retail trading is the primary driver of the negative idiosyncratic volatility-return relation. When we account for this noise, a positive volatility-return relation emerges. Thus, the volatility preferences of retail investors provide a simple and an intuitive economic explanation for the idiosyncratic volatility puzzle.

The rest of the paper is organized as follows. Section II briefly discusses the related literature. In Section III, we examine the volatility preferences of retail investors and attempt to identify the determinants of those preferences. In Section IV, we conduct the main asset pricing tests and examine the idiosyncratic volatility-return relation, conditional upon the degree of retail trading. In Section V, we conduct several tests to examine the robustness of our results and entertain alternative explanations for our findings. Section VI concludes with a summary and a brief discussion.

II. Related Research

A rapidly growing literature empirically examines whether and how idiosyncratic volatility is priced in the stock market. Using a time-series analysis, Goyal and Santa-Clara (2003) show that the average stock variance has forecasting power for market returns, and about 85% of average stock variance is idiosyncratic. Based on this evidence, they conclude that idiosyncratic risk levels can predict aggregate market-level returns. However, Bali, Cakici, Yan, and Zhang (2005) argue that aggregate idiosyncratic risk has no reliable and significant predictive power for market returns.

Turning to the pricing of idiosyncratic risk at the individual stock level, Ang, Hodrick, Xing, and Zhang (2006) first document the puzzling result that high idiosyncratic volatility stocks tend to earn low average returns in the future. Several studies have attempted to resolve this idiosyncratic volatility puzzle by employing alternative methods for estimating idiosyncratic volatility.

For example, Malkiel and Xu (2006) follow a portfolio-based approach to minimize errors-in-variables problems and find a positive volatility-return relation. Bali and Cakici (2008) find that the negative volatility premium is non-existent in equal-weighted idiosyncratic volatility portfolios, although Doran, Jiang, and Peterson (2008) find that the idiosyncratic volatility premium is negative even in equal-weighted portfolios if January returns are excluded. Spiegel and Wang (2005) and Fu (2008) use EGARCH type models to capture time-variation in idiosyncratic volatility and find a positive volatility-return relation. Other studies decompose idiosyncratic volatility into expected and unexpected components (e.g., Diavatopoulos, Doran, and Peterson (2006), Chua, Goh, and Zhang (2007)) and examine the relation between unexpected volatility and average future returns.

More recent papers provide alternative perspectives on the idiosyncratic volatility debate. For instance, Jiang, Xu, and Yao (2008) examine whether high idiosyncratic volatility proxies for future earnings shocks. Kapadia (2007) and Boyer, Mitton, and Vorkink (2008) use cross-sectional and expected idiosyncratic skewness measures to examine whether the idiosyncratic volatility puzzle is induced by investors' skewness preferences. They show that with idiosyncratic skewness controls, the negative idiosyncratic volatility premium becomes weaker but it is still significantly negative. Frieder and Jiang (2007) decompose the total volatility into upside volatility and downside volatility and show that only stocks with high upside volatility earn low returns. However, they also document the puzzling result that stocks with higher downside volatility fail to earn higher future returns.

Despite these previous attempts, the original puzzle identified by AHXZ remains because these studies use empirical frameworks that are different from the AHXZ method. In this paper, we follow the AHXZ method closely and show that the idiosyncratic volatility puzzle can be resolved even within their original empirical framework. The novelty of our paper is the exclusive focus on the heterogeneity in the idiosyncratic volatility preferences of investors and the resulting volatility-induced investor clienteles. We investigate how the existence of those clienteles affects the volatility-return relation and provide a simple as well as intuitive economic explanation for the idiosyncratic volatility puzzle identified in AHXZ.

Previous research shows that investor clienteles exist and affect corporate decisions such as dividend payouts and stock splits. The focus in those studies is typically on tax-induced investor clienteles (e.g., Litzenberger and Ramaswamy (1979), Allen, Bernardo, and Welch (2000), Graham and Kumar (2006)). The investor clientele in our paper is defined on the basis of investors' volatility preferences and we show that the preferences and trading behavior of volatility-seeking retail investors can explain the idiosyncratic volatility puzzle.

III. Volatility Preferences of Retail Investors

III.A. Data Sources

We use data from several sources. First, for the 1983 to 2000 time period, we obtain stock-level measures of retail trading from the Institute for the Study of Security Markets (ISSM) and the Trade and Quote (TAQ) databases, where we use small-sized trades (trade size $\leq$ \$5,000) to proxy for retail trades. Like Barber, Odean, and Zhu (2008), we use the ISSM/TAQ data only until 2000. The introduction of decimalized trading in 2001 and order-splitting by institutions due to lower trading costs imply that trade size would not be an effective proxy for retail trading after 2000. We also obtain the portfolio holdings and trades of a sample of individual investors from a large U.S. brokerage house for the period from 1991 to 1996.[5]

Next, we obtain daily and monthly split-adjusted stock returns, stock prices, and shares outstanding for all traded firms from the Center for Research on Security Prices (CRSP). The book

[5] See Barber and Odean (2000) for additional details about the retail investor dataset and Barber, Odean, and Zhu (2008) or Hvidkjaer (2008) for additional details about the ISSM/TAQ dataset, including the procedure for identifying small trades.

value of equity and the book value of debt are obtained from COMPUSTAT. Following the related idiosyncratic volatility studies, we restrict the sample to firms with CRSP share codes 10 and 11. We obtain the monthly Fama-French factor returns and monthly risk-free rates from Kenneth French's data library.[6] Both the daily and the monthly data range from January 1983 to December 2000. For each stock, we also compute the book-to-market ratio using the book equity value from COMPUSTAT. Last, we obtain quarterly institutional ownership measures for all stocks using Thomson Financial's Institutional (13F) holdings data and obtain analyst coverage data from Thomson Financial's I/B/E/S data set.

III.B. The Retail Trading Proportion (RTP) Measure

Our main measure is the "retail trading proportion" (RTP). For each stock in each month t, we compute the stock's retail trading proportion as the ratio of the total month-t buy- and sell-initiated small trades (trade size below $5,000) dollar volume and the total stock trading dollar volume in the same month. Ideally, we would like to observe the trades of all retail investors but, unfortunately, such detailed retail trading data are not available for an extended time-period. Therefore, we use the buy- and sell-initiated small-trades as a proxy for retail trading. Several recent studies have adopted a similar identification strategy (e.g., Battalio and Mendenhall (2005), Malmendier and Shanthikumar (2007), Barber, Odean, and Zhu (2008), Hvidkjaer (2008)).

To ensure that our RTP variable reflects retail preferences, we compare RTP with actual retail holdings and trading data from a large U.S. discount brokerage house. Figure 1 shows the excess portfolio weight and the excess trading weight for RTP sorted portfolios. The excess weight reflects the difference between the actual portfolio weight in the aggregate retail investors' portfolio based on the brokerage data and the market portfolio constructed using all CRSP stocks. The sample period averages of the excess weights are shown in the figure. The excess trading weight is defined in an analogous manner using the total trading volume (sum of buy and sell volumes) measure. Figure 1 shows that both the portfolio and trading weights in the brokerage sample increase with the level of RTP. Retail investors in the discount brokerage house considerably overweight and trade more stocks that have higher RTP.

For greater accuracy, we also estimate Fama-MacBeth and cross-sectional regressions. The dependent variable in these regressions is RTP and the independent variables are the portfolio weight and trading weight obtained using the actual holdings and trades of retail investors at the

[6]The data library is available at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/`.

discount brokerage house. The regression results are reported in Table I. We find that both the portfolio weight and the trading weight variables are strongly positively correlated with the RTP measure. This evidence indicates that buy- and sell-initiated small trades volume is higher among stocks that are held and traded by the sample of brokerage retail investors. These comparisons with brokerage data for the 1991 to 1996 sub-period indicate that our RTP measure captures the stock preferences of retail investors.

III.C. Characteristics of RTP Sorted Portfolios

To further examine the ability of the RTP measure to capture retail preferences and to gain additional insights into the stock preferences of retail investors, we sort stocks into deciles based on their monthly RTP levels. Table II reports the mean stock characteristics of RTP sorted portfolios for the 1983 to 2000 time period. We also report the average RTP levels for the ten decile portfolios. The average RTP level ranges from less than 1% for the lowest RTP decile portfolio to over 60% for the highest RTP decile portfolio.

Consistent with our conjecture that RTP captures retail preferences, we find that stock's institutional ownership, market capitalization and stock price all decrease monotonically with RTP. For instance, stocks in the top three RTP decile all have average price below $10, and average market value below $100 million dollars. Together, the top five RTP deciles represent only less than 10% of the total stock market capitalization. The stocks in the highest RTP decile have an average institutional ownership of only 3.01%, with 57.72% of stocks having IO below 5%. In contrast, the average IO for the lowest RTP decile is 50.78%, and only 3.19% stocks in this decile have IO below 5%.

Although the level of IO declines monotonically across the RTP deciles, the magnitude of the correlation between RTP and IO is not very high. The average correlation between RTP and 1−IO is 0.193 when we compute the cross-sectional correlation each quarter and then take the average across all quarters. This correlation is even lower (only 0.049) when we first compute time-series correlations between RTP and 1−IO for each stock and then obtain the average. These comparisons indicate that RTP is not merely a transformation of 1−IO.[7]

Examining other stock characteristics of RTP sorted portfolios, we find that stocks with high fraction of retail trading have higher book-to-market ratios, lower analyst coverage and lower past

[7]See Hvidkjaer (2008) for additional comparisons between the ISSM/TAQ small-trades data and the 13F institutional holdings data. The key conclusion from his analysis is also that the ISSM/TAQ small-trades data do not merely proxy for 1−IO.

returns. In fact, the highest RTP decile stocks earn an average of −11.46% over the past twelve months, while the lowest RTP decile stocks earn an average of 31.62%. This evidence is consistent with retail investors being contrarians and their willingness to bet on stock price reversals.

We also find that the average idiosyncratic volatility (IVOL) and idiosyncratic skewness (ISKEW) increases monotonically across RTP portfolios. Similar to Ang, Hodrick, Xing, and Zhang (2006), we obtain month-t estimate of a stock's idiosyncratic volatility as the standard deviation of the residual obtained by fitting a four-factor model (Fama-French three factors plus a momentum factor) to its daily returns during month t. Idiosyncratic skewness is computed using the Harvey and Siddique (2000) method and is defined as the scaled measure of the third moment of the residual obtained by fitting a two-factor ($RMRF$ and $RMRF^2$) model to daily returns from previous six months. $RMRF$ is the market return, excess over the risk-free rate.

Table II shows that stocks in the highest RTP decile have an average IVOL of 41.51%, while those in the lowest RTP decile have an average IVOL of only 11.66%. Similarly, the average idiosyncratic skewness for the highest RTP decile portfolio is 0.745, which is almost twice the average idiosyncratic skewness of 0.386 for the lowest RTP decile portfolio. These estimates indicate that the amount of retail trading is greater among both high idiosyncratic volatility and high idiosyncratic skewness stocks, which are likely to be perceived as speculative stocks. In Section III.E, we provide additional evidence along multiple dimensions to demonstrate that the RTP measure indeed captures speculative retail preferences.

III.D. Volatility Preferences of Retail and Institutional Investors

In Table III, we directly examine whether retail investors over-weight high IVOL stocks in their portfolios and exhibit a greater propensity to trade those stocks. Each month, we construct IVOL sorted portfolios and compute the weights of those portfolios in the market portfolio constructed using all CRSP stocks. The averages of those monthly expected weights (if retail investors in aggregate hold the market portfolio) are reported in column (1) of Table III. In columns (2) and (3), we report the actual weights allocated to IVOL portfolios by brokerage investors in their portfolio holdings and trading activities. The trading weight is the ratio of the trading volume of stocks in the IVOL portfolio to the total volume of all trades by brokerage investors. In column (4), we obtain trading weights using the small-trades data from ISSM/TAQ. In columns (5) to (7), we report the excess weights measured as the actual weights in column (2) to (4) minus the expected market weights in column (1).

The sorting results indicate that retail investors exhibit a greater propensity to hold and trade high idiosyncratic volatility stocks. For instance, they under-weight the lowest IVOL decile portfolio by 18.50% and over-weight the highest IVOL decile portfolio by 4.37% (see column (5)). The excess weights from the brokerage data and the ISSM/TAQ data portray a similar picture.

To better quantify the heterogeneity in the volatility preferences of investors, we examine the idiosyncratic volatility preferences of retail and institutional investors separately. We estimate monthly Fama-MacBeth cross-sectional regressions of RTP on a set of stock characteristics, including idiosyncratic volatility. The results reported in Table IV indicate that RTP is significantly positively related to idiosyncratic volatility. In contrast, the excess weight of a stock in the aggregate portfolio of 13F institutional investors is significantly negatively related to idiosyncratic volatility. Thus, institutions as a group, under-weight high idiosyncratic volatility stocks, while retail investors over-weight those stocks.

In contrast to our focus on investors' idiosyncratic volatility preferences, previous studies examine institutional investors' preference for total volatility, but the results are inconclusive. Using a two-year sample of mutual funds (1991 and 1992), Falkenstein (1996) finds that mutual funds display a preference for high-volatility stocks. Sias (1996) finds that higher levels of institutional ownership for NYSE listed securities, over the period from 1977 to 1991 are associated with higher contemporaneous stock return volatility. However, using data for the 1980 to 1996 sample period, Gompers and Metrick (2001) do not detect a significant relation between institutional holdings and stock return volatility.

III.E. Volatility Preferences and Retail Speculation

Although we find strong evidence of retail preferences for idiosyncratic volatility, it is not clear why retail investors find high idiosyncratic volatility stocks attractive. One important reason could be that they perceive high IVOL stocks as instruments for speculative trading. If this conjecture is correct, the level of RTP should be high for other stocks that are also likely to be viewed as speculative instruments, including stocks with low prices, high idiosyncratic skewness levels, and non-dividend paying status. Further, RTP should be even higher for high IVOL stocks when they also possess these additional speculative characteristics.

To test this conjecture, we estimate monthly Fama-MacBeth regressions of RTP on a set of stock characteristics, including several measures that could capture the speculative content of a stock. In addition, we define interaction terms using price, idiosyncratic volatility, and idiosyncratic skewness measures to better characterize the speculative nature of stocks. The independent variables have been standardized to have a mean of zero and a standard deviation of one. This transformation allows us to compare the coefficient estimates directly within and across regression specifications.

The RTP regression estimates are reported in Table IV. We note that RTP exhibits persistence. The lagged RTP has a strong positive coefficient estimate.[8] Importantly, we find that RTP is considerably higher for stocks with higher IVOL levels and more positive idiosyncratic skewness. However, the coefficient estimate of ISKEW is only about one-tenth of the magnitude of IVOL. This evidence indicates that idiosyncratic volatility is a more important determinant of RTP than idiosyncratic skewness. Table IV estimates also show that low priced stocks tend to have high RTP. Furthermore, RTP is higher for non-dividend paying stocks, although the coefficient estimate of the dividend paying dummy becomes statistically insignificant when we account for other stock characteristics (see column (3)).

Examining the coefficients of the interaction terms, we find that High IVOL $\times$ Low Price and High IVOL $\times$ High ISKEW interaction terms have significantly positive coefficient estimates. Thus, stocks with higher IVOL have even higher RTP if they are both low priced and have high idiosyncratic skewness. Similarly, the significantly positive estimate of High ISKEW $\times$ Low Price interaction dummy indicates that high skewness stocks have higher levels of RTP if they also have low prices. The interaction term estimates are consistent with our conjecture and indicates that RTP levels are even higher for stocks with speculative characteristics.

To further test the idea that RTP captures the speculative preferences of retail investors, we examine the characteristics of the retail investor clientele of high RTP stocks. The hypothesis is that the clientele characteristics of high RTP stocks should be similar to the characteristics of investors who are more likely to engage in speculative or gambling-motivated trading (e.g., younger, male, less educated, and low-income investors), as documented in Kumar (2008).[9] Similar to Graham and Kumar (2006), using the trades of retail investors from a large U.S. brokerage house from 1991 to 1996, we measure the average characteristics of investors who trade the stock during the six-year sample period. We measure several stock-level investor characteristics such as age, income,

[8] We also examine the persistence of RTP using a transition matrix. We measure the percentage of stocks that belong to a specific RTP quintile in month t remain in the same quintile in the following month. We find that more than 70% of stocks classified in the lowest or the highest RTP quintile remain in their respective quintiles in the following month. In the middle three quintiles, about 45% of stocks retain their quintile memberships in two consecutive months. Even after three months, 65% and 40% of stocks in the extreme and middle quintiles retain their quintile memberships, respectively.

[9] It is very likely that investors with these characteristics are less likely to own stocks (e.g., Campbell (2006)). We argue that, conditional upon stock market participation, those investors are more likely to hold speculative stocks.

education level, gender, religion, race/ethnicity, and location. Using these clientele characteristics, we estimate a cross-sectional regression in which the sample-period average RTP for a stock is the dependent variable and the clientele characteristics of the stock are the independent variables. The results are reported in Table V. In column (1), we report estimates from a specification that only includes characteristics that are available in the brokerage data and in column (2) we consider additional characteristics defined using measures associated with investor's location.

We find that stocks with high levels of RTP have younger clienteles with lower income, lower education levels, and non-professional occupations. Those stocks also have a greater proportion of male and single investors and have relatively less diversified clienteles. Moreover, the RTP is high for stocks that are held by urban investors and those who reside in areas with higher per-capita lottery expenditures. Both these geographical characteristics are associated with a greater propensity to speculate and gamble. These demographic characteristics along with the religious and racial/ethnic characteristics of high RTP stocks are very similar to the characteristics of investors who are more attracted toward speculative and lottery-type stocks (Kumar (2008)). These cross-section regression results suggest that RTP is a good proxy for speculative retail preferences.

When we consider an alternative measure of retail trading that captures the direction of trading (i.e., the buy-sell imbalance or BSI), we find very different results (see Table V, column (3)).[10] Stocks with higher levels of BSI do not have clientele characteristics that are similar to the characteristics of investors who find speculative stocks attractive. The BSI regression estimates indicate that RTP rather than BSI is a more appropriate proxy for speculative trading. This evidence also indicates that speculative investors are not merely accumulating the shares of the stocks they like. Rather, they actively buy and sell those stocks and derive additional utility from the process of trading itself. For instance, speculative trading could be a source of entertainment (Dorn and Sengmueller (2006)) or provide extra utility to sensation seekers (Grinblatt and Keloharju (2008)). In sum, the level of trading rather than the direction of trading appears to be a more appropriate proxy for speculative retail preferences.

Taken together, the results in Tables IV and V indicate that retail investors are likely to actively trade high idiosyncratic volatility stocks due to their strong speculative preferences. Along multiple dimensions, we find that stocks with high RTP levels tend to be more speculative. These results further support our conjecture that RTP captures speculative preferences of retail investors.

[10]Like RTP, BSI is computed using the ISSM/TAQ data, where we use small-sized trades to proxy for retail trades. BSI is defined as $(B-S)/(B+S)$, where B is the total monthly buy-initiated small-trades volume and S is the total monthly sell-initiated small-trades volume measured in dollars.

III.F. Volatility Preferences and Realization Utility

Apart from their speculative motives, another reason why some retail investors might prefer high volatility stocks is that they derive additional utility from realizing gains on the stocks they own. Barberis and Xiong (2008a) present a model of portfolio choice with realization utility in which investors' propensity to realize gains (losses) is stronger (weaker) among stocks with higher idiosyncratic volatility. Because of these features, realization utility investors exhibit greater disposition effect among stocks with higher idiosyncratic volatility. Further, they suggest that the propensity to realize winners would be higher when there is more uncertainty about the true valuation of a stock.

In this section, we use the brokerage data to test these key theoretical predictions of the Barberis and Xiong (2008a) model. Specifically, we estimate pooled OLS regressions with year fixed effects, where the dependent variable in various specifications is one of the following measures: (i) proportion of gains realized (PGR), (ii) proportion of losses realized (PLR), and (iii) the ratio PGR/PLR. The three measures are computed for each stock at the end of each year using the portfolio holdings and trades of all brokerage investors. PGR is the proportion of gains realized and is defined as the ratio of the number of realized "winners" (stock positions where an investor experiences a gain) and the total number of winners (realized + paper). PLR is the proportion of losses realized and is defined in an analogous manner. Additional details on these measures are available in Odean (1998). The main independent variable of interest is the idiosyncratic volatility level of the stock. Several other stock characteristics are employed as control variables, and they are defined in the caption of Table VI. All independent variables are measured during year $t-1$.

The panel regression estimates are reported in Table VI. Consistent with the empirical predictions of the Barberis and Xiong (2008a) model, we find that there is a positive relation between idiosyncratic volatility and PGR (see columns (1) and (2)). Furthermore, consistent with their predictions, we find that investors' propensity to realize losses is lower for stocks with higher idiosyncratic volatility (see column (3)). When we use the PGR/PLR ratio measure (i.e., a measure of the stock-level disposition effect), we find stronger disposition effect for stocks with higher idiosyncratic volatility. These results support the empirical predictions of the Barberis and Xiong (2008a) model, which provides one mechanism for generating an affinity for volatility.

Our estimates in Table VI also indicate that investors' propensity to realize gains (losses) is higher (lower) for stocks with higher idiosyncratic skewness and, consequently, the disposition effect is stronger for those stocks. The idiosyncratic skewness coefficient estimates are qualitatively similar to the coefficient estimates of idiosyncratic volatility and further suggests that investors with speculative preferences exhibit a stronger desire to experience realization utility. The effects of idiosyncratic skewness on PGR, PLR and disposition effect are weaker in magnitude than the corresponding effects of idiosyncratic volatility. The relative magnitudes of the two coefficient estimates are consistent with our earlier finding that the level of speculative retail trading is more strongly related to idiosyncratic volatility than idiosyncratic skewness (see Table IV).

IV. Volatility Preferences and Volatility Premium

Our results so far show that some retail investors are attracted toward high idiosyncratic volatility stocks. This volatility preference reflects retail investors' tendency to engage in speculative trading, seek utility by realizing gains on the stocks they own, and derive sensation or entertainment value through trading of high volatility stocks. In this section, we study the potential pricing effects of retail investors' idiosyncratic volatility preferences.

IV.A. Main Testable Hypotheses

We test three related asset pricing hypotheses. First, we examine the relation between the retail trading proportion (RTP) of a stock and its future average return. If high RTP stocks have clienteles that derive additional non-wealth utility from holding and trading those stocks, all else equal, investors would be willing to pay a higher price for high RTP stocks. Subsequently, these stocks would earn lower returns. The greater the amount of speculative retail trading in a stock, the more overvalued it tends to be, and hence the lower would be the subsequent return. Of course, arbitrageurs would attempt to exploit the overvaluation of stocks with high retail trading proportion (e.g., using short positions). But, high arbitrage costs would limit the effectiveness of arbitrage activities and, consequently, the mispricing might not be completely eliminated and could persist. We examine empirically whether the RTP effect is stronger among stocks that are more costly to arbitrage.

These arguments lead to our first asset pricing hypothesis:

> **Hypothesis 1:** *Stocks with greater intensity of speculative retail trading, as reflected by high levels of retail trading proportion (RTP), earn lower average future returns.*

Furthermore, the negative RTP-return relation is stronger when arbitrage costs are higher.

Building upon our first hypothesis, we examine the relation between idiosyncratic volatility and future returns, conditional upon the level of retail trading. We have shown that RTP is positively related to idiosyncratic volatility (Table IV). If RTP is a negative predictor of future returns (Hypothesis 1), then it is important to account for the RTP effect in studying the volatility-return relation. Our conjecture is that the relation between idiosyncratic volatility and stock returns depends crucially on the existence of volatility-seeking retail clienteles. More precisely, we posit the following relation:

> **Hypothesis 2:** *For stocks with high retail trading proportion, future returns decrease with idiosyncratic volatility. In contrast, for stocks with low retail trading proportion, volatility-return relation is insignificant or future returns increase with idiosyncratic volatility.*

In our third hypothesis, we examine the retail trading proportion-return relation and the idiosyncratic volatility-return relation, conditional upon various speculative stock characteristics and arbitrage cost proxies. We conjecture that among stocks with strong speculative characteristics (e.g., high idiosyncratic skewness), RTP is more likely to reflect speculative retail trading. If RTP negatively predicts future stock returns because of speculative retail investors' willingness to pay a premium to hold and trade stocks they like, then the RTP-return relation should be stronger (i.e., more negative) among stocks with stronger speculative characteristics, especially if those stocks are also more difficult to arbitrage.

Further, we have shown that retail investors prefer volatility more when idiosyncratic skewness is higher (columns (2) and (3) of Table IV). Thus, if the negative volatility-return relation is induced by volatility-seeking speculation, it should be stronger among stocks with stronger speculative characteristics when we do not account for the influences of speculative retail trading.

To summarize, our third hypothesis posits that:

> **Hypothesis 3:** *The negative retail trading proportion-return relation is stronger among stocks with stronger speculative characteristics. Furthermore, higher idiosyncratic volatility stocks earn even lower future returns when they have strong speculative characteristics. Both relations are stronger among stocks with higher arbitrage costs.*

IV.B. RTP and Average Returns: Univariate Sorts

To begin, we test the first hypothesis using a portfolio-based approach. At the end of each month, we form RTP quintile portfolios and compute their equal-weighted and value-weighted returns. Table VII reports the characteristics and performance of those RTP sorted portfolios for January 1983 to December 2000 sample period. In Panel A, we report the main performance estimates and, for robustness, in Panel B we present the performance estimates for different sub-periods and sub-samples.

The sorting results indicate that the RTP-return relation is economically significant. For instance, the lowest RTP quintile earns a value-weighted mean monthly return of 1.765%, while the highest RTP quintile earns a large negative monthly return of −3.231%. There is also a large negative spread between the equal-weighted average returns of high and low RTP quintiles. The characteristic-adjusted performance estimates or the four-factor (Fama-French three factors plus a momentum factor) alphas portray a very similar picture.

The risk-adjusted performance estimates in Table VII, Panel B indicate that the RTP-return relation is robust. Excluding stocks with price below $5 or restricting the analysis to stocks with low arbitrage costs (idiosyncratic volatility, our proxy for arbitrage costs, is in the bottom three deciles) or high institutional ownership (IO in the highest three deciles) weakens the result. Nevertheless, high RTP stocks still significantly under-perform low RTP stocks in all these sub-samples. The results are also similar when we restrict the sample to the first half of the sample period (1983 to 1991) or exclude January returns.[11] We also find that the profits of RTP-based trading strategies are not limited to a few periods. Figure 3 plots the raw monthly return difference between the low RTP and high RTP portfolios and the 12-month moving average of the monthly return differentials. The low RTP stocks outperform high RTP stocks consistently, as the return differential is positive for most of the months during our sample. In particular, our results are not driven by the internet bubble period of the late 1990s.

The economic magnitudes of the abnormal returns of some RTP portfolios are large. For instance, the high RTP portfolio earns a mean monthly risk-adjusted return of −4.095%.[12] However, it is difficult to trade on this finding, because stocks in the high RTP quintile have very low market

[11] Given the full-sample (1983-2000) results and the 1983-1991 sub-period estimates, it is clear that the results would be similar and somewhat stronger for the 1992-2000 period. For brevity, we do not report those results.

[12] AHXZ also report instances of such extreme portfolio returns. For example, the monthly three-factor alpha of high idiosyncratic volatility quintile portfolio is −2.66% if those stocks have performed poorly over the past 12 months.

capitalizations (see Table II), and face high transaction costs and short sales constraints.

Table VII Panel C reports the performance estimates of RTP-based trading strategies that could be potentially realized. We exclude all stocks that are priced below $10 and restrict the stock sample based on short-sales constraints and arbitrage costs. We sort stocks into two groups (top half and bottom half) based on their retail trading proportion to ensure that each group captures a meaningful fraction of the market. The results indicate that low RTP stocks still earn significantly higher returns than high RTP stocks. The raw (risk-adjusted) performance differential is 1.814% (1.673%) per month when we only use the "Minimum Stock Price = $10" filter. When we restrict the sample further and consider stocks that are shortable,[13] the raw (risk-adjusted) performance differential reduces to 1.192% (1.051%) per month but remain economically significant. These estimates are similar when we consider stocks that have lower arbitrage costs (idiosyncratic volatility in the three lowest deciles).

Taken together, the RTP sorting results provide strong support to Hypothesis 1. Stock's retail trading proportion is negatively related to average future stock return. This relation is stronger among stocks that are more difficult to arbitrage. When arbitrage costs are low, stocks with low RTP still significantly outperform high RTP stocks by over 1% per month.

IV.C. RTP and IVOL Double Sorts

In our second set of tests, we examine the performance of RTP-IVOL double sorted portfolios to gather support for our second hypothesis. The evidence also provides additional support for the first hypothesis. Table VIII, Panel A reports the average returns of IVOL sorted portfolios as well as the portfolios double-sorted along RTP and IVOL dimensions. Table VIII, Panel B reports the excess returns of the same set of portfolios relative to Fama-French three factors plus a momentum factor. The results are similar when we examine raw or risk-adjusted portfolio returns. Hence, for brevity, we focus the discussion primarily on the raw returns reported in Table VIII, Panel A.

The first column of Table VIII, Panel A shows that when we sort stocks using IVOL only, there is a negative relation between idiosyncratic volatility and the average monthly returns in the following month. The lowest IVOL quintile earns an average monthly return of 1.39%, while the highest IVOL quintile earns an average monthly return of −0.16%. These results are consistent with the AHXZ evidence.

[13]Stocks with any short-interest are identified as shortable. See Purnanandam and Seyhun (2007) for details on the short-interest data we use, which are available only for the 1991 to 2000 period.

More interestingly, when we examine the returns of RTP-IVOL double-sorted portfolios, we find that average returns decline monotonically with RTP. Within each IVOL quintile, consistent with Hypothesis 1, high RTP stocks on average earn significantly lower returns than the low RTP stocks. For instance, when IVOL is low (quintile 1), the low and the high RTP portfolios earn monthly returns of 1.66% and -1.14%, respectively. The difference of -2.80% is statistically significant (t-statistic $= -15.24$). This evidence indicates that the ability of retail trading proportion to predict the returns in the following month is not somehow mechanically induced by the negative volatility-return relation highlighted in Ang, Hodrick, Xing, and Zhang (2006, 2008).

Examining the RTP-return relation across IVOL sorted portfolios, we find that the RTP effect is stronger among stocks with higher idiosyncratic volatility. For example, the high RTP stocks on average under-perform low RTP stocks by 2.80% per month in the lowest IVOL quintile, but by 7.64% in the highest IVOL quintile. To the extent that arbitrage activities are less effective among stocks with high idiosyncratic volatility (e.g., Pontiff (1996), Shleifer and Vishny (1997), Wurgler and Zhuravskaya (2002)), this result is consistent with stock mis-pricing underlying the RTP-return relation.

Furthermore, consistent with Hypothesis 2, the double sorting results indicate that the volatility-return relation changes sign across the extreme RTP categories. We find that the negative volatility-return relation is concentrated in higher RTP quintiles. In contrast, in lower RTP quintiles, there is a positive volatility-return relation. For stocks with intermediate level of retail trading, future stock returns are not significantly related to idiosyncratic volatility. Taken together, the results from RTP and IVOL based double sorts indicate that the level of retail trading proportion is an important determinant of the idiosyncratic volatility-return relation.

To graphically illustrate the relation between RTP, idiosyncratic volatility, and average returns, in Figure 2, we plot the idiosyncratic volatility premium (the difference between the value-weighted returns of high and low idiosyncratic volatility portfolios) for the ten RTP decile portfolios. We find that the idiosyncratic volatility premium is positive in the first five (low level) RTP deciles and these five categories represent a large proportion (over 90%) of the market (see Table I, column (1)). In only a relatively small (less than 10%) segment of the market with high levels of retail trading, the idiosyncratic volatility premium is insignificant or significantly negative. This evidence indicates that the intensity of speculative retail trading, as captured by RTP, strongly influences the idiosyncratic volatility-return relation.

IV.D. Robustness of RTP-IVOL Double Sorting Results

We conduct several tests to examine the robustness of the RTP-IVOL double sort results and report the results in Table VIII Panel C. For brevity, we only report the performance estimates for the High IVOL − Low IVOL portfolio. In the first robustness check, to ensure our results are not driven entirely by very small, low priced, hard-to-arbitrage stocks, we exclude stocks with price below $5. We still find that high IVOL stocks significantly out-perform low IVOL stocks when RTP levels are low, and the pattern reverses when RTP levels are high.

Kaniel, Saar, and Titman (2008) propose that risk-averse individual investors provide liquidity to institutions that demand immediacy. Thus, prices fall as institutions sell to individuals one week and rebound the next. Like Barber, Odean, and Zhu (2008), we use only trades initiated by retail investors and, thus, our measure is designed to exclude limit orders. It implies that losses of speculative retail investors are not due to the limit orders of individual investors being opportunistically picked off by institutional investors. Nevertheless, to control for the possibility suggested in Kaniel, Saar, and Titman (2008), in our second robustness test, we skip a month and re-estimate the returns of RTP-IVOL double sorted portfolios. Again, the results do not change materially relative to the full-sample results reported in the last row of Table VIII Panel B.

In the next two tests, we consider only larger stocks and stocks that can be shorted. Even when we exclude all NASDAQ stocks from the sample or only consider stocks that have short interest, we find that high IVOL stocks significantly outperform low IVOL stocks when RTP levels are low, and the pattern reverses when RTP levels are high. The last two robustness tests show that the same patterns hold in subperiods or after excluding January returns.

To examine whether the RTP-conditional idiosyncratic volatility spread are potentially attainable, we re-estimate the performance differentials of IVOL-sorted portfolios for stock sub-samples in which arbitrage costs are likely to be lower and micro-structure effects should not materially affect the trading profits. Specifically, we only consider stocks that are priced at $10 or above and form four RTP-IVOL sorted portfolios ensuring each portfolio has a significant market capitalization. After conditioning upon the level of RTP, we obtain the monthly average returns for low (bottom half) and high (top half) IVOL portfolios. We only report the results for low (bottom third or bottom half) RTP stocks because although the trading profits for high RTP portfolios look big on paper (see Panel B), they are unlikely to be realized due to various trading frictions.

The results are reported in Table VIII Panel D. We find that among the bottom third RTP stocks, those with high IVOL earns an average monthly return of 2.622%, while those with low

IVOL earns an average monthly return of 1.635%. The 0.987% difference is economically significant. The four-factor alpha for the difference portfolio is slightly higher (= 1.022%). Even among the bottom half RTP stocks, the monthly raw and the risk-adjusted return difference between the high IVOL and the low IVOL portfolios is 0.664% and 0.758% respectively, which are still economically significant.

Overall, we find that our results that support Hypothesis 2 are robust and economically significant. Below, we present further evidence on the robustness of our results in a multivariate framework using Fama-MacBeth cross-sectional regressions.

IV.E. Fama-MacBeth Cross-Sectional Regression Estimates

To accurately characterize the influence of retail volatility clienteles on the volatility-return relation, we follow the Fama-MacBeth regressions approach in Ang, Hodrick, Xing, and Zhang (2008). The dependent variable in these regressions is the monthly stock return. The set of independent variables includes the variables used in Ang, Hodrick, Xing, and Zhang (2008) (e.g., stock characteristics and factor exposures) as well as the retail trading proportion variable.

The Fama-MacBeth regression results are reported in Table IX. In column (1), we present the results for the 1980 to 2003 period to show that our results are very similar to the evidence in Ang, Hodrick, Xing, and Zhang (2008) when we use the exact specification and the exact time-period used in that study. Consistent with their evidence, we find that idiosyncratic volatility has a strong negative coefficient estimate (coefficient estimate = -0.491, t-statistic = -4.35). Even when we re-estimate the coefficients for the sample and the time-period for which the retail trading data are available (1983 to 2000), we find very similar results (see column (2)). In particular, IVOL maintains its strong negative coefficient estimate (coefficient estimate = -0.479, t-statistic = -4.59).

Table IX, column (3) reports the regression estimates when we include the our key variable (i.e., RTP) as an additional explanatory variable. We find that the volatility-return relation changes dramatically in the presence of RTP. The IVOL coefficient estimate now becomes significantly positive (coefficient estimate = 0.393, t-statistic = 3.69). This evidence highlights the critical role of speculative retail trading in shaping the idiosyncratic volatility-return relation.

We also find that RTP has a strong negative coefficient estimate (estimate = -1.541, t-statistic = -10.69). Thus, stocks with a larger proportion of retail trading in a month earn lower returns in the following month. These results are very similar when we estimate the regression with the total volatility measure (see column (4)). The negative coefficient estimate for RTP in these predictive regressions again confirms Hypothesis 1.

To further investigate how RTP influences the volatility-return relation, we estimate the Fama-MacBeth regressions with High RTP $\times$ High IVOL and Low RTP $\times$ High IVOL interaction terms in the specification (see column (5)). The "high" and the "low" categories are defined as the top and the bottom RTP quintiles, respectively. We find that, all else equal, high IVOL stocks earn -0.469% lower monthly returns (t-statistic $= -8.80$) when those stocks also have high levels of retail trading. In contrast, when high IVOL stocks have low levels of retail trading, they earn 0.247% higher monthly returns (t-statistic $= 6.59$). These estimates are consistent with the predictions of Hypothesis 2 and indicate that the idiosyncratic volatility-return relation depends critically upon the proportion of retail trading in the stock.

For robustness, in columns (6) and (7), we report the Fama-MacBeth regression estimates for sub-sample of stocks with low (bottom one-third) and high (top one-third) RTP, respectively. Consistent with the signs of the estimates of the High RTP $\times$ High IVOL and Low RTP $\times$ High IVOL interaction terms in column (5), we find that the idiosyncratic volatility-return relation is negative in the high RTP sub-sample, but positive in the low RTP sub-sample. These results provide additional support for Hypothesis 2.

IV.F. Results in Skewness Sub-Samples

In the next set of tests, we gather support for our third hypothesis, which posits that the volatility-return relation would be stronger among stocks with stronger speculative characteristics. Motivated by the arguments in Barberis and Huang (2008), we use idiosyncratic skewness as the key indicator of speculative stock characteristic. We regress monthly stock returns on the lagged idiosyncratic volatility and retail trading proportion for five idiosyncratic skewness (ISKEW) based sub-samples. Given our asset pricing hypotheses, the magnitudes of both IVOL and RTP coefficient estimates are expected be stronger in higher ISKEW quintiles.

The estimation results are summarized in Table X. We find that the coefficient estimate of IVOL declines and becomes more negative as we move from low to high skewness quintiles. In the lowest skewness quintile, the IVOL coefficient estimate is -0.089 and is statistically insignificant (t-statistic $= -0.83$). In contrast, the IVOL coefficient estimate is strongly negative in the highest skewness quintile (coefficient estimate $= -0.588$, t-statistic $= -4.77$). This evidence supports Hypothesis 3 and indicates that the negative volatility-return relation is stronger among speculative stocks with higher skewness.

When we include RTP in the regression specifications, we find that the IVOL coefficient estimate is significantly positive in all five sub-samples. Furthermore, we find that the coefficient estimate of RTP is significantly negative and becomes more negative as we move from low to high skewness quintiles. For example, in the lowest ISKEW quintile, RTP has a coefficient estimate of -0.878 (t-statistic $= -5.94$), but in the highest ISKEW quintile, it is significantly stronger (coefficient estimate $= -1.781$, t-statistic $= -9.53$). Thus, consistent with Hypothesis 3, the evidence indicates that the level of retail trading has a stronger influence on average returns of speculative stocks that have higher skewness.

The RTP and IVOL coefficient estimates in ISKEW sub-samples become stronger when we restrict the sample to lower priced stocks (Price $<$ \$10). These lower priced stocks are likely to be have higher levels of speculative trading and higher arbitrage costs. The stronger coefficient estimates reflect the combined effects of greater speculation and higher arbitrage costs. Again, the evidence is consistent with Hypothesis 3.

V. Robustness Checks and Alternative Explanations

In this section, we present results from several Fama-MacBeth regressions designed to check the robustness of the baseline results reported in Table IX. The robustness test results are summarized in Table XI.

V.A. Alternative Estimation Period

In the first robustness test, unlike all our previous analyses, we estimate both the idiosyncratic volatility and the retail trading proportion variables using the returns and trading data over the previous three months instead of only one month. We find that the Fama-MacBeth regression estimates are similar to the baseline estimates (compare Table XI, column (1) to Table IX, column (3)).

V.B. BSI: An Alternative Measure of Retail Trading

In the second robustness test, we examine whether an alternative retail trading measure, the monthly buy-sell imbalance, yields similar results as the retail trading proportion variable. Barber, Odean, and Zhu (2008) use the same data as our study to examine whether persistent buying or selling by individual investors affect stock prices and forecast future stock returns.[14] They construct a buy-sell imbalance (BSI) measure, which is defined as the ratio of (Buy-initiated small-trades volume − Sell-initiated small-trades volume) and total small-trades volume. BSI is related to our RTP measure, but there are significant differences. A stock can have low level of retail trading, but the retail trades could be mostly on the same side, which implies a large (either positive or negative) buy-sell imbalance. There can also be lots of retail trading in a stock (high RTP) which is quite balanced on the buy and sell sides (low buy-sell imbalance). The average cross-sectional correlation between RTP and BSI is only 0.082.

Empirically, RTP and BSI behave very differently in terms of their ability to forecast stock returns. Both Kaniel, Saar, and Titman (2008) and Barber, Odean, and Zhu (2008) show that stocks heavily bought by individuals one week reliably outperform the market the following week. In contrast, we find RTP negatively forecasts stock return over the next month.

When we replace RTP with BSI in the regression specification, we find that BSI has a significantly positive coefficient estimate (see column (2) of Table XI Panel A). This evidence is consistent with the findings in Barber, Odean, and Zhu (2008) and indicates that stocks with greater retail buying interest earn higher returns in the following month. More importantly, we find that the IVOL coefficient estimate remains significantly negative with little change in magnitude when BSI is included as a control in the return-idiosyncratic volatility regression. When we include both RTP and BSI as regressors (see column (6) of Table XI Panel A), the RTP effect is little changed in terms of both economic and statistical significance of forecasting stock returns. Overall, the effects of BSI and RTP on returns are largely independent.

These results with the BSI measure are very different from our previous results with the RTP measure. The differences are evident in the relation between RTP and average future returns and also in the manner in which RTP affects the idiosyncratic volatility-return relation. Thus, the level of retail trading that reflects retail investors' speculative preferences rather than the trading imbalance is the critical determinant of the volatility-return relation.

[14]Other related studies include Hvidkjaer (2008), Dorn, Huberman, and Sengmueller (2008), and Kaniel, Saar, and Titman (2008). Kaniel, Saar, and Titman (2008) use only NYSE trades that includes limit orders.

V.C. Short-Term Return Reversals

The results in Tables II and IV indicate that RTP is strongly negatively related to past stock returns. To examine the possibility that RTP captures the short-term return reversal effect (e.g., Jegadeesh (1990), Lehmann (1990)), we replace RTP with past one-month return in the regression specification. We find that the previous month return has a strong negative coefficient estimate and it weakens the negative IVOL estimate (see column (3) of Table XI Panel A). Nevertheless, the IVOL estimate is still significantly negative (coefficient estimate = -0.366, t-statistic = -3.28). This result indicates that short-term return reversal is partially behind the puzzling negative idiosyncratic volatility-return relation, but it does not provide a complete answer to the puzzle. In addition, column (6) of Table XI, Panel A shows that the predictive power of RTP for future returns persists with little change in economic magnitude when we explicitly control for the past one-month return.[15]

V.D. Role of Institutional Ownership

Thus far we have interpreted RTP as capturing the amount of speculative trading by retail investors. Another possible explanation for why RTP negatively forecasts future returns of stocks is that institutions are more informed and better at identifying stocks that would do well. So stocks they shun away from (which would have high RTP) have lower average future returns. To entertain this alternative interpretation, we replace RTP with institutional ownership level and institutional change measures in the regression specification (see Table XI, column (4)). We find that both institutional measures have insignificant coefficient estimates and IVOL continues to have a significantly negative coefficient estimate. Thus, our baseline results cannot be explained by the stock preferences or trading activities of institutional investors.

V.E. Idiosyncratic Skewness

In column (5) of Table XI, we control for idiosyncratic skewness because it is related to both RTP and IVOL. Theoretical models have been proposed in which idiosyncratic skewness is negatively related to returns (e.g., Mitton and Vorkink (2007), Barberis and Huang (2008), Brunnermeier,

[15]Huang, Liu, Ghee, and Zhang (2007) use a different specification and different sample period (1963 to 2004) and show that the volatility-return is insignificant when past one-month return is used as a control variable.

Gollier and Parker (2007)). While we do find that the ISKEW coefficient estimate in the return regression is significantly negative, it does not change the sign, the magnitude, or the statistical significance of the IVOL coefficient estimate. This evidence indicates that the idiosyncratic volatility and the idiosyncratic skewness effects on stock returns are largely distinct phenomena.

In the last specification (see column (6)), we include all five new variables in the regression specification along with the RTP measure. We find that the magnitudes of the RTP and IVOL coefficients are not materially affected by these additional controls. Like our previous evidence, we find that RTP is significantly negatively related to future stock returns, and the IVOL coefficient estimate is significantly positive in the presence of RTP.

V.F. Effect of Leverage

Johnson (2004) argues that the negative relation between idiosyncratic volatility and future average returns would be stronger when leverage is high. Although the evidence in AHXZ is inconsistent with this conjecture, we examine whether the strong RTP effects we document are related to leverage. We estimate the Fama-MacBeth return regressions for five leverage quintiles, where following Johnson (2004), we define leverage as book value of debt divided by the book value of debt plus market value of equity.

The results are reported in Table XI, Panel B. We find that in all five leverage sub-samples, RTP has a strong negative coefficient estimate and IVOL has a significantly positive coefficient estimate. These estimates are very similar to the full sample results shown earlier. Most notably, in the presence of RTP, even for the high leverage category, there is a strong positive relation between IVOL and returns. There is little variations across different leverage subsamples in the coefficients of RTP and IVOL. Overall, the leverage-based sub-sample results indicate that the RTP measure does not capture effects that can be attributed to leverage.[16]

In unreported tests, similar to the evidence in AHXZ, we find that the leverage variable has an insignificant negative coefficient estimate, while the leverage-idiosyncratic volatility interaction term has a positive but statistically insignificant coefficient estimate. More importantly, inclusion of these two additional variables does not materially change the coefficient estimates of RTP and IVOL.

[16]We also experiment with a direct measure of distress risk. Using the Shumway (2001) default risk measure, we find that the coefficient estimates of IVOL and RTP are not significantly influenced when we use the default probability measure as an additional control variable in our Fama-MacBeth regressions.

V.G. Effect of Firm Age

We also experiment with other regression specifications that include firm age (the time since the firm first appeared in the CRSP data) or recent IPO dummy (set to one if the firm went public within the past one year). In untabulated results, we find that the relation between future returns and RTP as well as IVOL get amplified among "younger" firms, where the RTP-return relation is affected to a greater degree.

The IVOL $\times$ Young Firm interaction term has a significantly negative coefficient estimate (estimate $= -0.172$, t-stat $= -3.04$). The RTP $\times$ Young Firm interaction term also has a significantly negative coefficient estimate (estimate $= -0.445$, t-stat $= -8.43$). When both interaction terms are included in the same specification, the sign of IVOL $\times$ Young Firm interaction term turns positive (estimate $= 0.125$, t-stat $= 2.08$), while the RTP $\times$ Young Firm interaction term maintains its negative coefficient estimate (estimate $= -0.474$, t-stat $= -8.24$) Overall, the evidence indicates that RTP influences the IVOL-return relation for both "young" and "mature" firms, although the RTP effect is stronger among younger firms.

V.H. Extended Sample Period

Although the RTP measure would be a more noisy proxy for retail trading beyond 2000, in untabulated results we find that the main results are qualitatively very similar but weaker when we extend the sample period to 2005. Both the full sample results for the 1983 to 2005 period and the 2001-05 sub-sample results are very similar to the reported results.[17]

Collectively, the robustness test results summarized in Table XI indicate that the baseline results in Table IX are stable. Both the strong negative RTP-return relation and the strong influence of RTP on the strength and the direction of the idiosyncratic volatility-return relation appear novel and robust.

VI. Summary and Conclusion

This paper studies the puzzling negative relation between idiosyncratic volatility and expected returns identified in Ang, Hodrick, Xing, and Zhang (2006, 2008). The main finding is that the relation between idiosyncratic volatility and expected returns crucially depends on the existence

[17]We thank Paul Koch for providing the small-trades data from TAQ for the 2001-05 period.

of investor clienteles induced by heterogeneity in their volatility preferences. The speculative preferences of volatility-seeking retail clienteles can successfully explain the puzzling negative relation between idiosyncratic volatility and expected returns.

The first part of the paper documents that institutional investors underweight high idiosyncratic volatility stocks, while retail investors exhibit a strong preference for high idiosyncratic volatility stocks and other stocks with strong speculative characteristics. Due to their volatility preferences, retail investors hold and actively trade high idiosyncratic volatility stocks. Retail investors also actively trade speculative stocks because those stocks offer them greater opportunities for experiencing higher levels of realization utility (Barberis and Xiong (2008a)). Stocks whose trading are dominated by retail investors have strong speculative characteristics, and the characteristics of the retail clienteles of these stocks are remarkably similar to the characteristics of investors who are attracted to speculative and lottery-type stocks (Kumar (2008)).

The second part of the paper tests the pricing impact of retail investors' speculative preferences. In particular, we examine whether speculative retail preferences influence the relation between idiosyncratic volatility and expected returns. We hypothesize that investors attracted toward high volatility stocks and other stocks with strong speculative characteristics are willing to pay a higher price and require lower average returns for investing in those stocks. Thus, stocks with high retail trading proportion tend to be overvalued and have lower subsequent returns. Our empirical findings strongly support this hypothesis. Several alternative hypotheses we consider cannot explain the robust and novel finding that high retail trading proportion predicts lower future stock returns.

The asset-pricing tests also indicate that the "noise" generated by speculative retail investors distorts the volatility-return relation. We show that the puzzling negative idiosyncratic volatility-return relation is concentrated in the segment of the market that is dominated by retail investors and among stocks that have strong speculative characteristics. Among stocks with low retail trading proportion, future average stock returns increase with idiosyncratic volatility.

Taken together, these results indicate that our speculative retail trading proxy is a strong predictor of future returns and provides an intuitive explanation for the idiosyncratic volatility puzzle. Future research may find it fruitful to examine whether our speculative retail trading measure can successfully explain other related asset pricing puzzles such as the negative relation between distress risk and stock returns.

References

Allen, F., A. E. Bernardo, and I. Welch, 2000, "A Theory of Dividends Based on Tax Clienteles," *Journal of Finance*, 55, 2499–2536.

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman, 2004, "Why Constrain Your Mutual Fund Manager?," *Journal of Financial Economics*, 73, 289–321.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259–299.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2008, "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence," *Journal of Financial Economics*, Forthcoming.

Badrinath, S. G., G. D. Gay, and J. R. Kale, 1989, "Patterns of Institutional Investment, Prudence, and the Managerial "Safety-Net" Hypothesis," *Journal of Risk and Insurance*, 56, 605–629.

Bali, T. G., and N. Cakici, 2008, "Idiosyncratic Volatility and the Cross-Section of Expected Returns," *Journal of Financial and Quantitative Analysis*, 43, 29–58.

Bali, T. G., N. Cakici, X. Yan, and Z. Zhang, 2005, "Does Idiosyncratic Risk Really Matter?," *Journal of Finance*, 60, 905–929.

Barber, B. M., and T. Odean, 2000, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55, 773–806.

Barber, B. M., T. Odean, and N. Zhu, 2008, "Do Noise Traders Move Markets?," *Review of Financial Studies*, Forthcoming.

Barberis, N., and M. Huang, 2008, "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, Forthcoming.

Barberis, N., and W. Xiong, 2008a, "Realization Utility," Working Paper (February), Princeton University and Yale School of Management.

Barberis, N., and W. Xiong, 2008b, "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation," *Journal of Finance*, Forthcoming.

Battalio, R. H., and R. R. Mendenhall, 2005, "Earnings Expectations, Investor Trade Size, and Anomalous Returns Around Earnings Announcements," *Journal of Financial Economics*, 77, 289–319.

Boyer, B., T. Mitton, and K. Vorkink, 2008, "Expected Idiosyncratic Skewness," Working Paper (January), Brigham Young University.

Brav, A., M. Brandt, J. Graham, and A. Kumar, 2008, "The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes?," Working Paper (March), Duke University and University of Texas at Austin.

Brunnermeier, M. K., C. Gollier, and J. A. Parker, 2007, "Optimal Beliefs, Asset Prices and the Preference for Skewed Returns," *American Economic Review*, 97, 159-165.

Campbell, J. Y., 2006, "Household Finance," *Journal of Finance*, 61, 1553–1604.

Chua, C. T., J. Goh, and Z. Zhang, 2007, "Expected Volatility, Unexpected Volatility, and the Cross-section of Stock Returns," Working Paper (December), Singapore Management University.

Clotfelter, C. T., and P. J. Cook, 1989, *Selling Hope: State Lotteries in America*, Harvard University Press, Cambridge, MA.

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho, 2002, "Who Underreacts to Cash-Flow News? Evidence from Trading between Individuals and Institutions," *Journal of Financial Economics*, 66, 409–462.

Daniel, K. D., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," *Journal of Finance*, 52, 1035–1058.

Del Guercio, D., 1996, "The Distorting Effect of the Prudent Man Law on Institutional Equity Investments," *Journal of Financial Economics*, 40, 31–62.

Diavatopoulos, D., J. S. Doran, and D. R. Peterson, 2006, "The Information Content in Implied Idiosyncratic Volatility and the Cross-Section of Stock Returns: Evidence from the Option Markets," Working Paper (April), Florida State University.

Doran, J. S., D. Jiang, and D. R. Peterson, 2008, "Gambling in the New Year? The January Idiosyncratic Volatility Puzzle," Working Paper (March), Florida State University.

Dorn, D., and G. Huberman, 2007, "Preferred Risk Habitat of Individual Investors," Working Paper (October), Columbia University and Drexel University.

Dorn, D., G. Huberman, and P. Sengmueller, 2008, "Correlated Trading and Returns," *Journal of Finance*, 63, 885–920.

Dorn, D., and P. Sengmueller, 2006, "Trading as Entertainment," Working Paper (October), Drexel University.

Falkenstein, E. G., 1996, "Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings," *Journal of Finance*, 51, 111–135.

Feng, L., and M. S. Seasholes, 2005, "Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?," *Review of Finance*, 9, 305–351.

Frieder, L., and G. J. Jiang, 2007, "Separating Up From Down: New Evidence on the Idiosyncratic Volatility - Return Relation," Working Paper (March), University of Arizona.

Fu, F., 2008, "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*, Forthcoming.

Gompers, P. A., and A. Metrick, 2001, "Institutional Investors and Equity Prices," *Quarterly Journal of Economics*, 116, 229–259.

Goyal, A., and P. Santa-Clara, 2003, "Idiosyncratic Risk Matters," *Journal of Finance*, 58, 975–1007.

Graham, J. R., and A. Kumar, 2006, "Do Dividend Clienteles Exist? Evidence on Dividend Preferences of Retail Investors," *Journal of Finance*, 61, 1305–1336.

Grinblatt, M., and M. Keloharju, 2008, "Sensation Seeking, Overconfidence, and Trading Activity," *Journal of Finance*, Forthcoming.

Harvey, C. R., and A. Siddique, 2000, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263–1295.

Huang, W., Q. Liu, S. R. Ghee, and L. Zhang, 2007, "Return Reversals, Idiosyncratic Risk and Expected Returns," Working Paper (February), University of Hawaii at Manoa.

Hvidkjaer, S., 2008, "Small Trades and the Cross-section of Stock Returns," *Review of Financial Studies*, Forthcoming.

Jegadeesh, N., 1990, "Evidence of Predictable Behavior of Security Returns," *Journal of Finance*, 45, 881–898.

Jiang, G. J., D. Xu, and T. Yao, 2008, "The Information Content of Idiosyncratic Volatility," *Journal of Financial and Quantitative Analysis*, Forthcoming.

Johnson, T. C., 2004, "Forecast Dispersion and the Cross Section of Expected Returns," *Journal of Finance*, 59, 1957–1978.

Jones, C., and M. Rhodes-Kropf, 2004, "The Price of Diversifiable Risk in Venture Capital and Private Equity," Working Paper (July), Columbia University.

Kaniel, R., G. Saar, and S. Titman, 2008, "Individual Investor Trading and Stock Returns," *Journal of Finance*, 63, 273–310.

Kapadia, N., 2007, "The Next Microsoft? Skewness, Idiosyncratic Volatility, and Expected Returns," Working Paper (November), Rice University.

Kumar, A., 2008, "Who Gambles in the Stock Market?," Working Paper (April), McCombs School of Business, University of Texas at Austin.

Lehmann, B. N., 1990, "Fads, Martingales, and Market Efficiency," *Quarterly Journal of Economics*, 105, 1–28.

Lintner, J., 1965, "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47, 13–37.

Litzenberger, R. H., and K. Ramaswamy, 1979, "Dividends, Short Selling Restrictions, Tax-Induced Investor Clienteles and Market Equilibrium," *Journal of Finance*, 35, 469–482.

Malkiel, B. G., and Y. Xu, 2006, "Idiosyncratic Risk and Security Returns," Working Paper (March), School of Management, University of Texas at Dallas.

Malmendier, U., and D. M. Shanthikumar, 2007, "Are Small Investors Naive About Incentives?," *Journal of Financial Economics*, 85, 457–489.

Merton, R. C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42, 483–510.

Mitton, T., and K. Vorkink, 2007, "Equilibrium Under-diversification and the Preference for Skewness," *Review of Financial Studies*, 20, 1255–1288.

Odean, T., 1998, "Are Investors Reluctant to Realize Their Losses?," *Journal of Finance*, 53, 1775–1798.

Pontiff, J., 1996, "Costly Arbitrage: Evidence from Closed-End Funds," *Quarterly Journal of Economics*, 111, 1135–1151.

Purnanandam, A., and N. Seyhun, 2007, "Shorts and Insiders," Working Paper (July), University of Michigan.

Ross, S. A., 1976, "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

Shapira, Z., and I. Venezia, 2001, "Patterns of Behavior of Professionally Managed and Independent Investors," *Journal of Banking and Finance*, 25, 1573–1587.

Sharpe, W. F., 1964, "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 19, 425–442.

Shefrin, H. M., and M. Statman, 1985, "The Disposition to Sell Winners Too Early and Ride Losers Too Long," *Journal of Finance*, 40, 777–790.

Shefrin, H. M., and M. Statman, 2000, "Behavioral Portfolio Theory," *Journal of Financial and Quantitative Analysis*, 35, 127–151.

Shleifer, A., and R. W. Vishny, 1997, "The Limits of Arbitrage," *Journal of Finance*, 52, 35–55.

Shumway, T., 2001, "Forecasting Bankruptcy More Accurately: A Simple Hazard Model," *Journal of Business*, 74, 101–124.

Sias, R. W., 1996, "Volatility and the Institutional Investor," *Financial Analysts Journal*, 52, 13–20.

Spiegel, M. I., and X. Wang, 2005, "Cross-Sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk," Working Paper (September), Yale School of Management.

Wurgler, J., and E. Zhuravskaya, 2002, "Does Arbitrage Flatten Demand Curves for Stocks?," *Journal of Business*, 75, 583–608.

TABLE I

Small Trades Volume, Retail Holdings, and Retail Trading

This table reports the estimation results for monthly Fama-MacBeth (first three columns) and cross-sectional (last three columns) regressions, where the dependent variable is the proportion of retail trading (RTP) in a stock. It is defined as the ratio of the total buy- and sell-initiated small trades dollar volume and the total market dollar trading volume. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. The independent variables are measures of portfolio holdings and trading of a sample of individual investors at a large U.S. discount brokerage house. The excess portfolio weight in a certain month t is defined as the month-t weight of the stock in the aggregate portfolio of individual investors minus the month-t weight of the stock in the aggregate market portfolio. The aggregate individual investor portfolio is obtained by combining the portfolios of all investors in the sample and the market portfolio is defined by including all stocks available in the CRSP database. The excess trading weight is defined in an analogous manner using the total trading volume (sum of buy volume and sell volume) measure. In columns (1)-(3), we use the Pontiff (1996) methodology to correct Fama-MacBeth standard errors for potential serial correlation. In columns (4)-(6), we use the sample period averages of all variables and estimate one cross-sectional regression. The t-statistics for the coefficient estimates are shown in smaller font below the estimates. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. Both the dependent and the independent variables are standardized so that each variable has a mean of zero and a standard deviation of one. The sample period is from January 1991 to November 1996. The salient numbers in the table are shown in bold.

	Fama-MacBeth			Cross-Sectional		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.031	−0.001	0.029	0.031	−0.002	0.032
	11.11	−1.53	9.96	2.97	−0.03	3.041
Excess Portfolio Weight	0.322		**0.305**	0.505		**0.516**
	9.19		8.80	18.43		17.604
Excess Trading Weight		0.121	**0.089**		0.091	**0.066**
		7.93	7.51		7.79	3.76
(Average) Number of Observations	4,178	4,178	4,178	7, 572	7,572	7,572
(Average) Adjusted R^2	5.68%	1.59%	6.76%	16.33%	1.07%	17.09%

TABLE II

Small-Trades Volume and Stock Characteristics

This table reports the mean stock characteristics, conditional upon the level of retail trading in the stock. The retail trading proportion (RTP) is defined as the ratio of the total monthly buy- and sell-initiated small trades (trade size below $5,000) dollar volume and the total market dollar trading volume during the same month. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. The following stock characteristic measures are reported: idiosyncratic volatility (annualized, standard deviation of the residual from a four-factor model (the three Fama-French factors and the momentum factor), where daily returns during a month are used to estimate the model), firm size (in billion dollars), the book-to-market (B/M) ratio, past 12-month return (12mRet), stock price, institutional ownership (IO), proportion of stocks in the portfolio with low (less than 5%) institutional ownership (Low IO), idiosyncratic skewness (ISKEW), and analyst coverage (ANCOV). Idiosyncratic skewness is defined as the scaled third moment of residuals from a factor model that contains market return over the risk-free rate ($RMRF$) and $RMRF^2$ as factors and is estimated using six months of daily returns. The average level of RTP for the RTP decile portfolios are reported inside the brackets in the first column. The second column reports the fraction of the market represented by the RTP decile portfolio. Each decile portfolio contains an average of 410 stocks. The sample period is from January 1983 to December 2000. The institutional holdings data are from Thomson Financial. The salient numbers in the table are shown in bold.

RTP Decile	FracMkt	IVOL	Size	B/M	12mRet	Price	IO	Low IO	ISKEW	ANCOV
Low (0.74%)	**39.26%**	**11.66%**	**$3.145**	**0.592**	**31.62%**	**$46.19**	**50.78%**	**3.19%**	**0.386**	**11.59**
D2 (1.66%)	24.56%	12.03%	$1.556	0.639	27.71%	$27.16	44.56%	3.23%	0.398	9.57
D3 (2.70%)	15.59%	13.07%	$0.956	0.675	26.14%	$23.20	38.65%	4.41%	0.402	7.29
D4 (4.09%)	9.06%	13.89%	$0.546	0.719	23.54%	$19.67	32.79%	6.59%	0.414	5.54
D5 (6.04%)	5.04%	15.12%	$0.299	0.739	20.94%	$16.83	27.74%	9.12%	0.437	4.18
D6 (8.86%)	2.95%	16.47%	$0.174	0.765	17.10%	$13.98	23.15%	13.09%	0.493	3.08
D7 (13.27%)	1.76%	18.54%	$0.103	0.785	13.13%	$11.20	18.93%	18.52%	0.560	2.14
D8 (20.57%)	0.98%	22.20%	$0.058	0.783	7.86%	$8.28	14.83%	26.54%	0.610	1.37
D9 (33.64%)	0.56%	28.50%	$0.032	0.729	1.62%	$5.44	10.92%	38.14%	0.678	0.79
High (63.75%)	**0.24%**	**41.51%**	**$0.014**	**0.871**	**−11.46%**	**$2.97**	**3.01%**	**57.72%**	**0.745**	**0.30**

TABLE III

Idiosyncratic Volatility Preferences of Retail Investors: Sorting Results

This table reports the portfolio holdings and trading levels of retail investors in idiosyncratic volatility (IVOL) sorted portfolios. The idiosyncratic volatility in month t is defined as the standard deviation of the residual from a four-factor model (the three Fama-French factors and the momentum factor), where daily returns from month t are used to estimate the model. Using this definition, IVOL deciles are constructed each month and, for each of the deciles, we compute the percentage weights of retail holdings and trading. The sample period averages of those weights are reported in the table. Three different weights are reported: expected weights, actual weights, and the excess (actual − expected) weights. The expected decile-weight is the total weight of the decile-stocks in the aggregate market portfolio. The market portfolio is defined by including all stocks available in the CRSP database. The actual decile-weight in a certain month t is defined as the month-t weight of decile-stocks in the aggregate portfolio of individual investors at a large U.S. discount brokerage house. The aggregate individual investor portfolio is obtained by combining the portfolios of all investors. In column (1), we report the expected weights. Column (2) contains the actual portfolio weights based on the stock holdings of all investors in the brokerage sample. Column (3) contains the actual weights based on the stock trades of all investors in the brokerage sample. In column (4), we use the small trades in the Institute for the Study of Security Markets (ISSM) and the Trade and Quote (TAQ) databases to proxy for retail trading. Columns (5)-(7) reports the difference between columns (2)-(4) and (1), respectively. The sample period is from January 1991 to November 1996. The salient numbers in the table are shown in bold.

		Actual Weight			Excess Weight		
IVOL Decile	(1) Expected Weight	(2)	(3)	(4)	(5)	(6)	(7)
D1 (Low)	**40.66**	**22.15**	**15.82**	**16.01**	**−18.50**	**−24.84**	**−24.65**
D2	26.68	16.68	15.05	15.61	−10.00	−11.63	−11.07
D3	13.10	10.65	12.18	11.18	−2.45	−0.92	−1.92
D4	8.09	9.27	12.04	10.41	1.17	3.95	2.32
D5	5.10	9.61	13.22	11.67	4.52	8.13	6.57
D6	2.83	7.85	11.19	10.29	5.02	8.37	7.47
D7	1.64	6.99	9.06	9.25	5.36	7.42	7.61
D8	1.09	6.83	6.47	7.76	5.74	5.38	6.67
D9	0.56	5.34	3.04	4.69	4.78	2.48	4.13
D10 (High)	**0.27**	**4.63**	**1.92**	**3.13**	**4.37**	**1.66**	**2.87**

TABLE IV

Speculative Stock Characteristics and Retail Trading: Fama-MacBeth Regression Estimates

This table reports Fama-MacBeth regression estimates, where the dependent variable is the retail trading proportion (RTP) measure. It is defined as the ratio of the total buy- and sell-initiated small trades dollar volume and the total market dollar trading volume. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. The main independent variables are: (i) idiosyncratic volatility, which is the standard deviation of the residual from a four-factor model (the three Fama-French factors and the momentum factor) to the stock returns time-series, (ii) idiosyncratic skewness, which is defined as the scaled measure of the third moment of the residual obtained by fitting a two-factor ($RMRF$ and $RMRF^2$) model to the daily returns from previous six months, and (iii) stock price. We also define interaction terms using these three measures, where "high" category refers to the top three deciles and "low" refers to the lowest three deciles. A dividend paying dummy (it is set to one if the stock pays dividend at least once during the previous one year) and a NASDAQ dummy are also used as potential indicators of speculative characteristics. Additionally, the following control variables are employed: (i) systematic skewness (the coefficient of $RMRF^2$ in the two-factor regression to estimate idiosyncratic skewness), (ii) market beta, (iii) firm size, (iv) book-to-market (B/M) ratio, (v) past twelve-month stock return, and (vi) volume turnover, which is the ratio of the number of shares traded in a month and the number of shares outstanding. In column (4), the dependent variable is the excess weight assigned to a stock in the aggregate institutional portfolio. The excess portfolio weight allocated to stock i in quarter t is defined as: $EW_{ipt} = \frac{w_{ipt} - w_{imt}}{w_{imt}} \times 100$, where, w_{ipt} is the actual weight assigned to stock i in portfolio p in quarter t and w_{imt} is the weight of stock i in the aggregate market portfolio in quarter t. We use the Pontiff (1996) methodology to correct Fama-MacBeth standard errors for potential serial correlation. The t-statistics, obtained using the corrected standard errors, are reported in smaller font below the estimates. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. The independent variables have been standardized (mean is set to zero and standard deviation is one) so that the coefficient estimates can be directly compared within and across specifications. The sample period is from January 1983 to December 2000. The salient numbers in the table are shown in bold.

TABLE IV (Continued)

Speculative Stock Characteristics and Retail Trading: Fama-MacBeth Regression Estimates

Dependent variable is retail trading proportion (RTP) of stock i in month t.

Variable	(1)	(2)	(3)	(4)
Intercept	3.708	3.614	3.558	0.035
	13.32	13.27	13.79	4.80
Idiosyncratic Volatility	**2.297**	**1.998**	**1.698**	−0.054
	8.59	7.13	7.83	−5.71
Idiosyncratic Skewness	**0.364**	**0.209**	**0.201**	−0.026
	4.26	3.59	3.11	−2.66
Stock Price	**−0.834**	**−0.772**	**−0.898**	0.344
	−11.53	−12.36	−11.41	10.55
Dividend Paying Dummy	−0.025	−0.026	0.005	0.118
	−2.63	−2.16	0.33	7.90
NASDAQ Dummy	0.037	0.038	−0.006	−0.098
	1.59	1.74	−0.27	−6.31
High IVOL × High ISKEW		**0.179**	**0.156**	−0.045
		5.22	4.13	−3.19
High IVOL × Low Price		**0.214**	**0.276**	−0.063
		5.39	7.37	−4.80
High IKEW × Low Price		**0.140**	**0.163**	−0.040
		3.54	3.90	−3.88
Systematic Skewness			0.021	0.022
			0.78	2.37
Market Beta			0.102	0.103
			1.23	8.09
Log(Firm Size)			−0.099	0.193
			−5.89	6.83
Book-To-Market Ratio			0.159	0.014
			5.70	2.08
Past 12-Month Stock Returns			−0.161	0.089
			−2.42	5.18
Volume Turnover			0.639	−0.028
			6.14	−3.08
Lagged RTP	4.936	4.875	4.678	
	13.13	13.04	10.44	
Average Number of Observations	3,795	3,795	3,752	4,596
Average Adjusted R^2	0.451	0.486	0.522	0.328

TABLE V

Characteristics of Retail Clientele and Proportion of Retail Trading: Cross-Sectional Regression Estimates

This table reports cross-sectional regression estimates, where the dependent variable is the average retail trading proportion (RTP) or the buy-sell imbalance (BSI) of a stock measured over the 1991 to 1996 brokerage sample period. RTP is defined as the ratio of the total buy- and sell-initiated small trades dollar volume and the total market dollar trading volume. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. BSI is defined as the ratio of (Buy-initiated small-trades volume − Sell-initiated small-trades volume) and total small-trades volume. The main independent variables are the following characteristics of the retail investor clientele that trades the stock: Age, Annual Income, Education, Professional Occupation, Gender (Proportion Male), Marital Status (Proportion Married), Proportion Catholic, Proportion African American, Proportion Hispanic, Proportion Foreign Born, Proportion Urban (located within 100 miles of the top 25 U.S. metropolitan regions), Average State-Level Lottery Sales, and Portfolio Concentration (normalized portfolio variance). The clientele characteristic is the equal-weighted average characteristic of retail investors who trade the stock during the brokerage sample period, where stocks with fewer than five trades are excluded from the sample. To measure these variables, we use data on the portfolio holdings, trading and demographics of individual investors at a large U.S. discount brokerage house over the 1991 to 1996 time period. The t-statistics are reported in smaller font below the estimates. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. The independent variables have been standardized (mean is set to zero and standard deviation is one). The demographic characteristics of investors in the brokerage sample are measured a few months after the end of the sample period (June 1997) and are provided by Infobase Inc.

TABLE V (Continued)

Investor Characteristics and Retail Trading: Cross-Sectional Regression Estimates

Variable	(1):RTP×100	(2):RTP×100	(3):BSI×100
Intercept	8.371	8.375	−2.267
	22.12	20.56	−4.49
Age	−0.494	−0.539	0.208
	−2.41	−2.52	0.40
Income	−1.111	−0.781	0.074
	−5.33	−3.50	0.14
Professional Dummy	−0.466	−0.280	0.918
	−2.08	−1.63	1.63
Proportion Male	1.749	1.774	−1.299
	6.23	5.93	−1.82
Proportion Married	−1.264	−1.181	0.002
	−4.55	−4.08	0.22
Portfolio Concentration	2.277	2.228	−3.054
	9.96	9.39	−5.26
Education		−1.539	1.853
		−6.63	3.29
Proportion Catholic		1.455	1.700
		4.93	2.38
Proportion African American		0.356	0.077
		1.78	0.14
Proportion Hispanic		0.587	−0.114
		2.06	−0.19
Proportion Foreign Born		0.134	0.652
		0.56	1.86
Proportion Urban		0.474	0.937
		2.69	1.38
Average State-Level Lottery Sales		1.113	−1.603
		5.18	−3.05
Number of Stocks	6,231	5,925	5,925
Adjusted R^2	0.032	0.059	0.011

TABLE VI

Idiosyncratic Volatility and the Propensity to Sell Winners and Losers: Panel Regression Estimates

This table reports panel regression (pooled OLS with year fixed effects) estimates, where the dependent variable is one of the following three measures in year t for a given stock: (i) proportion of gains realized (PGR, Column (1) and (2)), (ii) proportion of losses realized (PLR, Column (3)), and (iii) disposition effect (DE) defined as PGR/PLR (Column (4). PGR is the proportion of gains realized and is defined as the ratio of the number of realized "winners" (stock positions where an investor experiences a gain) and the total number of winners (realized + paper). PLR is the proportion of losses realized and is defined in an analogous manner. To ensure that the measures are less noisy, stocks with fewer than ten trades during a year are excluded. The main independent variables are idiosyncratic volatility and idiosyncratic skewness. The idiosyncratic volatility measure is the variance of the residual obtained by fitting a four-factor model (the three Fama-French factors and the momentum factor) to the stock returns time-series. Idiosyncratic skewness is defined as the scaled measure of the third moment of the residual obtained by fitting a two-factor ($RMRF$ and $RMRF^2$) model to the daily returns from previous six months. $RMRF$ is the market return, excess over the risk-free rate. Both measures are estimated for each stock each month using daily returns data. Additionally, the following control variables are employed: (i) monthly volume turnover, which is the ratio of the number of shares traded in a month and the number of shares outstanding, (ii) firm age, which is the number of years between since the stock first appears in the CRSP database, (iii) market beta, (iv) firm size, (iii) book-to-market (B/M) ratio, (iv) past twelve-month stock return, (v) a dividend paying dummy, which is set to one if the stock pays dividend at least once during the previous one year, (vi) institutional ownership in the stock, (vii) a NASDAQ dummy, (viii) stock price, (ix) bid-ask spread, and (x) analyst coverage, which is defined as the number of analysts covering the stock during the past year. All independent variables are measured during the year $t-1$. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. Both the dependent and the independent variables have been standardized (mean is set to zero and standard deviation is one) so that the coefficient estimates can be directly compared within and across specifications. To account for potential serial and cross correlations in errors, we compute firm and year clustered standard errors. The t-statistics, obtained using the corrected standard errors, are reported in smaller font below the estimates. The salient numbers in the table are shown in bold.

TABLE VI (Continued)

Idiosyncratic Volatility, Idiosyncratic Skewness, and the Propensity to Sell Winners and Losers: Panel Regression Estimates

Variable	(1):PGR	(2):PGR	(3):PLR	(4):DE
Idiosyncratic Volatility	**0.224**	**0.185**	**−0.052**	**0.160**
	15.72	12.99	−3.55	8.48
Idiosyncratic Skewness	**0.150**	**0.133**	**−0.043**	**0.087**
	10.63	9.59	−2.58	9.21
Monthly Turnover		0.103	0.081	0.049
		8.84	3.91	3.65
Firm Age		0.010	0.014	−0.009
		2.53	2.59	−1.13
Market Beta		0.029	−0.016	0.018
		4.61	−1.05	1.72
Log(Firm Size)		−0.171	−0.101	−0.142
		−6.74	−5.14	−4.61
Book-To-Market Ratio		0.001	−0.022	−0.003
		0.064	−1.39	−0.47
Past 12-Month Stock Returns		0.032	0.085	−0.043
		1.97	4.85	−2.57
Dividend Paying Dummy		−0.009	−0.008	−0.003
		−0.48	−0.44	−0.14
NASDAQ Dummy		−0.030	−0.002	−0.010
		−1.74	−0.113	−0.55
Institutional Ownership		−0.030	−0.007	−0.024
		−1.55	−0.36	−1.62
Stock Price		−0.094	0.018	−0.029
		−3.26	0.61	−2.09
Bid-Ask Spread		−0.108	−0.124	0.053
		−4.48	−5.07	3.52
Log(1 + Analyst Coverage)		0.041	0.011	0.035
		1.61	1.06	1.55
Number of Observations	13,856	13,077	13,077	13,077
Adjusted R^2	0.058	0.136	0.051	0.069

TABLE VII

Risk Exposures and Performance of RTP Sorted Portfolios

This table reports the characteristics and performance of RTP sorted portfolios. The quintile portfolios are formed at the end of each month using the monthly RTP break-points. The monthly retail trading proportion (RTP) is the ratio of the retail trading volume and the total dollar trading volume in the market, where the retail trading volume is the sum of buy- and sell-initiated small trades dollar volume in the Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases. Panel A reports the main performance estimates, including the raw average monthly returns (both value-weighted and equal-weighted), characteristic-adjusted returns computed using the Daniel, Grinblatt, Titman, and Wermers (1997) methodology, factor exposures, four-factor alphas, and the adjusted R^2 obtained from time-series regressions of RTP-sorted portfolio returns on the three Fama-French factors plus the momentum factors. Panel B reports the performance estimates for different sub-periods and sub-samples. In the first sub-sample, we exclude stocks priced below \$5. In the second sub-sample, we skip a month between the portfolio formation month and performance measurement month. The low arbitrage cost sub-sample contains stocks that are in the bottom three deciles, where idiosyncratic volatility is the proxy for arbitrage costs. The high institutional ownership sub-sample contains stocks with IO in the highest three deciles. In Panel C, we report the raw and risk-adjusted performance estimates for RTP-based trading strategies that are potentially attainable. Shortable stocks are those that have positive short-interest and, as before, the low arbitrage cost sub-sample contains stocks that are in the bottom three IVOL sorted deciles. The t-statistics for the coefficient estimates are shown in smaller font below the estimates. Only stocks with CRSP share code 10 and 11 are included in the analysis. The sample period is from January 1983 to December 2000. The salient numbers in the table are shown in bold.

Panel A: Full-Sample Estimates

	RTP Quintile					
Measure	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	Low−High
Mean Monthly Return (EW)	2.964	1.944	1.416	0.763	−1.048	4.012
Std. Dev. EW Return	4.99	4.72	4.83	5.37	6.67	4.78
Mean Monthly Return (VW)	**1.765**	0.924	0.112	−1.008	−3.231	**4.996**
Std. Dev. VW Return	4.37	4.26	5.09	6.38	7.46	5.86
Characteristic-Adjusted Return	**0.415**	−0.248	−0.877	−1.505	−2.767	**3.182**
	8.76	−3.87	−7.92	−11.64	−13.09	14.12
Alpha	**0.490**	−0.312	−1.006	−2.093	−4.095	**4.585**
	7.72	−3.55	−8.45	−12.07	−16.09	16.14
RMRF Exposure	1.036	0.986	0.991	1.071	1.072	−0.036
	36.12	25.36	33.72	25.01	17.05	−0.52
SMB Exposure	−0.097	0.079	0.555	0.892	1.125	−1.223
	−4.93	2.87	14.96	16.49	14.18	−13.81
HML Exposure	0.071	0.118	0.027	−0.054	−0.186	0.258
	2.83	3.37	0.58	−0.79	−2.84	2.28
UMD Exposure	−0.091	−0.080	−0.112	−0.141	−0.304	0.213
	−5.19	−3.26	−3.38	−2.93	−4.30	2.69

Continued . . .

Panel B: Alpha Estimates for Sub-Samples

Sample	RTP Quintile Q1 (Low)	Q2	Q3	Q4	Q5 (High)	Low−High
Price ≥ $5	0.539	−0.172	−0.724	−1.624	−3.658	4.198
	7.43	−2.13	−5.97	−10.96	−10.64	11.33
Price ≥ $5, Skip One Month	0.596	−0.317	−0.804	−1.550	−3.552	4.148
	7.61	−4.13	−6.72	−10.64	−11.21	12.04
Price ≥ $5, Low Arb Costs	**0.368**	−0.118	−0.486	−0.826	−1.788	**2.156**
	3.69	−0.93	−3.37	−5.48	−6.47	7.01
Price ≥ $5, High IO	0.531	−0.202	−0.751	−1.788	−2.844	3.375
	7.26	−2.44	−5.93	−7.05	−6.70	15.66
1983-1991 Sub-Period	0.314	−0.224	−0.714	−1.583	−3.683	3.997
	4.29	−2.67	−7.47	−9.82	−15.03	15.11
Exclude January Returns	0.533	−0.248	−0.963	−2.143	−4.363	4.896
	4.29	−2.67	−7.47	−9.82	−15.03	15.11

Panel C: Performance of Potentially Implementable RTP-Based Trading Strategies

Portfolio	Mean Ret	Alpha	Factor Exposures RMRF	SMB	HML	UMD
Minimum Stock Price = $10						
RTP Bottom Half	1.517	0.255	1.009	−0.078	0.079	−0.082
		4.81	36.98	−4.74	3.74	−5.55
RTP Top Half	−0.297	−1.418	0.973	0.502	−0.071	−0.065
		−10.01	17.82	11.36	−1.27	−1.66
Difference	**1.814**	**1.673**	0.036	−0.580	0.150	−0.017
		10.23	0.89	−11.37	2.31	−0.36
Minimum Stock Price = $10, Shortable Stocks Only						
RTP Bottom Half	**1.516**	**0.226**	1.008	−0.175	0.220	−0.115
		3.34	40.41	−8.29	8.19	−6.13
RTP Top Half	**0.324**	**−0.825**	0.985	0.089	0.175	−0.191
		−8.35	20.39	2.89	4.45	−7.11
Difference	**1.192**	**1.051**	0.023	−0.264	0.046	0.080
		9.81	0.88	−7.89	1.07	2.69
Minimum Stock Price = $10, Lower Arbitrage Costs						
RTP Bottom Half	1.475	0.212	0.972	−0.194	0.199	−0.086
		3.27	37.09	−9.03	7.29	−4.49
RTP Top Half	0.253	−0.856	0.882	0.165	0.230	−0.156
		−5.99	15.01	3.69	4.05	−3.92
Difference	**1.222**	**1.068**	0.091	−0.359	−0.030	0.069
		7.63	2.67	−8.38	−0.55	1.82

TABLE VIII

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Sorting Results

This table reports the raw (Panel A) and risk-adjusted (Panel B) value-weighted, average monthly portfolio returns for idiosyncratic volatility (IVOL) and retail trading proportion (RTP) sorted portfolios. Panel C reports the performance estimates for the High IVOL − Low IVOL portfolio for various sub-samples. In Panel D, we report the raw and risk-adjusted average monthly returns for RTP and IVOL based trading strategies that are potentially attainable. Each month, the portfolios are constructed by performing independent sorts using the IVOL and RTP estimates in the previous month. The idiosyncratic volatility in month t is defined as the standard deviation of the residual from a four-factor model (the three Fama-French factors and the momentum factor), where daily returns from month t are used to estimate the model. The monthly retail trading proportion is the ratio of the retail trading volume and the total dollar trading volume in the market, where the retail trading volume is the sum of buy- and sell-initiated small trades dollar volume in the Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases. In Panel C, the first sub-sample excludes stocks with price below $5. In the second sub-sample, we skip a month between portfolio formation date and the performance measurement month. In the next two sub-samples, we consider only larger and shortable stocks. In the next test, we replace RTP with IO and perform the sorts. In the last two tests, we report sub-period results and estimates excluding January returns. The Newey-West adjusted t-statistics for the alpha estimates are shown in smaller font below the estimates. The sample period is from January 1983 to December 2000. The salient numbers in the table are shown in bold.

Panel A: Mean Monthly Returns

		Retail Trading Proportion (RTP) Quintile				
IVOL Quintile	All	Q1 (Low)	Q2	Q3	Q4	Q5 (High)
Q1 (Low)	**1.390**	**1.664**	1.096	0.390	−0.104	**−1.142**
Q2	1.319	1.758	0.732	0.081	−0.999	−1.893
Q3	1.282	2.343	0.943	0.153	−1.003	−2.900
Q4	0.724	2.837	1.388	−0.214	−1.249	−3.622
Q5 (High)	**−0.156**	**3.967**	3.014	0.862	−0.969	**−3.677**
High−Low	**−1.545**	**2.302**	1.919	0.472	−0.865	**−2.536**
	−6.56	6.68	5.37	1.80	−3.36	−11.53

Continued . . .

Panel B: Four-Factor Alpha Estimates

IVOL Quintile	All	Retail Trading Proportion (RTP) Quintile Q1 (Low)	Q2	Q3	Q4	Q5 (High)
Q1 (Low)	**0.099**	**0.349**	−0.148	−0.729	−1.180	**−2.148**
	1.58	3.81	−1.13	−5.15	−6.22	−6.30
Q2	0.098	0.526	−0.467	−0.997	−2.078	−2.757
	1.19	4.28	−3.64	−6.29	−11.31	−11.01
Q3	0.081	1.092	−0.210	−0.891	−2.067	−3.859
	0.71	5.88	−1.12	−5.62	−9.26	−15.50
Q4	−0.421	1.648	0.276	−1.241	−2.224	−4.386
	−2.87	5.57	1.10	−5.32	−10.18	−14.74
Q5 (High)	**−1.203**	**2.627**	1.511	−0.575	−1.935	**−4.570**
	−4.38	4.02	2.65	−1.51	−5.32	−13.05
High−Low	**−1.302**	**2.278**	1.659	0.154	−0.755	**−2.422**
	−4.23	3.43	2.74	0.38	−1.84	−5.72

Panel C: High IVOL − Low IVOL Four-Factor Alpha Estimates for Sub-Samples

Sub-Sample	All	RTP Quintile Q1 (Low)	Q2	Q3	Q4	Q5 (High)
Price ≥ \$5	**−1.104**	**2.208**	1.503	0.718	−0.996	**−2.228**
	−4.04	5.50	3.97	2.24	−2.98	−4.61
Price ≥ \$5, Skip One Month	**−1.077**	**2.051**	1.281	0.689	0.049	**−1.671**
	−3.74	5.44	3.95	2.49	0.18	−4.99
All NYSE/AMEX Stocks	**−0.913**	**1.518**	0.492	−0.031	−1.128	**−2.139**
	−3.96	6.17	2.09	−0.14	−4.04	−5.51
Shortable NYSE/AMEX Stocks	**−0.471**	**0.944**	0.195	−0.620	−1.164	**−1.827**
(1991-2000 period)	−3.18	5.24	1.50	−3.22	−4.26	−6.03
1983-1991 Sub-Period	**−1.414**	**1.350**	0.812	−0.148	−0.967	**−2.582**
	−4.99	2.82	2.04	−0.27	−1.89	−6.10
Exclude January Returns	**−1.522**	**2.306**	1.786	−0.203	−0.812	**−2.715**
	−5.11	3.34	3.05	−0.49	−1.88	−6.44

Continued . . .

Panel D: Performance of Potentially Implementable Low RTP Trading Strategies

			Factor Exposures			
Portfolio	Mean Ret	Alpha	RMRF	SMB	HML	UMD
Minimum Stock Price = $10, RTP in Bottom Third						
IVOL Bottom Half	**1.635**	**0.355**	1.023	−0.146	0.133	−0.104
		5.34	32.27	−7.06	5.05	−5.63
IVOL Top Half	**2.622**	**1.377**	1.206	0.409	−0.426	−0.022
		6.41	12.71	6.11	−4.98	−0.37
Difference	**0.987**	**1.022**	**0.183**	**0.556**	**−0.559**	0.082
		4.85	3.52	8.46	−6.67	1.39
Minimum Stock Price = $10, RTP in Bottom Half						
IVOL Bottom Half	**1.578**	**0.311**	0.997	−0.138	0.140	−0.096
		4.42	37.42	−7.36	5.91	−5.78
IVOL Top Half	**2.242**	**1.069**	1.182	0.442	−0.395	−0.088
		5.01	14.70	7.32	−5.14	−1.62
Difference	**0.664**	**0.758**	**0.185**	**0.579**	**−0.535**	0.009
		4.09	3.84	9.54	−6.91	0.16

TABLE IX

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Fama-MacBeth Cross-Sectional Regression Estimates

This table reports the estimates from monthly Fama-MacBeth cross-sectional regressions, where the monthly stock return is the dependent variable. The main independent variables are idiosyncratic volatility (IVOL) and a measure of retail trading (RTP), both measured at the end of the previous month. The IVOL in month t is defined as the standard deviation of the residual from the factor model, where daily returns from month t are used to estimate the model. The monthly retail trading proportion (RTP) is the ratio of the retail trading volume and the total dollar trading volume, where the retail trading volume is the sum of buy- and sell-initiated small trades dollar volume in the Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases. Other independent variables include three factor exposures (market, small-minus-big (SMB), and high-minus-low (HML) betas) and three firm characteristics (firm size, book-to-market ratio, and past six-month return). The factor exposures are measured "contemporaneously", firm size and six-month returns are measured in the previous month, and the book-to-market measure is from six months ago. In column (1), we use the exact specification and the exact time-period (1980 to 2003) used in Ang, Hodrick, Xing, and Zhang (2008). All other regressions reported in this table use the time-period for which the retail trading data are available (1983 to 2000). In column (4) only, we use the total volatility measure instead of idiosyncratic volatility. In column (5), we add High RTP × High IVOL and Low RTP × High IVOL interaction terms in the specification. High (low) RTP × High IVOL is a dummy variable that takes value 1 for stocks that rank in the top (bottom) quintile by RTP and in the top quintile by IVOL. In column (6), we only include stocks that rank in the bottom one-third by RTP. In column (7), we only consider stocks with higher levels of retail trading (RTP in the top-third). The sample period is from January 1983 to December 2000 in all columns except the first column where we use the AHXZ sample period (January 1980 to December 2003). We follow the Pontiff (1996) methodology to correct the Fama-MacBeth standard errors for potential serial correlation. The t-statistics for the coefficient estimates are shown in smaller font below the estimates. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. The independent variables are standardized such that each variable has a mean of zero and a standard deviation of one. Only stocks with CRSP share code 10 and 11 are included in the analysis. The salient numbers in the table are shown in bold.

TABLE IX (Continued)

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Fama-MacBeth Cross-Sectional Regression Estimates

Dependent variable is the return of stock i in month t.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	1.554	1.513	1.468	1.438	1.492	2.592	−0.477
	4.27	3.23	3.51	3.66	3.58	6.84	−1.18
Total or	**−0.491**	**−0.479**	**0.393**	**0.408**	**0.464**	**0.247**	**−0.513**
Idiosyncratic Volatility	−4.35	−4.59	3.69	4.61	4.71	2.54	−3.79
AHXZ Variables							
Market Beta	1.093	1.019	1.085	1.043	0.953	0.966	1.026
	6.18	5.27	5.22	5.06	3.92	4.95	4.66
SMB Beta	−0.069	−0.081	−0.065	−0.039	−0.070	−0.048	0.174
	−0.75	−0.80	−0.58	−0.35	−0.51	−0.41	1.96
HML Beta	−0.616	−0.440	−0.465	−0.396	−0.350	−0.339	−0.515
	−2.92	−2.53	−2.48	−2.10	−1.50	−1.98	−2.74
Log(Firm Size)	−0.429	−0.387	−0.629	−0.702	−0.803	−1.181	−0.638
	−4.25	−3.93	−5.53	−5.99	−6.52	−8.31	−5.91
Book-To-Market	0.257	0.192	0.149	0.114	0.213	0.181	0.173
	2.42	3.45	2.57	2.50	3.98	3.57	3.10
Past Six-Month Return	−0.130	−0.167	−0.132	−0.124	−0.189	−0.104	−0.633
	−2.62	−1.07	−2.71	−3.26	−3.18	−1.18	−4.63
Retail Preference Proxies							
Retail Trading Proportion			**−1.541**	**−1.505**	**−1.273**		
			−10.69	−9.07	−9.68		
High RTP × High IVOL					**−0.469**		
					−8.80		
Low RTP × High IVOL					**0.247**		
					6.59		
Average Number of Stocks	4,765	3,738	3,503	3,503	3,503	1,168	1,168
Average Adjusted R^2	0.051	0.056	0.062	0.066	0.070	0.086	0.050

TABLE X

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Regression Estimates for Skewness Sub-Samples

This table reports the estimates from monthly Fama-MacBeth cross-sectional regressions for skewness-based sub-samples. We consider the baseline specification defined in Table IX. The monthly stock return is the dependent variable. The main independent variables are idiosyncratic volatility (IVOL) and a measure of retail trading (RTP), both measured at the end of the previous month. Idiosyncratic skewness is defined as the scaled measure of the third moment of the residual obtained by fitting a two-factor ($RMRF$ and $RMRF^2$) model to the daily returns from previous six months. The estimates are reported for (i) the AHXZ baseline specification and (ii) specification with the RTP measure added to the AHXZ specification. We also report the estimates for the sub-sample of stocks with price below $10. All independent variables have been standardized. The t-statistics for the coefficient estimates are reported in smaller font below the estimates. The salient numbers in the table are shown in bold.

	Idiosyncratic Skewness				
	Low	Q2	Q3	Q4	High
AHXZ specification					
Idiosyncratic Volatility	**−0.089**	−0.284	−0.392	−0.415	**−0.588**
	−0.83	−2.12	−2.98	−3.52	−4.77
AHXZ specification with RTP					
Idiosyncratic Volatility	0.235	0.275	0.364	0.465	0.426
	2.02	3.32	2.56	2.46	2.76
Retail Trading Proportion	**−0.878**	−1.109	−1.472	−1.726	**−1.781**
	−5.94	−8.53	−9.07	−8.51	−9.53
AHXZ specification with RTP (Stock Price < $10)					
Idiosyncratic Volatility	0.317	0.532	0.390	0.540	0.691
	2.47	3.37	3.03	3.45	3.89
Retail Trading Proportion	**−1.322**	−1.519	−1.590	−1.744	**−1.877**
	−8.62	−9.26	−9.23	−9.62	−9.47

(Coefficient estimates of AHXZ variables have been suppressed.)

TABLE XI

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Robustness Test Results

This table reports the estimates from monthly Fama-MacBeth cross-sectional regressions, where the monthly stock return is the dependent variable. In Panel A, we consider several variations on the baseline specification defined in Table IX. In Panel B, we estimate the baseline regression for leverage-based sub-samples. Following Johnson (2004), leverage is defined as the book value of debt divided by the book value of debt plus market value of equity. The main independent variables are idiosyncratic volatility (IVOL) and a measure of retail trading, both measured at the end of the previous month. In column (1), these two measures are obtained using the previous three months of data. Retail buy-sell imbalance is the ratio of buy-sell volume differential and the total buy- and sell-initiated retail trading volume. Institutional ownership is the level of institutional ownership at the end of the previous quarter and the change in institutional ownership is the quarterly change in institutional ownership level. Idiosyncratic skewness is the scaled measure of the third moment of the residual obtained by fitting a two-factor ($RMRF$ and $RMRF^2$) model to daily returns. Other independent variables (three factor exposures and three firm characteristics) are defined in Table VII. We follow the Pontiff (1996) methodology to correct the Fama-MacBeth standard errors for potential serial correlation. The t-statistics for the coefficient estimates are shown in smaller font below the estimates. To ensure that extreme values are not affecting the results, we winsorize all variables at their 0.5 and 99.5 percentile levels. The independent variables are standardized such that each variable has a mean of zero and a standard deviation of one. Only stocks with CRSP share code 10 and 11 are included in the analysis. Also, stocks with fewer than 15 daily returns in a month are excluded in that month. The sample period is from January 1983 to December 2000. The institutional holdings data are from Thomson Financial. The salient numbers in the table are shown in bold.

TABLE XI (Continued)

Idiosyncratic Volatility, Retail Trading, and Stock Returns: Robustness Test Results

Panel A: Estimates from Extended Specifications

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	1.514	1.458	1.477	1.519	1.498	1.531
	3.37	3.25	3.31	4.11	3.31	3.66
Idiosyncratic Volatility	**0.481**	**−0.548**	**−0.366**	**−0.483**	**−0.457**	**0.416**
	4.09	−4.52	−3.28	−4.16	−4.17	4.28
Retail Trading Proportion	−1.839					−1.522
	−11.44					−10.55
Retail Buy-Sell Imbalance		0.215				0.225
		5.22				6.01
Past One-Month Return			−0.840			−0.803
			−8.06			−7.82
Institutional Ownership				0.031		0.029
				1.55		1.44
Δ*Institutional Ownership*				−0.020		−0.018
				−0.42		−0.40
Idiosyncratic Skewness					−0.145	−0.209
					−4.77	−3.44
(Coefficient estimates of AHXZ variables have been suppressed.)						
Average Number of Stocks	3,427	3,503	4,735	3,290	4,706	3,290
Average Adjusted R^2	0.066	0.054	0.053	0.052	0.053	0.071

Panel B: Estimates from Leverage Sub-Samples

	Leverage				
Variable	Low	Q2	Q3	Q4	High
AHXZ specification					
Idiosyncratic Volatility	**−0.365**	**−0.342**	**−0.412**	**−0.385**	**−0.379**
	−3.07	−2.52	−3.14	−3.12	−3.59
AHXZ specification with RTP					
Idiosyncratic Volatility	**0.328**	**0.181**	**0.227**	**0.359**	**0.421**
	2.48	1.61	2.43	2.64	3.89
Retail Trading Proportion	−1.583	−1.377	−1.354	−1.398	−1.698
	−9.93	−8.83	−8.46	−9.23	−10.19
(Coefficient estimates of AHXZ variables have been suppressed.)					

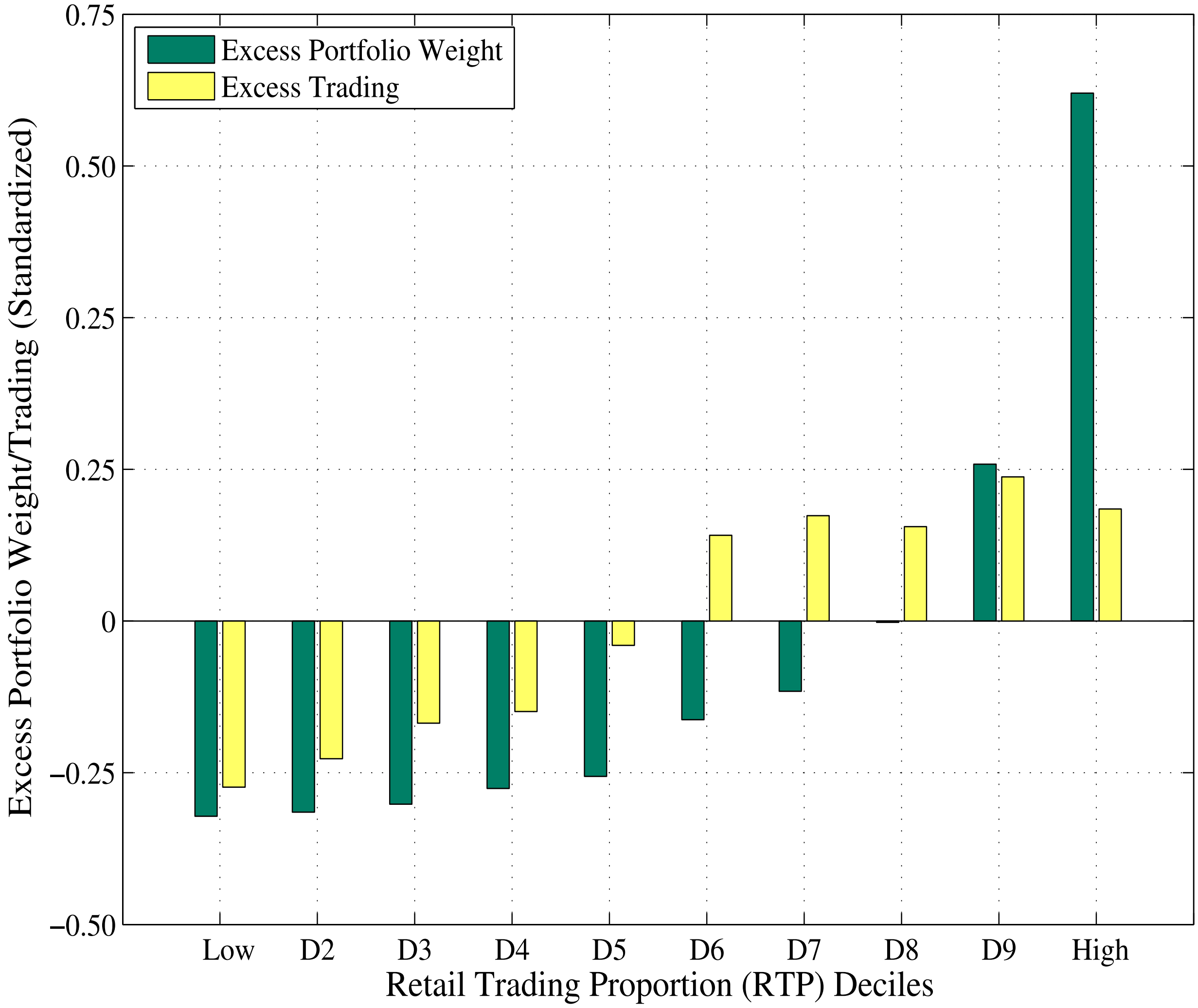

Figure 1. Small trades, retail holdings, and retail trades. This figure shows the average portfolio weights and retail trading weights in the brokerage data, conditional upon the level of retail trading proportion of the stock. The retail trading proportion (RTP) is defined as the ratio of the total buy- and sell-initiated small trades (trade size below $5,000) dollar volume and the total market dollar trading volume in the ISSM/TAQ data. The excess weight is defined as the difference between actual and expected weights. The expected decile-weight is the total weight of the decile-stocks in the aggregate market portfolio. The market portfolio is defined by including all stocks available in the CRSP database. The actual decile-weight in a certain month t is defined as the month-t weight of decile-stocks in the aggregate portfolio of individual investors. The aggregate individual investor portfolio is obtained by combining the portfolios of all investors. The weights are standardized (mean is set to zero and the standard deviation is one) to facilitate comparisons between the two weight measures. The sample period averages of those weights are shown in the figure. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. The sample period is from January 1991 to November 1996.

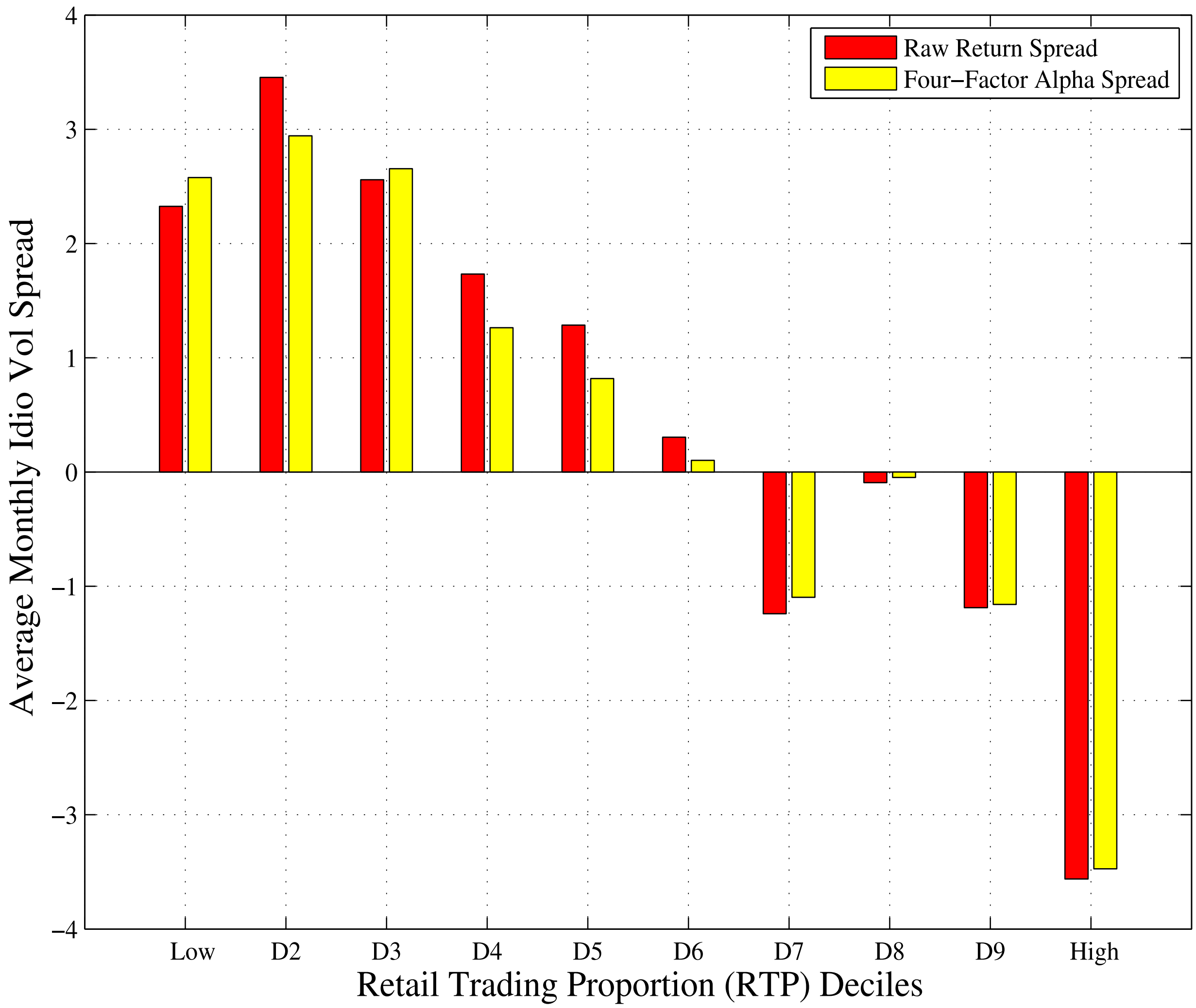

Figure 2. Retail trading and idiosyncratic volatility spread. This figure shows the average monthly idiosyncratic volatility (IVOL) spread, conditional upon the level of retail trading in the stock. The retail trading proportion (RTP) of a stock is defined as the ratio of the total buy- and sell-initiated small trades (trade size below $5,000) dollar volume and the total dollar trading volume for that stock. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. The idiosyncratic volatility spread is the difference between the value-weighted, monthly returns of high (top three deciles) and low (bottom three deciles) idiosyncratic volatility portfolios. Both the raw and the risk-adjusted spreads are plotted. Each month, the portfolios are constructed by performing independent sorts using the IVOL and retail trading proportion (RTP) estimates in the previous month. The idiosyncratic volatility in month t is defined as the standard deviation of the residual from a four-factor model (the three Fama-French factors and the momentum factor), where daily returns from month t are used to estimate the model. The sample period is from January 1983 to December 2000.

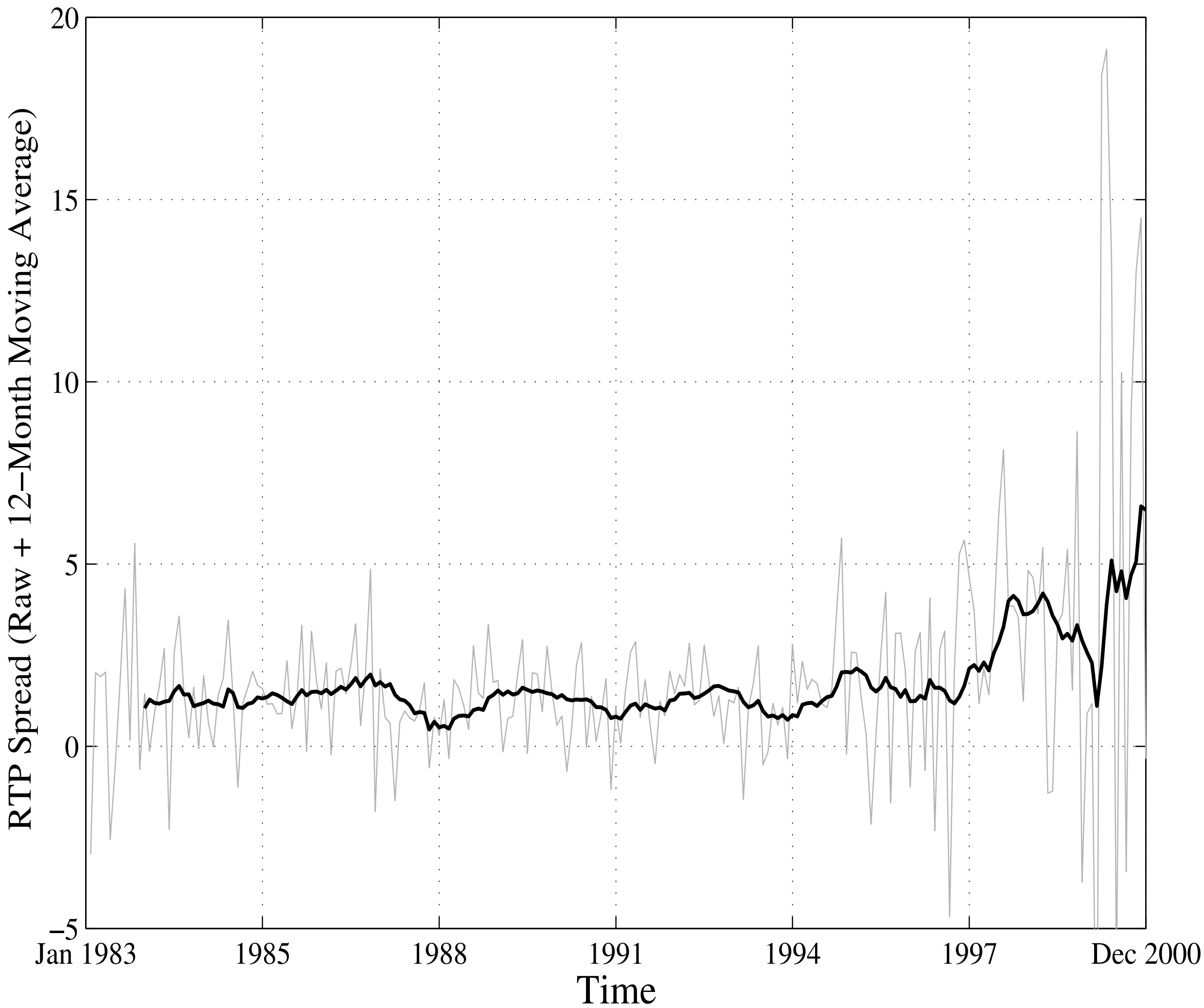

Figure 3. Time-series variation in the performance of RTP-based trading strategy. This figure shows the retail trading proportion (RTP) spread, defined as the difference in the value-weighted average monthly returns (in percentage) of low (bottom half) RTP and high (top half) RTP portfolios. Both the raw and the 12-month backward moving average time series are plotted. The retail trading proportion (RTP) is defined as the ratio of the total buy- and sell-initiated small trades (trade size below $5,000) dollar volume and the total market dollar trading volume. The small trades data are from Institute for the Study of Security Markets (ISSM) and Trade and Quote (TAQ) databases, where small trades are used as a proxy for retail trades. RTP portfolios are constructed each month by sorting on RTP estimates in the previous month. In any given month, stocks priced below $10 are excluded from the sample. The sample period is from January 1983 to December 2000.