INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Asset Pricing and Fund Investment Anomalies

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

by

Tobias Jacob Moskowitz

1998
The dissertation of Tobias Jacob Moskowitz is approved.

Richard Roll

Pedro Santa-Clara

Olivier Ledoit

Didier Sornette

Mark Grinblatt, Committee Chair

University of California, Los Angeles

1998
To Bonnie, whose unconditional support, advice, patience, and love, made this possible.

To my mother and father, for their support, advice, and ability to listen.
# Table of Contents

1 Industry Factors as an Explanation for Momentum in Stock Returns ................................................................. 1  
1.1 Introduction .................................................................................................................................................. 1  
1.2 Data Description and Industry Returns ........................................................................................................ 4  
1.3 Motivation ..................................................................................................................................................... 6  
1.3.1 Return Generating Process .................................................................................................................... 6  
1.4 Isolating Sources of Momentum Trading Profits ............................................................................................ 8  
1.4.1 Momentum Investment Strategies ........................................................................................................ 8  
1.5 Robustness of Industry Momentum Strategies .............................................................................................. 23  
1.5.1 Industry Momentum Trading Strategies ............................................................................................ 23  
1.5.2 The 1-Month, 1-Month Strategy and the Lead-Lag Effect ................................................................. 25  
1.5.3 Industry Momentum Independent of Individual Stock Momentum .................................................. 30  
1.6 Efficiency of Industry Momentum Strategies ................................................................................................. 32  
1.6.1 Diversification of Industry Momentum Strategies .............................................................................. 33  
1.6.2 Industry Momentum Factors ................................................................................................................ 35  
1.7 Conclusion .................................................................................................................................................... 39  
1.8 Appendix A: Return Adjustment Procedures ............................................................................................... 41  
1.8.1 Risk-Based Models ................................................................................................................................ 41  
1.8.2 Characteristic-Based Models ................................................................................................................ 43
3.2.2 Specification Test ............................................. 128

3.3 Time Series Analysis of Geographic Equity Preference .......... 131
  3.3.1 Tests on U.S. Mutual Fund Domestic Equity Holdings ... 131
  3.3.2 Cross-Section of Funds and Geographic Equity Preference 132

3.4 Relation Between Distance and Fund Turnover .................. 137

3.5 Effect of Local Bias on Fund Performance ........................ 139
  3.5.1 Performance Results of Local versus Distant Holdings ... 139
  3.5.2 Fund Characteristics and Local Performance .............. 142

3.6 Relation Between Local Bias, Information, and the Cross-Section
  of Expected Returns ............................................. 143
  3.6.1 Zero-Cost Portfolio Comparisons ............................ 144
  3.6.2 Expected Return Prediction of Local Ownership ............ 146
  3.6.3 A Link between Local Ownership and Asset Pricing Anomalies? ............................................. 149

3.7 Conclusion .................................................. 152
LIST OF FIGURES

1.1 Ex-Post Efficient Frontiers of Industry and ‘Random’ Industry Portfolios ......................................................... 52

2.1 Distribution of Trading Strategy Profits ........................................ 113

2.2 Distributions for Zero-Intercept F-test Statistic Under the Null (CAPM), Risk-Based. and Non-Risk Alternatives ........ 114

2.3 GARCH vs. Rolling Regression-Estimated Market β’s .............. 115

2.4 Relation Between Conditional Risk and Past Performance ........ 116

3.1 Local Bias Over Time: Difference Between Market and Actual Average Distances ................................................. 154

3.2 City Size and Local Bias .......................................................... 154

3.3 Cumulative Profits on Zero Cost Portfolios Over Time (1980:3 - 1994:4) ................................................................. 156

3.4 Cumulative Risk-Adjusted Profits on Zero Cost LMD Portfolio .... 158

vii
# List of Tables

1.1 Description and Summary Statistics of Industries .......................... 45  
1.2 Average Monthly Raw Excess and Abnormal Returns of Industries .... 46  
1.3 Momentum Profits for Individual Equities, Industries, and Random Industries ............................................................... 47  
1.4 Industry Momentum Trading Profits ....................................... 48  
1.5 Impact of Lead-Lag Effects on Industry Momentum Profits .......... 49  
1.6 Industry Momentum Independent of Individual Stock Momentum 50  
1.7 Diversification of Industry Momentum Strategies ....................... 51  
1.8 Average Monthly Returns of Industry Momentum Factor Loading and Past Return Sorted Portfolios ........................................ 53  

2.1 Size, BE/ME, and Momentum Profits Attributed to Risk ............... 106  
2.2 Common Volatility Tests of Size, BE/ME, and Momentum Portfolios 107  
2.3 Minimizing Risk of Size, BE/ME, and Momentum Strategies .......... 108  
2.4 Non-Nested Cross-Sectional Regression Tests of Factor Models 109  
2.5 Predicting Future Volatility ................................................. 110  
2.6 Remaining Volatility ......................................................... 111  
2.7 Risk of High and Low Past Return Industries ............................. 112  

3.1 Results of Tests for Local Bias Among U.S. Mutual Funds .......... 155  
3.2 Results of Tests for Local Bias Across Fund Investment Objective 156  
3.3 Results of Tests for Local Bias Across Fund Types ..................... 157  
3.4 Covariance Between Turnover and Distance .............................. 158
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5 Performance Differentials of Local and Distant Fund Portions</td>
<td>159</td>
</tr>
<tr>
<td>3.6 Regression of Adjusted Returns of Fund Holdings on Degree of Local Bias</td>
<td>160</td>
</tr>
<tr>
<td>3.7 Performance Measures of Local Investment Sorted Portfolios</td>
<td>161</td>
</tr>
<tr>
<td>3.8 Average Characteristics of the Local Investment Sorted Portfolios</td>
<td>161</td>
</tr>
<tr>
<td>3.9 Regression of DGTW Adjusted LO5-LO1 Value-Weighted Portfolio on Carhart Factors</td>
<td>162</td>
</tr>
<tr>
<td>3.10 Individual Security Fama-MacBeth Regressions</td>
<td>162</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

Many people have contributed greatly to the work that follows and I apologize if I omit anyone. In particular, I would like to thank several of my colleagues, Felipe Aguerrevere, Tony Bernardo, Michael Brennan, Ty Callahan, Bhagwan Chowdhry, Gordon Delianedis, Jonathan Howe, John McConnell, Sheridan Titman, Walt Torous, Ivo Welch, and Dave Wessells, for stimulating interesting discussions and commenting on my work. I thank Kent Daniel, Mark Grinblatt, Sheridan Titman, and Russ Wermers for graciously supplying data used in part of this dissertation and thank seminar participants at the Fall 1997 UCLA Finance seminar series for their comments. I am also grateful to the doctoral office, particularly Steve Lippman for his additional financial support, and Lydia Heyman who undoubtedly saved me hours of administrative work. In addition, I wish to thank Josh Coval, who patiently listened to half-baked ideas and came up with a few of his own. The third chapter of this dissertation is largely the result of our joint effort. I would also like to thank Bonnie Clark for her suggestions and advice, and whose support and patience made this possible. Finally, I am indebted to my committee members Richard Roll, Pedro Santa-Clara, Olivier Ledoit, and Mark Grinblatt for painstakingly reviewing my work and prompting interesting and challenging discussions. Their efforts, support, and guidance is greatly appreciated and enabled the following project to be completed. In particular, I owe Mark Grinblatt, my chairman, an immense debt of gratitude for stimulating my interest in the field, teaching me everything I know about finance, and advising, mentoring, and fostering my research in the area. Thank you all.
VITA

1971        Born. West Lafayette, IN. USA.

1992        Elected to Phi Beta Kappa. Beta Gamma Sigma. Phi Kappa Phi, and Golden Key, Purdue University.

1993        B.S. with distinction in Industrial Management and Industrial Engineering, Purdue University.

1994        M.S. in Management. Krannert Graduate School of Management, Purdue University.

Abstract of the Dissertation

Asset Pricing and Fund Investment Anomalies

by

Tobias Jacob Moskowitz

Doctor of Philosophy in Management

University of California, Los Angeles, 1998

Professor Mark Grinblatt, Chair

This dissertation explores two empirical puzzles in finance. The first puzzle is the momentum effect in stock returns. Specifically, equity returns seem to be positively autocorrelated over the 3-12 month investment horizon, and a strategy which buys stocks which have performed well in the past, and sells stocks which have performed poorly (e.g., a momentum investment strategy), appears to be highly profitable, generating about 12% per year per dollar long on a zero investment portfolio. These findings seem to violate the notion of efficient markets, and much of the recent literature has attempted to explain these phenomena via behavioral or cognitive biases among investors.

In the first essay of my dissertation, I identify the source of these momentum profits, and find that industry effects are largely responsible. That is, the extent to which stocks continue to perform well (or poorly) is mainly due to the stock belonging to the industry that happens to have performed well (poorly). In addition, industry momentum strategies appear to be highly profitable, and seem to dominate momentum strategies based on individual stocks. In the second essay, I then try to identify whether the profits from momentum strategies are attributed to risk, and find little evidence that risk-based explanations are
driving the profitability of such strategies. In addition, I analyze the risk of other investment strategies designed to exploit various stock market anomalies. Specifically, I focus on the size and book-to-market equity anomalies, which appear to predict expected stock returns. Both size and book-to-market equity appear more reasonably related to systematic risk than momentum, and strategies designed to exploit these anomalies are more risky and less profitable than momentum investment strategies.

The second empirical puzzle my dissertation addresses is the "home bias phenomenon", which is the fact that investors invest close to 94% of their wealth in domestic securities, largely avoiding exposure to international markets. From a diversification standpoint, such behavior appears grossly inefficient. However, in my third essay, I argue that this bias for nearby investments is not a uniquely international phenomenon. Specifically, I find that professional U.S. investment managers exhibit a strong preference for locally-headquartered firms in their domestic portfolios. This preference is most pronounced in firms which are small, highly levered, and produce non-traded goods, and is more acute among small managers from remote locations. Furthermore, the performance of the local portion of investment manager portfolios is significantly greater than that of their non-local portion. adjusting for risk. The results are supportive of asymmetric information interpretations of the home bias puzzle, and offer a rational explanation for the propensity to invest locally. Finally, these results may offer an intriguing clue for cross-sectional asset pricing anomalies.
CHAPTER 1

Industry Factors as an Explanation for Momentum in Stock Returns

1.1 Introduction

The ability to forecast future expected returns based on past returns alone remains an unexplained empirical anomaly.\footnote{The anomalous strong autocorrelation in stock returns has been documented by (among others) DeBondt and Thaler (1985), Lo and MacKinlay (1988), Jegadeesh (1990), and Jegadeesh and Titman (1993).} This paper focuses on the positive persistence in stock return performance (or momentum effect) over intermediate investment horizons (3 to 12 months) found by Jegadeesh and Titman (1993) (JT), and explores various explanations for its existence. JT offer a behavioral explanation for momentum, claiming that positive return autocorrelation is driven by “delayed price reactions to firm-specific information” (p. 67). However, if investors persistently and irrationally underreact to firm-specific information, then rational investors can profit from their irrational counterparts. Indeed, JT generate large profits from their “relative strength” or momentum trading strategies, which invest in a portfolio of winning stocks (high past returns) funded by a short position in a portfolio of losing stocks (low past returns). In an economy with an infinite number of assets, a similar self-financing portfolio long high past return stocks and short low past return stocks (weighted to have a similar factor beta
configuration as the winner portfolio) could be created with zero risk and positive expected return. implying that asymptotic arbitrage opportunities would exist. In a large asset economy, therefore, prices of high past return stocks should be bid up and those of low past return stocks bid down until the profitability of such strategies disappears. unless 'irrational' investors exist who are willing to hold the opposite position (i.e., long losing stocks and short winning stocks), a wealth depleting strategy. Not surprisingly, therefore, proponents of rational asset pricing claim such anomalies are due to risk, and that momentum trading profits are merely compensation for that risk.

This paper attempts to identify the source of momentum profits by comparing the relative magnitudes of the components contributing to the profits. Identifying the source of these profits may shed light on whether momentum is explained by risk or behavioral hypotheses, and if behavioral explanations apply, then finding the source of momentum identifies the best trading strategy to exploit this anomaly. While JT claim that serial correlation in firm-specific return components is responsible for momentum, recently, Conrad and Kaul (1997) claim that momentum profits are "entirely due to the contribution of cross-sectional variance in mean returns to total profits." (p. 3) rather than to predictability in the time-series of returns, although they do not attempt to determine the source of this cross-sectional variation (rational or irrational). Isolating the components of momentum profits, we find that neither idiosyncratic components nor dispersion in unconditional mean returns explains the profitability of momentum strategies. We do, however, find strong evidence that industry components are generating momentum. Specifically, once returns are adjusted for industry effects, momentum in individual stocks largely disappears. Conversely, industries themselves

---

2 Even with short sales restrictions, the same arguement holds as rational investors would only be willing to hold high past return stocks, avoiding low past return stocks.
exhibit significant momentum, even when controlling for size, book-to-market equity (BE/ME), and momentum in individual stock returns.

These findings may be consistent with both behavioral and rational theories of valuation. However, from a behavioral perspective, such findings would offer a refinement of the type of information investors underreact to (i.e., industry information). From a rational asset pricing perspective, industries may have systematic influences on returns, suggesting that industry-related (time-varying) risk premia may exist, which drive momentum.

Initial evidence seems to favor a risk-based interpretation for momentum profits, however. As industry momentum strategies appear to be risky and diversification across industries does not seem to alleviate this risk. That is, naive strategies that invest in high past return industries and short low past return industries have the same variance as sophisticated strategies which optimally minimize variance for the same expected return. Specifically, we find that mean-variance efficient strategies formed in the prior period underperform (are less efficient than) naive momentum strategies in the next period. Furthermore, industry momentum prevails after controlling for size. BE/ME and momentum in individual equities, and industry momentum factor-mimicking portfolios are closer to multifactor or local mean-variance (LMV) efficient than Carhart’s (1997) PR1YR momentum factor (which does not account for industry effects), and explain a larger cross-section of expected returns. In addition, stocks with high loadings on the

---

3 Given a set of factor loadings and expected return, these portfolios have the smallest variance (i.e., for this “local” subset of feasible returns, the portfolio is mean-variance efficient). For further discussion of local mean variance efficiency see Grinblatt and Titman (1987).

4 These results are, in some sense, matched by Cohen and Polk (1995) who develop industry adjusted book-to-market measures that are LMV more efficient and explain the cross-section of expected returns better than Fama and French’s (1993) HML factor (a book-to-market measure that does not account for industry effects). Both sets of results indicate the pricing importance of industries. However, while Cohen and Polk demonstrate that intra-industry measures provide a more refined estimate of the relation between BE/ME and the cross-section of expected returns, our results indicate that industries themselves drive momentum, and that
industry momentum factor exhibit higher average returns than stocks with low loadings, indicating a premium exists for industry momentum. Although there is little evidence of unconditional industry risk premia, these results suggest that industry factors may have important, conditional asset pricing implications, and thus favor a risk-based interpretation for momentum. In addition, loadings on PRIYR do not seem to be related to expected returns, suggesting that the industry momentum factor better captures dispersion in average returns, and thus is preferred as a benchmark for evaluating investment performance.

The rest of the paper is organized as follows. Section 2 briefly describes the data and formation of industries. Section 3 presents the motivation for the paper, based on a simple return generating process, and discusses the various sources of momentum profits. Section 4 isolates the sources of momentum profits and examines whether momentum is driven by industry effects. Section 5 further analyzes conditional industry returns and documents strong industry momentum patterns independent from momentum in individual stocks. Section 6 examines the efficiency of momentum strategies and proposes an industry momentum factor for performance evaluation. Finally, Section 7 concludes the paper.

1.2 Data Description and Industry Returns

Using the CRSP and COMPSTAT data files, 20 value-weighted industry portfolios are formed every month from July, 1963 to July, 1995. Two digit Standard Industrial Classification (SIC) codes are used to form industry portfolios in order to maximize coverage of NYSE, AMEX, and NASDAQ stocks, while maintaining a manageable number of industries, and insure each industry contains a large number of stocks. Table 1.1 provides a description and summary statistics

*intra-industry momentum patterns are virtually non-existent.*
of the industry portfolios. The average number of stocks per industry is 230, and the fewest number of stocks in any industry (except for Railroads) at any time is over 25. Therefore, virtually all portfolios are well-diversified (aside from industry effects).

Table 1.2 reports the average monthly raw excess and abnormal returns of the 20 industries relative to the Fama and French (1993) (FF) three factor model, Carhart (1997) four factor model, and Daniel, Grinblatt, Titman, and Wermers (1997) (DGTW) characteristic-based model, which are described in detail in Appendix 1.8. The cross-section of industry returns are largely captured by the models, providing little evidence that unconditional industry risk premia exist. However, the Gibbons, Ross, and Shanken (1989) (GRS) F-tests that abnormal returns are jointly zero relative to the CAPM (for the excess returns column). FF, Carhart, and DGTW models, are rejected at the 5% significance level, even under the more general assumption that returns are distributed elliptically. Thus, the effects of size, BE/ME, and momentum do not fully explain the cross-section of industry returns. Imposing industries may have important pricing implications independent from these factors.\(^7\)

\(^3\)In fact, using conservative Bonferroni confidence intervals, which account for the fact that we examine 20 not necessarily independent tests, we fail to reject all of the univariate tests that industry abnormal returns are significantly zero.

\(^6\)Gecz\'y (1997) shows that the adjustment to account for elliptically distributed returns entails multiplying the GRS F-statistic by a parameter of multivariate adjusted, excess kurtosis.

\(^7\)It should be noted that DGTW adjust their BE/ME measures for industry effects since Cohen and Polk (1995) demonstrate that differing accounting conventions across industries add noise to the relation between BE/ME and expected returns. Thus, their return adjustment procedure may capture industry returns slightly better since they refine the relation between BE/ME and average returns.
1.3 Motivation

1.3.1 Return Generating Process

Consider the following multifactor linear process for stock returns.

\[
\tilde{r}_{jt} = \mu_j + \sum_{k=1}^{K} \beta_{jk} \tilde{F}_{kt} + \sum_{m=1}^{M} \theta_{jm} \tilde{\delta}_{mt} + \tilde{\epsilon}_{jt}
\]  

(1.1)

where \(\tilde{r}_{jt}\) is the return of stock \(j\) at time \(t\), \(\mu_j\) is its unconditional mean. \(\tilde{F}_{kt}\) are demeaned portfolio returns where the unconditional mean of those returns are taken out (think of these as demeaned size and BE/ME portfolios, or the demeaned FF "factor" portfolios). \(\beta_{jk}\) are the portfolio sensitivities. \(\tilde{\delta}_{mt}\) are correlated components of returns across assets orthogonal to the \(K\) portfolios which we can think of as industry components, taking out the effects of market, size, and BE/ME. \(\theta_{jm}\) are the industry sensitivities, and \(\tilde{\epsilon}_{jt}\) are idiosyncratic components which are uncorrelated across assets.\(^8\) Note, this return generating process is not an equilibrium model of stock returns (as in Merton (1973)), nor is it necessarily a version of the Arbitrage Pricing Theory of Ross (1976). Simply, this return process is employed to illustrate the decomposition of momentum profits into components we already know are related to average returns (i.e., size and BE/ME) and components that might be contributing to the profitability of momentum strategies.

Since the \(K\) portfolios are demeaned, and industries do not appear to have mean returns significantly different from zero once the market, size, and BE/ME are taken into account, the \(K\) portfolios, industry components, and idiosyncratic terms have zero mean and, we assume for simplicity, that they are uncorrelated.

\(^8\)Of course, the \(\delta\)'s and \(\theta\)'s are changing over time, but since the time-variation in loadings is likely to be swamped by cross-sectional dispersion in loadings on market, size, and BE/ME, for ease of notation and illustration, we assume these coefficients are time invariant.
with each other. More formally,

\[ E[\tilde{F}_{kt}] = 0, \quad \forall k; \quad E[\tilde{\delta}_{mt}] = 0, \quad \forall m; \quad E[\tilde{\epsilon}_{jt}] = 0, \quad \forall j; \]
\[ E[\tilde{F}_{kt}\tilde{F}_{lt-1}] = 0, \quad \forall k \neq l; \quad E[\tilde{F}_{kt}\tilde{\delta}_{mt-1}] = 0, \quad \forall k, m; \quad E[\tilde{F}_{kt}\tilde{\epsilon}_{jt-1}] = 0, \quad \forall k, j; \]
\[ E[\tilde{\delta}_{mt}\tilde{\delta}_{nt-1}] = 0, \quad \forall m \neq n; \quad E[\tilde{\delta}_{mt}\tilde{\epsilon}_{jt-1}] = 0, \quad \forall m, j; \quad E[\tilde{\epsilon}_{jt}\tilde{\epsilon}_{it-1}] = 0, \quad \forall j \neq i. \]

Positive autocorrelation or momentum in returns implies stocks which outperformed the average stock in the past, will outperform the average stock in the future. That is,

\[ E[(\tilde{r}_{jt} - \bar{r}_t)(\tilde{r}_{jt-1} - \bar{r}_{t-1})] > 0. \]  \hspace{1cm} (1.2)

where \( \bar{r}_t \) is the cross-sectional or equal-weighted average return of stocks at time \( t \) (a bar over a variable represents its cross-sectional average in this section). Equivalently, equation (1.2) represents a self-financing momentum investment strategy where \((\tilde{r}_{jt-1} - \bar{r}_{t-1})\) is the amount invested in stock \( j \) at time \( t \), funded by shorting the same amount in the equal-weighted portfolio.\(^9\) Based on the process of stock returns, momentum profits can be decomposed as follows,

\[ E[(\tilde{r}_{jt} - \bar{r}_t)(\tilde{r}_{jt-1} - \bar{r}_{t-1})] = (\mu_j - \bar{r})^2 + \sum_{k=1}^{K} (\beta_{jk} - \bar{\beta}_k)^2 \text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1}) \]
\[ + \sum_{m=1}^{M} (\theta_{jm} - \bar{\theta}_m)^2 \text{Cov}(\tilde{\delta}_{mt}, \tilde{\delta}_{mt-1}) + \text{Cov}(\tilde{\epsilon}_{jt}, \tilde{\epsilon}_{jt-1}). \]  \hspace{1cm} (1.3)

Averaging over all \( N \) stocks, momentum trading profits equal,

\[ = \sigma_\mu^2 + \sum_{k=1}^{K} \sigma_{\beta_k}^2 \text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1}) + \sum_{m=1}^{M} \sigma_{\theta_m}^2 \text{Cov}(\tilde{\delta}_{mt}, \tilde{\delta}_{mt-1}) + \frac{1}{N} \sum_{j=1}^{N} \text{Cov}(\tilde{\epsilon}_{jt}, \tilde{\epsilon}_{jt-1}), \]  \hspace{1cm} (1.4)

\(^9\)This is similar to the investment strategies proposed by Lo and MacKinlay (1990), Lehman (1990), and Conrad and Kaul (1997). Although our investment strategies differ from these, ours closely follows those of JT, who document a 0.95 correlation between the profits generated from their strategy and the one specified in equation (1.2).
where $\sigma^2_\mu$, $\sigma^2_{\delta_k}$, and $\sigma^2_{\delta_m}$ represent the cross-sectional variances of mean returns, portfolio loadings, and industry sensitivities, respectively.

This decomposition allows us to compare the relative magnitudes of these sources of profits. Conrad and Kaul (1997) claim that $\sigma^2_\mu$ is the sole source of momentum trading profits, while JT claim that serial covariation in firm-specific components ($\tilde{\epsilon}_j$) are primarily responsible. We contend, however, that serial covariation in industry components ($\tilde{\delta}_m$) are generating the trading profits. Isolating the various sources of momentum profits, we can compare the relative magnitudes of the components in equation (1.4) and their contribution to profits.

1.4 Isolating Sources of Momentum Trading Profits

1.4.1 Momentum Investment Strategies

We form winners minus losers momentum investment portfolios similar to JT, by ranking stocks based on their prior $L$-month returns and forming a zero-cost portfolio of the highest past $L$-month return stocks funded by shorting a portfolio of low past return stocks, holding these positions over the next $H$ months. Since JT focus much of their analysis on the $L = 6$ month lagged, $H = 6$ month holding period strategy, for brevity and ease of comparison, we will do the same.\textsuperscript{10} For example, the 6-month, 6-month strategy at time $t$ entails ranking stocks based on their $t - 6$ to $t - 1$ returns, and computing the value-weighted return of the highest 30% of stocks every month from $t$ to $t + 5$ minus the value-weighted return of the lowest 30% of stocks every month from $t$ to $t + 5$. That is, we maintain our selection of winning and losing stocks over the next 6 months. This

\textsuperscript{10}The results for a host of trading strategies were very similar, and are available from the author upon request.
procedure is then repeated at time $t + 1$, and so on. At time $t$, we are long the winner portfolio (consisting of 6 equally weighted strategies: the portfolio of winning stocks selected at times $t - 5, \ldots, t$) and short the loser portfolio (the equal-weighted average of the 6 losing stock portfolios chosen at times $t - 5, \ldots, t$). At the same time, we close out the position undertaken in month $t - 6$. Therefore, only 1/6 of the securities in the entire strategy are being revised every month. Since positions in the winners and losers portfolios are held constant over the $H$-month holding period, but one month returns are being computed, we do not need to correct for autocorrelation induced by overlapping holding period returns.

1.4.1.1 Raw Profits

Our trading strategy is similar to JT, except we employ 30% breakpoints rather than deciles (10% breakpoints) to identify winners and losers, and value-weight rather than equal-weight the securities within each of the winning and losing portfolios. The 30% breakpoints are employed to mitigate the influence of outliers on trading profits, and to increase the diversification of the portfolios, while value-weighting is employed because the DGTW return adjustment (used later) is based on matching stocks with value-weighted benchmarks. In addition, the use of value-weights weakens the influence of the size effect and diminishes microstructure influences on profits. These modifications, not surprisingly, impact the momentum profits. JT document an annual 12% return per dollar long for their 6-month, 6-month strategy, while our strategy (shown in Table 1.3) generates about 6%. The statistical significance of our strategy, however, is much stronger, indicating the variance of our strategy is lower. Equal-weighting stocks within the 30% categories (rebalancing monthly) produces profits of about 9.3% per year, indicating that value-weighting is quite important. For most of the tests
in the paper, value-weighting is employed.

1.4.1.2 Serial Covariation in the Factor Portfolios

Since the equal-weighted portfolio has negligible idiosyncratic risk, and since the equal-weighted average loading on $\delta_m (\bar{\theta}_m)$ will be very small $\forall m \in M$. the serial covariance of the equal-weighted portfolio can be expressed as,

$$\text{Cov}(\bar{r}_t, \bar{r}_{t-1}) = \sum_{k=1}^{K} \beta_k^2 \text{Cov}(\bar{F}_{kt}, \bar{F}_{kt-1})$$ (1.5)

where $t$ is a 6-month period. Therefore, we can isolate one of the components of momentum profits. Using raw equal-weighted market portfolio returns. we find, in accordance with JT. that $\text{Cov}(\bar{r}_t, \bar{r}_{t-1}) = -0.0001$. suggesting that serial covariation in the FF factor portfolios is not contributing to momentum profits. In addition, the serial covariance of each of the three FF factor-mimicking portfolios is $\text{Cov}([Mkt - r_f]_t, [Mkt - r_f]_{t-1}) = -0.00008$, $\text{Cov}(SMB_t, SMB_{t-1}) = 0.00007$, and $\text{Cov}(HML_t, HML_{t-1}) = 0.00004$, none of which are significantly different from zero. Furthermore, employing momentum strategies on the FF factor portfolios (by investing in the factor-mimicking portfolio which had the highest prior return and shorting the factor which had the lowest), produces negative profits of -0.0005 (t-stat = -0.42). Thus, serial covariation in these portfolios does not seem to be driving momentum trading profits.\(^{11}\)

1.4.1.3 Industry Momentum Profits

Aggregating stocks into industry portfolios largely eliminates firm-specific components ($\bar{\epsilon}$) of returns, since industries contain over 230 stocks on average. In addition, the cross-sectional variance of mean industry returns, $\sigma^2_{\mu_i}$, is only 0.00083.

\(^{11}\)Grundy and Martin (1997) similarly find no positive autocorrelation among the FF factors, and conclude, based on this, that risk-based explanations are not driving momentum.
compared to the estimated cross-sectional dispersion of mean stock returns of 0.011. and our failure to reject an F-test that mean industry returns are equal (F-stat = 0.825. with a p-value of 0.677) suggests that the cross-sectional dispersion in industry returns is small. Thus, the existence of industry momentum provides evidence against both Conrad and Kaul’s (1997) claim that dispersion in mean returns drives momentum profits and JT’s claim that firm-specific components drive momentum. Referring back to our model of returns (2.70), industry momentum trading profits can be expressed as.

\[ \frac{1}{20} \sum_{t=1}^{20} \text{E}[(\tilde{R}_{it} - \bar{r}_t)(\tilde{R}_{it-1} - \bar{r}_{t-1})] = \sigma_{\mu_t}^2 + \sum_{k=1}^{K} \sigma_{3f_{k}}^2 \text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1}) \]

\[ + \sum_{m=1}^{M} \sigma_{\delta_{tm}}^2 \text{Cov}(\tilde{\delta}_{mt}, \tilde{\delta}_{mt-1}), \]

where \( \tilde{R}_{it} \) is the return on industry \( I \) at time \( t \). Since \( \sigma_{\mu_t}^2 \) is small and \( \text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1}) < 0 \) (as demonstrated previously), the existence of industry momentum profits implies.

\[ \sum_{m=1}^{M} \sigma_{\delta_{tm}}^2 \text{Cov}(\tilde{\delta}_{mt}, \tilde{\delta}_{mt-1}) > 0. \]  

(1.7)

Sorting industry portfolios based on their past 6-month returns, and investing in the top three industries, while shorting the bottom three industries (holding this position for 6 months) produces average monthly profits (shown in Table 1.3) of 0.43%, identical in magnitude to those obtained from the momentum strategy for individual equities. Since the cross-sectional variation in mean industry returns is much smaller than that for individual stock returns, under Conrad and Kaul’s (1997) hypothesis we would expect industry momentum profits to be significantly smaller than those for individual equities. This, however, is not the case.\(^{12}\) In addition, standard errors on the profits are equally high, indicating

\(^{12}\)Of course, these profits are still only half those documented in JT. Again, however, this discrepancy is due to value-weighting rather than equal-weighting stocks as in JT. Although the
that the risks of both strategies are similar. Thus, industry momentum appears as strong as individual stock momentum, consistent with the hypothesis that industry factors generate momentum trading profits.

1.4.1.4 Size and BE/ME Adjusted Profits

Certainly, failing to control for the unconditional mean returns of stocks will exaggerate profits, however, we contend that cross-sectional variation in mean returns does not account for the entirety of momentum profits. Thus, controlling for $\sigma^2$, significant momentum profits should remain. To isolate the momentum effect, we adjust returns for the effects of size and BE/ME by first sorting stocks into size quintiles, and then within each size quintile, sorting stocks into BE/ME quintiles, where stocks are value-weighted within these groups. Stock $j$ is then matched with one of the 25 portfolios based on its size and BE/ME at time $t - 1$, and the return of the matched portfolio is subtracted from stock $j$'s return at time $t$.\(^\text{13}\) The resulting size and BE/ME characteristic-adjusted return is defined as,

$$\bar{r}_{jt} = \bar{r}_{jt} - \bar{R}_{t}^{SB_{j,t-1}}$$

where $\bar{r}_{jt}$ is the return on security $j$, and $\bar{R}_{t}^{SB_{j,t-1}}$ is the return on the size and BE/ME matched portfolio. Adjusting for size and BE/ME in this manner accounts for almost 83% of the cross-sectional variation in mean returns. Yet, as Table 1.3 shows, the momentum 6-month, 6-month strategy for holding period

industries themselves are equally weighted in the winning and losing portfolios. The industry portfolios are value-weighted sums of stock returns. When we employ equal-weighted industry portfolios in the industry momentum strategies, we obtain an average profit of .0081 per month or 10.2% per year (t-stat=7.71), which is about 90 basis points higher than the equal-weighted momentum strategy for individual equities. When examining the strongest industry momentum strategy, the 1-month, 1-month short-term strategy, we obtain profits over 12% per year, as large as those in JT. We will discuss these strategies in more detail in the next section.

\(^\text{13}\)This is similar to the DGTW adjustment method, but here returns are only adjusted for size and BE/ME effects, leaving momentum unaccounted for.
returns adjusted for size and BE/ME\textsuperscript{14} remain strong, producing abnormal mean profits of 0.29\% per month (with a highly significant t-stat of 3.34). which is about two thirds the size of the raw profits. Since size and BE/ME account for 83\% of the cross-sectional variation in returns, yet explain less than one third of momentum profits, cross-sectional dispersion in mean returns is not likely the source.

1.4.1.5 DGTW Adjusted Profits

Adjusting returns for size and BE/ME accounts for a large percentage of the dispersion in mean returns and are frequently cited as sources of systematic risk (see Fama and French (1993, 1996) and references therein). Therefore, we can view this adjustment as essentially “de-meaning” returns or accounting for their unconditional means and sensitivities to common factors. Furthermore, as just shown, this adjustment does not alleviate momentum profits. However, if industry components (δ’s) are the true momentum factors, then sorting stocks based on past returns will be a noisy measure of momentum. Therefore, matching stocks ‘naively’ with a portfolio of stocks with the same past return, will improperly adjust returns for momentum, although it may accurately adjust the extreme (highest and lowest past return) stocks. Controlling for unconditional mean returns, since the highest (lowest) past return stocks likely had high (low) realizations on both their industry component (δ) and their firm-specific component (ε), benchmark portfolios formed from past returns will correctly adjust the highest and lowest past return stocks for momentum, implying

\[
E[\tilde{r}_{jt} | \tilde{r}_{jt-1} \gg \tilde{r}_{t-1}] = 0 = E[\tilde{r}_{jt} | \tilde{r}_{jt-1} \ll \tilde{r}_{t-1}],
\]  

\text{(1.9)}

\textsuperscript{14}Prior raw returns are always used to form portfolios, so that the selection of stocks into the ‘winners’ and ‘losers’ categories is the same.
where \( \bar{r}_{jt}^* \) is the 'naively' momentum adjusted return of stock \( j \). Thus, adjusting returns by matching stocks 'naively' with similar momentum benchmarks should adjust the highest and lowest stocks for momentum, even though it may be an improper benchmark for momentum. The DGTW return adjustment method accomplishes this by controlling for the effects of size and BE/ME and then matching stocks 'naively' with portfolios of similar past returns. Table 1.3 reports the DGTW adjusted profits, which are not significantly different from zero, implying,

\[
E[(\bar{r}_{jt}^*)(\bar{r}_{jt-1} - \bar{r}_{t-1})] = 0. \tag{1.10}
\]

where \( \bar{r}_{jt}^* \) is the DGTW adjusted return and \( \bar{r}_t^* \approx 0 \) (\( \bar{r}_t^* = 0.0003 \) with a t-stat of 0.31). Since the equal-weighted portfolio exhibits no DGTW adjusted return, we conclude that the DGTW adjustment eliminates the premia associated with size and BE/ME, effectively removing the influence of the \( K \) portfolios, \( \bar{F}_k \). Furthermore, the DGTW return adjustment accounts for almost 90% of the cross-sectional variation in asset returns, thus capturing the influence of \( \sigma^2_\mu \) on profits as well.

However, since only the extreme stocks will be properly adjusted with 'naive' past return benchmarks, employing these 'naive' benchmarks on industry returns should not eliminate industry momentum profits. That is,

\[
\frac{1}{20} \sum_{i=1}^{20} E[(\bar{R}_{it}^* - \bar{r}_t^*)(\bar{R}_{it-1} - \bar{r}_{t-1})] = \sigma^2_\mu_i + \sum_{k=1}^{K} \sigma^2_{\beta_{ik}} \text{Cov}(\bar{F}_{kt}, \bar{F}_{kt-1}) \tag{1.11} \\
+ \sum_{m=1}^{M} \sigma^2_{\delta_{it}} \text{Cov}(\delta_{mt}, \delta_{mt-1}) > 0.
\]

where \( \bar{R}_{it}^* \) is the industry return composed of a weighted sum of \( \bar{r}_{jt}^* \)'s for all \( j \in I \). If \( \sigma^2_\mu_i \) and \( \sigma^2_{\beta_{it}} \) are small, then the last component is significantly positive.

Aggregating the individual DGTW adjusted stock returns, \( \bar{r}_{jt}^* \)'s, into industry portfolios, as shown in Table 1.3, the industry momentum profits are still
significant, producing average monthly profits of 0.20% (t-stat = 2.27). Thus, ‘naive’ past return benchmarks do not account for industry momentum profits, consistent with industry factors generating momentum. In addition, since raw industry returns exhibit little cross-sectional variation on average, and adjusting industries via DGTW accounts for over 90% of this variation, this provides an even stronger test for whether dispersion in mean returns is responsible for the profits. Furthermore, adjusting for size and BE/ME eliminates industry sensitivity to these common “factors” (which we know are not positively autocorrelated anyway). Finally, since industries also diversify away idiosyncratic components of returns, only serial covariation in the industry factors remains as a possible source of profits. Thus,

$$\frac{1}{20} \sum_{t=1}^{20} E[(\tilde{R}_{tm}^t)(\tilde{R}_{t-1} - \bar{r}_{t-1})] = \sum_{m=1}^{M} \sigma_{\bar{r}_{tm}}^2 \text{Cov}(\delta_{mt}, \delta_{mt-1}) > 0. \quad (1.12)$$

indicating that serial covariation in industry components is generating significant profits. This provides further evidence against both Conrad and Kaul’s (1997) hypothesis that cross-sectional dispersion in mean returns and JT’s claim that serial correlation in idiosyncratic components accounts for the entirety of momentum trading profits.

### 1.4.1.6 Industry Adjusted Profits

If industry factors drive momentum, then industry components should explain momentum in common stock returns. Matching stocks with industry benchmarks implies.

$$\frac{1}{N} \sum_{j=1}^{N} E[(\tilde{r}_{jt} - \tilde{R}_{jt})(\tilde{r}_{jt-1} - \bar{r}_{t-1})] = \sigma_{\mu}^2 + \sum_{k=1}^{K} \sigma_{\bar{r}_{kj}}^2 \text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1}) + \frac{1}{N} \sum_{j=1}^{N} \text{Cov}(\tilde{e}_{jt}, \tilde{e}_{jt-1}) \quad (1.13)$$

Thus, controlling for $\sigma_{\mu}^2$ and $\sigma_{\bar{r}_{kj}}^2$, if (1.13) is positive, then serial covariation in idiosyncratic components generates significant profits.
Two procedures for removing industry effects from individual stock returns are employed. The first simply subtracts the contemporaneous industry portfolio return, adjusted for size and BE/ME effects, from the return of each individual security in that industry (also adjusted for size and BE/ME effects). Thus, size, BE/ME, and industry adjusted returns are defined as,

\[ \tilde{\tilde{r}}_{jt}^{sb,I} = \tilde{r}_{jt}^{sb} - \tilde{R}_{I,t}^{sb} \quad \text{for } j \in I \]  

(1.14)

where \( \tilde{R}_{I,t}^{sb} \) is the size and BE/ME adjusted return on industry \( I \), to which stock \( j \) belongs at time \( t \).

In addition, we compute “pure industry” factors similar to Roll (1992), Heston and Rouwenhorst (1994), and Griffin and Karolyi (1995), by running a weighted least squares cross-sectional regression, at each point in time, of individual stock returns on 20 industry dummy variables.

\[ r_j = \alpha + \sum_{I=1}^{20} \gamma_I D_{I,j} + e_j \]  

(1.15)

where \( D_{I,j} \) is an industry dummy equal to one if stock \( j \) is in industry \( I \) and zero otherwise. Because each stock is in one industry, there is an identification problem when we use all 20 industry dummies, resulting in perfect multicollinearity between the regressors. Rather than choosing an arbitrary industry as the benchmark, however, and interpreting the regression coefficients as deviations from the benchmark industry, we employ the restriction that the value-weighted sum of the industry coefficients equal zero.

\[ \sum_{I=1}^{20} w_I \gamma_I = 0. \]  

(1.16)

where \( w_I \) is the value weight of industry \( I \) in the market portfolio.

Equation (1.15) is estimated subject to restriction (1.16) via weighted least squares, where the weights are the market capitalization of each stock. The
intercept is interpreted as the value-weighted market, and industry dummy coefficients ($\gamma$) as deviations from the value-weighted market or "pure" industry effects. Running this regression every month produces a time-series of pure industry returns, $\hat{\gamma}_{it}$. Thus, size, BE/ME, and pure industry-adjusted returns are defined as:

\[ \tilde{r}_{jt}^{sb, Pl} = \tilde{r}_{jt}^{sb} - \hat{\gamma}_{it}, \quad \text{for } j \in I \]  \hspace{1cm} (1.17)

where $\hat{\gamma}_{it}$ is the pure industry $I$ return (i.e., estimated coefficient from the cross-sectional regression) to which stock $j$ belongs at time $t$.

Table 1.3 reports the results for the 6-month, 6-month momentum trading strategies. As the table shows, momentum in individual stock returns is eliminated when returns are adjusted for industry effects or pure industry effects, and the adjustment appears to be even better than DGTW.

Referring back to our decomposition of momentum profits in equation (1.4), adjusting for size and BE/ME effects accounts for a large portion of the cross-sectional variation in returns ($\sigma_{\mu}^2$), and are frequently cited as sources of systematic risk. Thus, $\sigma_{\mu}^2$ and $\text{Cov}(\tilde{F}_{kt}, \tilde{F}_{kt-1})$ are eliminated as sources of momentum profits. In addition, adjusting for industry effects eliminates the third component of momentum profits, leaving serial covariation in residual terms as the only potential source of profits. Therefore, the unprofitability of industry-adjusted momentum trading strategies implies,

\[ E[(\tilde{r}_{jt}^{sb} - \tilde{R}_{jt}) (\tilde{r}_{jt-1} - \tilde{r}_{t-1})] = \text{Cov}(\tilde{\epsilon}_{jt}, \tilde{\epsilon}_{jt-1}) = 0. \]  \hspace{1cm} (1.18)

indicating that serial correlation in firm idiosyncracies is not a source of momentum profits. Thus, industry components seem to be driving momentum.
1.4.1.7 ‘Random’ Industry Portfolios

‘Random’ industry portfolios are formed by replacing every true stock in industry $I$ with another stock that had the same past return, but may or may not be in industry $I$. That is, $\forall j \in I$, we replace stock $j$ with stock $i$ such that $\tilde{r}_{it-1} = \tilde{r}_{jt-1}$. The probability that $i \in I$ can be expressed as:

$$P(i \in I|\tilde{r}_{it-1} = \tilde{r}_{jt-1}, i \neq j).$$

The expected return on this ‘replacement’ stock $i$ is,

$$E[\tilde{r}_{it}|\tilde{r}_{it-1} = \tilde{r}_{jt-1}, i \neq j],$$

and the expected return on the ‘random’ industry is therefore.

$$E[\tilde{R}_{it}] = \sum_{j \in I} w_j E[\tilde{r}_{it}|\tilde{r}_{it-1} = \tilde{r}_{jt-1}, i \neq j], \quad \sum_{j \in I} w_j = 1. \quad (1.21)$$

Since, ‘random’ industries contain stocks from various industries, they will contain a small fraction of stocks from the true industry, and therefore will not exhibit momentum if industry components are the true momentum factors. However, since ‘random’ industries have the same past returns as every true stock in the industry, they will exhibit momentum if other components drive momentum. Consequently, adjusting returns via ‘random’ industry benchmarks will not eliminate momentum trading profits in either individual stocks or in industries if industry factors drive momentum.

To illustrate this, if we control for the unconditional mean returns of stocks and their sensitivity to common factors (i.e., the FF factors), then the conditional expectation of $\tilde{r}_{it}$ will depend on the dispersion of $\tilde{\delta}$ and $\tilde{\epsilon}$ components. To see this, assume both $\tilde{\delta}$ and $\tilde{\epsilon}$ are distributed normally with mean zero and some variance (i.e., $\tilde{\delta} \sim N(0, \sigma_\delta^2)$ and $\tilde{\epsilon} \sim N(0, \sigma_\epsilon^2)$), and assume exposure to an industry factor is normalized to one if the stock is in the industry and zero otherwise. If the $\tilde{\delta}$'s are serially correlated, but the $\tilde{\epsilon}$'s are not (i.e., our model), then the expected return
of a candidate replacement stock is $\tilde{\delta}$ with some probability and zero otherwise. That is, conditional on the two stocks having the same momentum attribute (i.e., $\tilde{\delta} + \tilde{\epsilon}$), the expected return of the candidate replacement is

$$E[\tilde{\tau}_t|\delta_{t-1} + \epsilon_{t-1}] = E[\tilde{\delta}_t|\delta_{t-1} + \epsilon_{t-1}],$$

(1.22)

and assuming $\tilde{\delta}$ is perfectly serially correlated.

$$= E[\tilde{\delta}|\tilde{\delta} + \tilde{\epsilon}] = \frac{\sigma_{\tilde{\delta}}^2}{\sigma_{\tilde{\delta}}^2 + \sigma_{\tilde{\epsilon}}^2} [\tilde{\delta} + \tilde{\epsilon}].$$

(1.23)

As $\frac{\sigma_{\tilde{\delta}}^2}{\sigma_{\tilde{\epsilon}}^2} \rightarrow 0$. $\Rightarrow E[\tilde{\tau}_t|\delta_{t-1} + \epsilon_{t-1}] \rightarrow 0.$

(1.24)

Thus, when the cross-sectional variation in idiosyncratic components is much larger than that of industry components, the expected return on the candidate replacement stock approaches zero. Or, to put this another way, the probability that the past return of the replacement stock was due to $\tilde{\epsilon}$ rather than $\tilde{\delta}$ is large, and thus the expected return on the replacement stock is closer to zero, since, under our model, there is no autocorrelation in firm-specific disturbances. So, if dispersion in idiosyncratic components swamps that of industry components, then ‘random’ industry portfolios will not exhibit momentum and will not provide adequate benchmarks to adjust returns for momentum.

Under the JT hypothesis, however, there is positive autocorrelation in the $\epsilon$’s, but not the $\delta$’s. Since only $\tilde{\epsilon}$’s are serially correlated in this case, the expected return on the candidate replacement stock is the conditional expectation of $\tilde{\epsilon}$:

$$E[\tilde{\epsilon}|\tilde{\delta} + \tilde{\epsilon}] = \frac{\sigma_{\tilde{\epsilon}}^2}{\sigma_{\tilde{\delta}}^2 + \sigma_{\tilde{\epsilon}}^2} [\tilde{\delta} + \tilde{\epsilon}].$$

(1.25)

Thus, if cross-sectional variation in idiosyncratic components swamps dispersion in industry components, the expected return of the candidate replacement stock approaches the past return. In other words, if firm-specific disturbances are serially correlated, and if their cross-sectional variation is much greater than
that of industry components. random industries will primarily be comprised of stocks with the same prior $\tilde{\epsilon}$ realization as opposed to $\tilde{\delta}$. and will subsequently exhibit the same return next period. Therefore, the formation of industry and ‘random’ industry portfolios allows us to directly compare the JT hypothesis of serial covariation in idiosyncratic terms with our hypothesis of serial covariation in industry factors driving momentum.

Ranking all stocks in ascending order based on their prior 6-month returns, we form ‘random’ industry portfolios by replacing each stock in an industry with a stock that has the next highest momentum characteristic (6-month prior return) to that stock (and may or may not be in the same industry). In this way, ‘random’ industry portfolios have the same momentum attributes as the true industry, but contain stocks from various industries. For example, given $N$ stocks ranked in ascending order based on 6-month prior returns, stock $j$ belonging to industry $I$ is replaced with stock $j + 1$‘s return for all $j = 1, ..., N$. We also form ‘random’ industries by replacing stock $j$‘s return with stock $j - 1$‘s return (i.e., replace each stock with the stock ranked below it), and by replacing stock $j$‘s return with an equal-weighted return of the stocks ranked above and below it (i.e., replace $\bar{r}_{jt}$ with $\frac{\bar{r}_{j+1,t} + \bar{r}_{j-1,t}}{2}$).\textsuperscript{15}

Since there are many more stocks than industries, the cross-sectional dispersion in industry returns is much smaller than the cross-sectional dispersion in idiosyncratic terms. In fact, defining $\tilde{\xi}_j$ to be the residual return after taking out size, BE/ME, and industry effects (i.e., $\tilde{\xi}_{jt} = \bar{r}_{jt} - \bar{r}_{jt}^{sb,I}$), and defining $\tilde{\delta}$ to be the industry return adjusted for size and BE/ME effects (i.e., $\tilde{R}_{it}^{sb}$), the average monthly cross-sectional variance of idiosyncratic components ($\tilde{\sigma}_e^2$) over the 32 year period is 0.015964. compared to the average cross-sectional variance of industry

\textsuperscript{15To avoid endpoint problems, we simply replace stock $N$ with stock $N - 1$ and stock 1 with stock 2.}
components \((\sigma^2)\) of only 0.000424. a magnitude almost 38 times as large. Thus, \(\sigma^2\) is much smaller than \(\sigma^2\), suggesting that random industries will contain stocks with similar \(\hat{c}\) realizations rather than stocks from the same industry. In fact, the percentage of replacement stocks that are in the ‘correct’ industry is only 12.85% on average, and ranged from 9.25% to as much as 18.23% over the entire sample period. Therefore, from equation (1.24), random industry portfolios should not exhibit momentum and should be poor benchmarks for momentum if industry components are serially correlated, but the reverse will be true if idiosyncratic components drive momentum.

Table 1.3 reports the momentum profits from the random industries and from individual stocks adjusted via the random industry benchmarks (i.e., \(\bar{r}^{gb} = \bar{r}_{jt} - \bar{R}_{gt}\)). As shown in the table, momentum profits are non-existent for the random industries, and momentum profits for individual stocks are virtually unaltered by the random industry adjustment, consistent with the true industry being the important component behind momentum profits.

### 1.4.1.8 Industry Neutral Portfolios

Finally, we create two zero-cost portfolios similar to Carhart’s (1997) \(PR1YR\) individual stock momentum factor (see Appendix 2.7), where stocks are first sorted on past 6-month returns (from \(t-6\) to \(t-1\)) and the equal-weighted average return of the top 30% of stocks minus the bottom 30% of stocks is computed at time \(t\). The difference here is that (1) the 30% and 70% breakpoints used to form this portfolio are set within each industry, and (2) stocks are sorted based

\(^{16}\)For brevity, we only report results for the random industries generated by replacing each stock return with an equal-weighted average of the stocks ranked above and below it based on past returns. The results for other random industries were similar and are available from the author upon request.
on their past 6-month return in excess of their industry average over the same time period. We refer to the first portfolio as an ‘industry neutral’ portfolio, since low past return stocks are subtracted from high past return stocks within the same industry. Thus, the industry effect is neutralized from returns. As Table 1.3 demonstrates, the industry neutral portfolio produces mean profits of 0.0003 with an insignificant test statistic of 0.20, indicating, again, that once we account for industry effects, momentum in individual equities is virtually nonexistent. We refer to the second portfolio as an ‘excess industry’ portfolio, since we select our winning and losing stocks based on their past returns in excess of the industry benchmark. That is, we choose stocks that had the highest and lowest realizations in the past, independent from the influences of their industry. Thus, we do not expect to generate momentum profits if industry factors drive momentum. As Table 1.3 shows, the excess industry portfolio does not exhibit significant profits (mean = 0.0008, t-stat = 0.39), consistent with industry factors being the primary source of momentum.17

Finally, if past returns are a noisy measure of momentum, and industry components are a better estimate of true momentum, then going long the worst past performing stocks within the best past performing industries, and short the best past performing stocks within the worst past performing industries, should still yield significant profits. Conditioning on being in the three industries which performed the best over the last 6 months, we rank stocks within each of these industries based on their prior 6-month return, and form an equal-weighted portfolio of the bottom 30% of stocks within each of these three high past performing industries. Likewise, we form an equal-weighted portfolio of the top 30% of past 6-month return stocks belonging to each of the three worst performing indus-

17Note that ‘industry neutral’ and ‘excess industry’ portfolios use raw returns when computing profits. Thus, the insignificance of these two strategies is even more impressive.
tries. and subtract this portfolio return from the previous one. In other words, we are long the losers from the winning industries. and short the winners from the losing industries. This zero-cost portfolio should exhibit significant profits if industry effects drive momentum. and should produce significant negative profits if individual stock returns better identify momentum. As Table 1.3 shows, this portfolio produces positive and significant profits of 0.19% per month. indicating the importance of industries in explaining momentum in stock returns.

1.5 Robustness of Industry Momentum Strategies

Analysis of industry abnormal returns has thus far only pertained to the unconditional mean of those returns over time. We contend, however. that industries may have important pricing implications conditionally. That is. the unconditional mean abnormal return of an industry may be zero over time. yet still may have significant explanatory power over shorter time periods. Further examining the conditional first moments of industry returns and the robustness of industry momentum strategies is. therefore. the focus of this section of the paper.

1.5.1 Industry Momentum Trading Strategies

One of the puzzling features of momentum in individual equities is that it only appears over intermediate investment horizons (3-12 months), while returns tend to mean revert over the short-term (< 3 months) and long-term (3-5 years). Behavioralists reconcile this phenomenon by claiming investors overreact to short-

---

18 This may be the case when industries grow over the prior sub-period to where they comprise a significant portion of the economy, thus imposing non-diversifiable influences on asset prices. Or it may be the case that high (low) past return industries load higher (lower) on systematic risk factors in the subsequent period. Moskowitz (1997b) explores these issues and finds evidence of the latter.
period news, underreact to intermediate term news (see JT), and overreact, once again, to long-term information (see DeBondt and Thaler (1985)). However, since industry factors seem to drive momentum, we might expect short-term positive autocorrelation in returns to exist if these industry factors are related to risk. Therefore, this section compares industry momentum patterns with those for individual stocks, and determines the robustness of industry momentum strategies.

Ranking the 20 industries based on their \( L(= 1, 6, 12, 24) \)-month lagged returns, we form portfolios of the highest and lowest past performing industries, holding them for \( H(= 1, 6, 12, 24, 36) \) months, rebalancing monthly. Three sets of industry momentum trading strategies are employed: IMS1 merely invests in the highest past return industry and shorts the lowest past return industry; IMS2 computes the equal-weighted return of the three highest past return industries and subtracts the equal-weighted return of the three lowest past return industries; and IMS3 computes the equal-weighted return of all industries with past returns greater than the cross-sectional mean prior \( L \) month return and subtracts the equal-weighted return of all industries with past returns less than the average prior return.\(^{19}\) The latter two strategies are employed to determine if momentum is present in all industries, or if it is merely driven by outliers (i.e., one or two “hot” or “distressed” industries). The results from the industry momentum trading strategies are reported in Table 1.4.

The three industry momentum strategies all produce similar patterns in returns. IMS1 generates the highest mean returns over intermediate investment horizons, and lowest mean returns over longer horizons (24 and 36 months). As we include more industries in our ‘winning’ and ‘losing’ portfolios, via strategies IMS2 and IMS3, the profits from the zero-cost portfolios decline, but are still

\(^{19}\)The industries are given equal weight in the industry momentum strategies, but stocks within each industry are value-weighted.
strong. This suggests that momentum in industries is not driven by just the highest and lowest past performing industries, but is present in all industries. Furthermore, the reduction in the mean returns of the IMS2 and IMS3 strategies relative to IMS1 is consistent with the fact that the better (worse) an industry performed in the past, the better (worse) it will perform subsequently over intermediate time horizons. Likewise, the better (worse) an industry performed in the past, the worse (better) it will perform over long time horizons. This is consistent with the findings of JT for individual stock returns, where momentum profits are strong over intermediate holding periods (3-12 months), but diminish beyond a year.

1.5.2 The 1-Month, 1-Month Strategy and the Lead-Lag Effect

However, unlike the profits from individual equity momentum, the strongest industry momentum strategy is the 1-month lagged, 1-month holding period strategy. This is contrary to the findings of Jegadeesh (1990), who documents short-term return reversals in individual stocks, but is consistent with a risk premium explanation for industry momentum. That is, if time-varying industry risk premia exist, then last month’s return will be a strong predictor of the return next period. In addition, if time-variation in industry returns is simply a function of the behavior of individual stock returns, then short-term industry return reversals should be present. Thus, industry return patterns appear somewhat unique, having a component independent from patterns generated by individual stocks.

1.5.2.1 Bid-Ask Bounce

Another possibility, however, is that the one month return reversal for individual stocks is due to microstructure effects (such as bid-ask bounce), which are alle-
viated by forming industry portfolios. To test this conjecture, we skip a month before computing the 1-month holding period returns, so that industries and stocks are selected based on their prior returns from the end of month $t - 3$ to $t - 2$, while holding period returns are computed in month $t$. Industry momentum persists and is quite strong, producing mean raw profits of 0.75% ($t$-stat = 3.89), and DGTW adjusted profits of 0.57% ($t$-stat = 3.33). Reversal effects for individual stocks are no longer prevalent, however, as the relative strength strategy produces positive profits of 0.25% ($t$-stat = 1.41). Thus, microstructure effects may be partly responsible for the short-term reversal effect in individual stocks, although Jegadeesh (1990) finds that such effects do not explain the reversal return effect entirely. Thus, other explanations such as overreaction to short-term information may hold for very short investment horizons, particularly if the profitability of short horizon trading strategies is limited by the tradability of assets and the high turnover such strategies require.

1.5.2.2 Lead-Lag Effect

Alternatively, we may be picking up the well-known lead-lag effect that large stock returns lead small stock returns for as much as several months. However, since industry portfolios are value-weighted, lead-lag effects should largely be alleviated. In addition, the random industries, which are also value-weighted, do not exhibit momentum, even for the 1-month, 1-month strategy (Table 1.6), and forming random industries by first ranking stocks based on size, and replacing stocks with securities of similar size also generates insignificant 1-month, 1-month

---

20 Examining further the influence of microstructure on momentum and contrarian investment strategies is beyond the scope of this paper. For further discussion of these issues, see Conrad and Kaul (1989), Jegadeesh and Titman (1995a), Kaul and Nimalendran (1990), and Lo and MacKinlay (1990).

momentum profits of 0.21% (t-stat = 0.92).\textsuperscript{22}

However, if intra-industry lead-lag effects are much stronger than inter-industry lead-lag effects, then examining the momentum profits of the 'random' industries is not a fair comparison.\textsuperscript{23} To test this conjecture, we compute 1-month, 1-month momentum strategies on individual stocks within each industry and compare the profits to those from 1-month, 1-month individual stock strategies computed within each 'random' industry. Since 'random' industries contain stocks from various industries, if the intra-industry lead-lag relation is stronger than inter-industry lead-lag effects, then the 1-month, 1-month strategy within industries should produce stronger return reversals (i.e., more negative momentum profits) than the same strategy within 'random' industries. Specifically, we sort stocks within each industry based on their prior month return and set 30% and 70% breakpoints within each industry. We then go long the highest 30% prior 1 month return stocks and short the lowest 30%, both value-weighting and equal-weighting stocks within these categories. This strategy is similar to our 'industry-neutral' strategy described in Section 1.4.1.8, except for the 1-month, 1-month strategy. The same procedure is conducted for our 'random' industries, where stocks are first assigned to a random industry by replacing every true stock in industry \( I \) with the stock ranked above it based on last month's return. Again, 30% and 70% breakpoints are set within each random industry and both value-weighted and equal-weighted winners minus losers momentum profits are computed. As Table 1.5 shows, the 1-month, 1-month profits within the true industries are significantly stronger (i.e., more negative) than those within random industries.

\textsuperscript{22}Interestingly, the percentage of replacement stocks in the random industry portfolios of similar size that are from the same industry is 11.44% and ranges from 8.46% to 15.02% over the sample period. This percentage is less than that for random industries of similar past returns, indicating that past returns and industry effects are more highly correlated than firm size and industry.

\textsuperscript{23}Thanks to Sheridan Titman for pointing this out.
indicating that lead-lag effects are magnified within industries. In other words, the Jegadeesh (1990) short-term reversal effect is more prevalent within industries. In addition, we also create ‘random’ industries by replacing every true stock in industry I with the stock ranked above it based on last month’s market capitalization (size). The resulting 1-month, 1-month profits within these ‘size generated random’ industry portfolios. shown in Table 1.5. are also significantly less negative than the intra-industry profits.\textsuperscript{24}

Thus, our earlier comparison of ‘random’ industry 1-month, 1-month profits to the true industry profits is not a fair comparison, and therefore does not alleviate concerns about the lead-lag effect driving these profits. In order to address the impact of the lead-lag relation on our 1-month, 1-month industry strategy, therefore, we decompose the IMS2 1-month, 1-month profits into components related to size. Specifically, at time $t$, we select the three industries which performed the best and the three which performed the worst over the past month. Then, within each industry, we sort stocks into quintiles based on their market capitalization at time $t - 1$. Instead of computing the return of the three high past performing industries minus the three low past performing industries, we now compute the return of the highest size quintiles within each of the three high past performing industries, and subtract the return of the highest size quintiles within each of the three lowest past performing industries. That is, we restrict securities to be the largest 20% of stocks within each industry, essentially redefining industries as only containing the largest stocks within the industry. These stocks are then both value-weighted and equal-weighted so that their weights sum to one. If lead-lag effects are primarily driving the 1-month, 1-month IMS2 profits, then restricting the industries to contain only the largest stocks should produce

\textsuperscript{24}Examining further the differences between inter-industry and intra-industry lead-lag relationships is beyond the scope of this paper and is left for further research.
significantly reduced profits. As Table 1.5 shows, however, the IMS2(1,1) profits for the largest stocks are of the same order of magnitude as the IMS2(1,1) profits (Table 1.4), which employ all stocks in the industry.\textsuperscript{25} This suggests that lead-lag effects are not materially affecting our profits. However, conducting the same analysis, but restricting securities to be the smallest 20% of stocks within each industry, we find that profits substantially increase (as shown in Table 1.5), indicating that a significant lead-lag relation may be present. Since smaller stocks have higher average returns, however, this discrepancy may be due to the size premium. Computing the large stock and small stock 1-month, 1-month industry momentum profits using DGTW adjusted returns, we see (in Table 1.5) that the risk-adjusted profits are much closer. Thus, once we account for risk, the large and small stock strategies exhibit similar profits, indicating that lead-lag effects are not driving our results.\textsuperscript{26}

Finally, to further assess the impact of the lead-lag effect on the IMS2(1,1) profits, we decompose the 1-month holding period return of this strategy into size quintiles as described above. Specifically, we restrict industries to only contain stocks within a particular size quintile, and compute 1-month, 1-month industry momentum profits assuming the holding period returns are only comprised of stocks within that particular size category. As before, the winning and losing industries are still chosen based on the entire industry return in the previous period, so that the same industries are selected for our IMS2 strategy. The difference here, is that we do not rebalance stocks within each quintile to sum to one, so that we can capture the contribution of each size category to total profits.

\textsuperscript{25}Asness and Stevens (1996) also document a strong 1-month industry momentum effect among 49 industries using 4 digit SIC codes, and show that this strategy remains significant even when they restrict the sample to the most liquid stocks.

\textsuperscript{26}Jegadeesh and Titman (1995b) similarly find that the short-term reversal effect in individual stock returns, and the contrarian profits that can be generated from these, are not explained by a lead-lag effect where small stocks follow the movements of large stocks.
As Table 1.5 demonstrates, the smallest stocks contribute less than 1% to total value-weighted profits, while the largest stocks comprise over 75% of these profits. Controlling for risk via the FF, Carhart, and DGTW risk adjustment procedures, the story is even more clear, as the size premium no longer confounds the influence of small stocks, and, not surprisingly, the contribution of small stocks to risk adjusted profits is even smaller. This result seems obvious since stocks within an industry are value-weighted, and reconfirms our earlier conjecture that value-weighting largely alleviates the impact of lead-lag effects on profits.

In addition, we compute the contribution of the size quintiles to total profits when stocks are weighted equally (again, here the weights do not sum to one within a size quintile, in fact, they sum to 0.20). As Table 1.5 demonstrates, the smallest quintile does not contribute substantially more to total profits than the largest quintile (22.68% vs. 14.58%), suggesting that lead-lag effects are not significantly impacting profits. Furthermore, since these are raw returns, the contribution of small stocks to total profits is confounded by the size premium. Accounting for risk via the FF, Carhart, and DGTW models, the influence of the smallest quintile is even weaker. Thus, even when stocks are equally weighted within industries, the intra-industry lead-lag effect does not seem to be substantially contributing to profits. Therefore, lead-lag effects unlikely explain our findings.

1.5.3 Industry Momentum Independent of Individual Stock Momentum

Finally, as a robustness check on our industry momentum strategies and our risk adjustment procedures, we regress the DGTW adjusted profits on the Carhart factors and subtract the random industry momentum profits from these profits
as well. This will capture the remaining effects of size, BE/ME, and individual stock momentum left over from the DGTW adjustment to ensure our results are not driven by poor risk adjustment.

More generally, since industry returns are a linear combination of individual stock returns, adjusting them via DGTW controls for momentum (and size, and BE/ME effects) in individual stock returns. Table 1.6 reports the DGTW adjusted IMS2 trading profits, which are still very profitable, suggesting that industries have a time-varying component to their returns independent of, and unexplained by, momentum (and size and BE/ME effects) in individual stock returns. This provides further evidence in favor of industry components generating positive autocorrelation in returns. Regressing the DGTW adjusted IMS2 profits on \( PR1YR \) and on the Carhart factors should capture any remaining effects of size, BE/ME, or individual stock momentum left over from the DGTW adjustment, and will also account for nearly all of the cross-sectional dispersion in industry mean returns, providing further evidence against Conrad and Kaul's (1997) hypothesis if industry momentum is still prevalent. As shown in Table 1.6, the intercepts from these regressions remain significantly positive, and the economic significance of the trading profits hardly declines.\(^{27}\) For example, for the \( L = 12, H = 9 \) strategy (DGTW adjusted profits of 0.22% per month), only 3 basis points are captured by \( PR1YR \), and only 4 basis points are captured by the Carhart model.\(^{28}\) Thus, industry momentum does not appear to be a consequence of improper return adjustment from DGTW, and size, BE/ME, and individual stock momentum fail to account for the conditional mean returns of industries, suggesting that industries may have conditional importance for pricing.

\(^{27}\)To conserve space we only report several of the profitable IMS2 strategies. Results for other strategies were similar and are available from the author upon request.

\(^{28}\)Although \( PR1YR \) does not capture industry momentum profits, the coefficients on \( PR1YR \) are highly significant, consistent with industry momentum driving individual stock momentum.
Finally, subtracting the IMS2 trading profits of the 'random' industry portfolios from the DGTW adjusted true industry profits, as a final robustness test, we find no significant reduction in the economic or statistical significance of our industry momentum trading strategies. Therefore, the true industry component of returns appears to be the key source of momentum profits, and controlling for individual stock momentum and cross-sectional variation in returns does not seem to account for these profits.

1.6 Efficiency of Industry Momentum Strategies

Our results may be consistent with both behavioral and rational theories of valuation, however. Under behavioral theories, investors persistently and irrationally underreact to industry information, thereby providing rational investors with a profit opportunity. Under rational theories, time-varying industry risk premia may exist, and the persistence in stock return performance may be driven by the stock's covariation with these economy-wide factors. Thus, apparent profits from momentum investment strategies may be compensation for risk. The question, therefore, is whether industry factors represent compensation for risk, or whether we have discovered another apparent abnormality in capital markets. At issue is whether industry effects are diversifiable, and whether they have systematic influence on asset prices. Specifically, if industries exhibit time-varying risk premia, then

\[ E[\tilde{r}_{jt}] = r_{ft} + \sum_{k=1}^{K} \beta_{jk} \lambda_{kt} + \sum_{m=1}^{M} \theta_{jm} \eta_{mt} \]  \hspace{1cm} (1.26)

where \( \lambda_{kt} \) are the risk premia on the \( K \) factors, and \( \eta_{mt} \) are the risk premia on the \( M \) industry factors. Under the behavioral hypothesis, \( \sum_{m=1}^{M} \theta_{jm} \eta_{mt} = 0, \forall t \). Since
industries do not seem to exhibit risk premia unconditionally (see Table 1.2), we assert that they may exhibit conditional premia. That is, $\eta_{mt}$ may be non-zero at certain times. We can effectively capture the influence of these conditional effects by employing an industry momentum factor in place of the sum of the industry components in (1.26).

1.6.1 Diversification of Industry Momentum Strategies

In this subsection, we wish to determine the efficiency (in a mean-variance sense) of industry momentum strategies. That is, can we increase the Sharpe ratio of our momentum strategies by including more industries in our portfolios? If industry diversification benefits are large, then strategies that diversify across industries will beat our simple industry momentum strategies on average. Conversely, if we cannot reduce the variance of our strategies, then such strategies may be inherently risky, possibly due to sources of systematic variation.

1.6.1.1 Ex Post Efficient Strategies

As an indication of the diversification benefits of investing across industries, we compare the ex post unconditional efficient frontier of our 20 industry portfolios to those of four sets of 'random' industry portfolios.\textsuperscript{29} Since random industries con-

\textsuperscript{29}Random industries are formed by replacing stocks with another security of similar past return (defined as the past 6 month return from $t - 9$ to $t - 3$, to avoid the 1-month reversal effect). Ranking stocks based on their past returns, four sets of value-weighted random industries are formed by replacing the true stock with 1) the stock ranked above it, weighted by the market capitalization of the true stock ('true' weight), 2) the stock ranked above it, weighted by the market capitalization of the replacement stock ('replacement' weight), 3) the stock ranked below it, using the 'true' weights, and 4) the stock ranked below it, using the 'replacement' weights. Since true and random industries contain the same number of securities, the relative efficiency of industry and random industry portfolios is solely determined by the benefits of diversifying across industries. Furthermore, the use of 'true' and 'replacement' weights ensures that the weighting scheme is not driving the results.
tain stocks from various industries, if industry diversification benefits are large, random industry portfolios will have lower variances than the true industries on average. Although not reported for brevity, the variances of the random industries are not significantly smaller than those of the true industries, indicating that reduction in variance by diversifying across industries is minimal. In addition, the distance between the efficient frontier formed from random industries and the random industries themselves should be much smaller than the distance between the efficient frontier of true industries and the true industries themselves, if diversification across industries is important.

Figure 1.1 plots the ex post efficient frontiers for the true and various random industries. The random industries do not appear to be more efficient than the true industries, suggesting that the diversification benefits of investing across our industries are small. Thus, we will unlikely improve the efficiency of our industry momentum strategies. In addition, we highlight the unconditional mean and variance of the IMS2 1-month, 1-month and 12-month, 12-month strategies. As the figure shows, the IMS2(1,1) strategy is very close to the ex-post efficient frontier of industries. In addition, we plot $PR1YR$, the individual stock momentum factor, which is formed based on past return sorts rather than past industry return sorts. As shown in the figure, $PR1YR$ appears highly inefficient relative to the industry momentum factors, suggesting that the industry momentum factor may be a better proxy for some true, unknown source of systematic risk, and thus should be preferred for use in performance evaluation.

1.6.1.2 Ex Ante Efficient Strategies

However, as investors, we are interested in forming efficient strategies ex ante. Therefore, using rolling conditional covariance matrix estimates, by computing
the sample covariance matrix of industry returns using the past 36 months of returns, we compute the minimum variance and tangency portfolios of the 20 industries at time $t-1$, and compare the time $t$ returns (and Sharpe measures) of these strategies to the IMS2(1,1) strategy. In addition, we compute the efficient portfolio at time $t-1$ with the same mean return as the IMS2(1,1) strategy, referred to as eff-IM and compare its subsequent performance. In other words, we are testing whether we can improve upon our simple IMS2 trading strategy by forming efficient portfolios in the prior period and evaluating their subsequent performance out of sample. Since an estimate of the mean return of industries is needed to compute the portfolio weights of the tangent and eff-IM portfolios, we employ the sample mean of the industries up to time $t-1$. Table 1.7 demonstrates that these portfolios do not significantly outperform the IMS2(1,1) strategy in the next period. Thus, the ex ante efficient strategies are quite similar in mean and risk to the simple IMS2 strategy, indicating that this simple strategy is close to efficient, and may, therefore, provide a good benchmark for performance evaluation.

1.6.2 Industry Momentum Factors

Given the efficiency of our industry momentum portfolios, the inability of the FF, Carhart, and DGTW models to explain conditional industry returns, and the fact that industry factors appear to be a more precise estimate of momentum, we propose employing industry momentum measures for performance evaluation. Two such measures are employed to better capture cross-sectional variation in expected returns: the zero-cost returns from the IMS2(1,1) strategy, and IMS2(12,12) strategy. As shown in Figure 1.1, the Sharpe ratios of these portfolios are much higher than the Sharpe ratio of $PR1YR$, suggesting that these
industry momentum measures are (locally mean-variance) more efficient than \textit{PR1YR}. Furthermore, since the industry momentum measures are not spanned by the Carhart factors (see Table 1.6), incorporating them in place of \textit{PR1YR} in the Carhart factor model may capture greater cross-sectional variation in asset returns. Indeed, this is the case as we fail to reject the GRS F-test that industry intercepts are zero relative to the FF factors plus IMS2(1,1) and the FF factors plus IMS2(12,12) (F-stats = 1.59 and 1.60, respectively). In addition, the minimum variance and tangent portfolios formed from the FF factors plus the IMS2(1,1) variable exhibit higher Sharpe ratios than the minimum variance and tangent portfolios formed from the FF factors plus \textit{PR1YR} (0.316 and 0.368 versus 0.289 and 0.319), indicating that IMS2(1,1) expands the efficient frontier beyond that for \textit{PR1YR}. Furthermore, placing stocks into 25 past return sorted portfolios, based on 6-month lagged returns, and computing the value-weighted return of each portfolio at every time \(t\), we regress these portfolios on the FF factors plus \textit{PR1YR} (i.e., the Carhart model) and on the FF factors plus IMS2(1,1). The GRS F-stat from the first set of regressions is 2.59 (p-value = 0.0000), and from the second set is 2.13 (p-value = 0.0015), indicating that the IMS2(1,1) variable better captures cross-sectional variation in returns generated by momentum. Similar results were obtained when we replace IMS2(1,1) with IMS2(12,12) in the regression, producing a GRS F-stat of 2.30 (p-value = 0.0005). Thus, although we cannot fully explain the cross-sectional differences in returns generated by these 25 portfolios, our industry momentum measures do seem to capture greater cross-sectional variation in returns than does \textit{PR1YR}, indicating that our measures may provide better benchmarks.
1.6.2.1 A Premium for Industry Momentum?

If industry momentum is related to risk, then a stock's sensitivity to an industry momentum factor should be a positive predictor of future expected returns. In other words, a premium for industry momentum should exist. Following Daniel and Titman (1997), we examine whether the cross-section of expected returns are determined by momentum characteristics or momentum factor loadings, by applying their methodology to momentum. That is, we examine whether the past return characteristic better explains the cross-section of expected returns than the covariation of stocks with a momentum factor. The results of this analysis will determine whether the momentum factor represents systematic risk which captures the dispersion in expected returns generated by momentum, and whether a premium exists for this factor. More generally, the results will provide evidence either further supporting characteristic-based models, or favoring factor models of asset returns.

Using the prior 36 months of returns, we estimate factor loadings on the industry momentum factor by regressing the returns of each stock over the past 36 months on the FF factors plus the IMS2(12,12) profits. Stocks with less than 24 months of prior return history are excluded, and the remaining stocks are sorted into quintiles based on their estimated coefficients on IMS2(12,12). Within each of these IMS2(12,12) coefficient quintiles, stocks are then sorted into quintiles based on their prior 12 month returns, and are weighted by the inverse of their residual variance from the pre-formation regression, in order to minimize the influence of stocks with largely unexplained returns, and to reduce the idiosyncratic variance of the portfolios.\textsuperscript{30} The same analysis is repeated twice by substituting

\textsuperscript{30}Since we are testing whether industry momentum sensitivity is related to expected returns, and whether it captures the cross-sectional dispersion in returns generated by momentum, following Berk's (1997) critique of the Daniel and Titman (1997) methodology, we first sort
the IMS2(1,1) variable and Carhart's PR1YR variable for IMS2(12,12) in the prior 36-month (or pre-formation) regression.

Table 1.8 reports the average returns of the factor loading and past return characteristic sorted portfolios. As the table shows, prior return sorted portfolios exhibit statistically insignificant variation within factor loading groups, although the economic significance of the past return characteristic sorted portfolios remains high. Thus, controlling for industry momentum factor sensitivity does not account for the relation between past returns and expected returns. In other words, consistent with Daniel and Titman (1997), a characteristic model appears to better capture the cross-sectional variation in common stock returns generated by momentum. However, stocks with high loadings on the industry momentum factor exhibit significantly greater average returns than stocks with low loadings (as evidenced by the last row in the table), indicating that covariation with the industry momentum factor is positively related to expected returns. Therefore, a premium appears to exist for the industry momentum factor, consistent with a risk-based explanation for momentum. Thus, at least a portion of the momentum effect appears to be related to risk.\textsuperscript{31} However, no such premium is apparent for PR1YR, suggesting that the Carhart factor is a poor proxy for the true momentum factor in the economy, consistent with our earlier claims that industry momentum measures may be better benchmarks for performance evaluation than 'naive' past return proxies.

\textsuperscript{31}Moskowitz (1997) shows that momentum is partly due to systematic risk, but that the magnitude of momentum profits appear too high to be explained by risk alone. This is consistent with the findings here.
1.7 Conclusion

The fact that industry effects seem to be the primary cause of momentum in stock returns may have significant consequences for performance evaluation. For instance, the persistence in mutual fund performance documented by Grinblatt and Titman (1991), Hendricks, Patel, and Zeckhauser (1993), Grinblatt, Titman, and Wermers (1995), Wermers (1997), Carhart (1997), and many others may be much less prevalent once industry effects are accounted for. The “hot hands” often attributed to fund managers may, in fact, largely be due to institutional investors shifting their holdings away from industries/sectors that experienced recent underperformance and towards sectors with high past returns, a potentially risky strategy.

In addition, new insights on the role of industries in the international economy may be provided by examining the conditional influence of industries in an international setting. Recently, Rouwenhorst (1997) documents that momentum strategies are highly profitable in Europe, suggesting that a common global factor may be driving this phenomenon. One candidate for such factors are industries. Thus, extending the results in this paper to the international economy is a natural next step. In addition, analyzing the conditional influence of industries may help resolve the debate on the role of industrial structure for international diversification strategies, argued by Roll (1992), Heston and Rouwenhorst (1994), Griffin and Karolyi (1995), and most recently Arshanapalli, Doukas, and Lang (1997). Evaluating the diversification benefits within versus across industries (both domestically and internationally) may also be an interesting topic for future research.

This paper also highlights the importance of accounting for industry effects when assessing abnormal performance, measuring the cost of equity capital (see
Fama and French (1997)), or evaluating the financial impact of an event. To that end, the impact of industries on contrarian investment strategies and on the lead-lag relation between large and small firms may provide interesting avenues for further research.

Finally, the potential existence of conditional industry risk premia may question the usefulness of current performance evaluation techniques which identify factors presumed to price assets over the entire sample period. This paper argues that failure to account for conditional industry influences implies abnormal profits can be obtained from momentum investment strategies. However, a determination of whether these conditional influences represent risk premia is, at this point, premature. Moskowitz (1997) analyzes these issues in depth and finds evidence that conditional and time-varying industry risk premia appear to exist, indicating that momentum profits are partly compensation for risk. However, only a fraction of momentum profits can truly be attributable to risk, as the premium generated from momentum strategies appears too high to be explained by risk alone.
1.8 Appendix A: Return Adjustment Procedures

1.8.1 Risk-Based Models

The Fama and French (1993) and Carhart (1997) models specify factor-mimicking portfolios that are correlated with unknown economy-wide factors (as in the APT of Ross (1976)) or that represent the changing investment opportunity set (as in Merton (1973)). Under the three factor FF model, regressing the time-series of excess returns on the three factor-mimicking portfolios should produce an intercept not significantly different from zero if the model explains returns:

\[ R_{t,t} - r_{f,t} = \alpha_t + \beta_{1,t}[MKT_t - r_{f,t}] + \beta_{2,t}SMB_t + \beta_{3,t}HML_t + e_{t,t}. \quad (1.27) \]

where \( R_{f,t} \) is the return on industry \( I \) at time \( t \) (\( \forall I = 1, \ldots, 20 \)), \( MKT \) is the return on the value-weighted market portfolio, \( r_f \) is the three-month Treasury Bill rate, \( SMB \) is a zero-cost portfolio that is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, and \( HML \) is the difference between the returns of a portfolio of high book-to-market equity (BE/ME) stocks and a portfolio of low BE/ME stocks.

Abnormal returns relative to the Carhart model are defined similarly, with the addition of another factor-mimicking portfolio, \( PR1YR \), which is the return on a portfolio of high prior year return stocks minus a portfolio of low prior year return stocks. Appendix 2.7 describes the construction of the factor-mimicking portfolios.
1.8.1.1 Factor Portfolio Construction

$SMB$ and $HML$ are constructed in the same manner as Fama and French (1993). After ranking all NYSE firms by their BE/ME\textsuperscript{32} ratios at the end of year $t - 1$ and their market capitalization at the end of June of year $t$, we form 30 percent and 70 percent breakpoints for BE/ME and a 50 percent breakpoint for market capitalization. Beginning in July of year $t$,\textsuperscript{33} all NYSE, AMEX, and NASDAQ stocks are placed into the three book-to-market groups (High, Medium, and Low) and two size groups (Small and Big) based on the breakpoints. Six value-weighted portfolios are then formed based on the intersection between the size and book-to-market groups. The $HML$ returns are defined as: $(r_{HB} + r_{HS} - r_{LB} - r_{LS})/2$, and the $SMB$ returns are defined as: $(r_{HS} + r_{MS} + r_{LS} - r_{HB} - r_{MB} - r_{LB})/3$. Only firms that have been listed on COMPUSTAT for at least two years prior to portfolio formation and have prices on CRSP in December of year $t - 1$ and June of year $t$ are included.

$PR1YR$ is constructed in the same manner as Carhart (1997) by forming an equally-weighted portfolio of firms with the highest 30 percent prior 12-month returns, lagged one month (i.e., from $t - 12$ to $t - 1$), and subtracting from this an equally weighted portfolio of firms with the lowest 30 percent prior 12-month returns. The resulting zero-cost portfolio is re-formed monthly.

\textsuperscript{32}Book value of equity is defined as stockholder's equity plus deferred taxes and investment tax credits. These numbers are obtained from COMPUSTAT. BE/ME is calculated by dividing the most recent book value in year $t - 1$ by the number of shares outstanding times price at the end of year $t - 1$, obtained from CRSP. All firms with BE/ME values less than zero are excluded.

\textsuperscript{33}As in Fama and French (1993), the end of June is used as the portfolio formation date to ensure that book values of equity are publicly available from annual reports at the time portfolios are formed.
1.8.2 Characteristic-Based Models

Since Daniel and Titman (1997) find that characteristics of stocks, rather than sensitivities of stock returns to factor-mimicking portfolios, better explain the cross-sectional variation in returns, we employ the characteristic-based return adjustment of Daniel, Grinblatt, Titman and Wermers (1997) (DGTW). DGTW argue that since investors seem to require a premium for stocks with certain characteristics, securities with similar characteristics should have the same price. Therefore, they match a stock with a diversified portfolio of securities with similar size, BE/ME, and momentum characteristics, and subtract the return of this benchmark portfolio from the stock’s return. The abnormal return of industry $I$ is the value-weighted sum of the individual DGTW-adjusted returns of stocks belonging to industry $I$ at time $t$. The DGTW return adjustment procedure is briefly described in Appendix 1.8.2.1.

1.8.2.1 DGTW Return Adjustment Procedure

DGTW assign stocks to one of five categories based on the prior period’s market capitalization, then within each of these groups divide stocks into five BE/ME categories, and then into five 12 month prior return groups. The breakpoints used for each of the three characteristics are based on NYSE stocks only. Value-weighted returns are then computed for each group of stocks at time $t$, creating 125 portfolio returns. Each stock is then matched with one of the 125 portfolios based on its characteristics at time $t - 1$. The abnormal return for stock $j$ is defined as the return on the stock minus the return on the matched portfolio at time $t$ (i.e., $\tilde{r}_{jt} - \tilde{R}_{i,t-1}^b$, where the latter is the month $t$ return of the matched
characteristic-based portfolio for stock $j$).\textsuperscript{34}  

\textsuperscript{34}These returns were graciously supplied by Kent Daniel, Mark Grinblatt, Sheridan Titman, and Russ Wermers, and correspond to the Jan., 1973 to July, 1995 time period. For further description of the methodology and construction of their characteristic-based performance measure, see Daniel, Grinblatt, Titman, and Wermers (1997).
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining</td>
<td>10-14</td>
<td>279.86 (123)</td>
<td>0.165</td>
<td>4.50%</td>
<td>258.89</td>
</tr>
<tr>
<td>Food and Beverage</td>
<td>20</td>
<td>126.89 (100)</td>
<td>0.159</td>
<td>4.53%</td>
<td>725.13</td>
</tr>
<tr>
<td>Textile and Apparel</td>
<td>22-3</td>
<td>117.03 (84)</td>
<td>0.152</td>
<td>6.73%</td>
<td>94.47</td>
</tr>
<tr>
<td>Paper Products</td>
<td>26</td>
<td>53.21 (41)</td>
<td>0.177</td>
<td>2.32%</td>
<td>735.52</td>
</tr>
<tr>
<td>Chemical</td>
<td>28</td>
<td>208.15 (126)</td>
<td>0.172</td>
<td>9.85%</td>
<td>730.91</td>
</tr>
<tr>
<td>Petroleum</td>
<td>29</td>
<td>37.49 (26)</td>
<td>0.169</td>
<td>8.94%</td>
<td>350.22</td>
</tr>
<tr>
<td>Construction</td>
<td>32</td>
<td>56.44 (39)</td>
<td>0.140</td>
<td>0.96%</td>
<td>260.22</td>
</tr>
<tr>
<td>Primary Metals</td>
<td>33</td>
<td>85.11 (66)</td>
<td>0.130</td>
<td>1.95%</td>
<td>308.10</td>
</tr>
<tr>
<td>Fabricated Metals</td>
<td>34</td>
<td>119.38 (76)</td>
<td>0.163</td>
<td>1.21%</td>
<td>181.95</td>
</tr>
<tr>
<td>Machinery</td>
<td>35</td>
<td>274.16 (143)</td>
<td>0.152</td>
<td>7.38%</td>
<td>388.12</td>
</tr>
<tr>
<td>Electrical Equipment</td>
<td>36</td>
<td>311.60 (165)</td>
<td>0.180</td>
<td>5.57%</td>
<td>283.81</td>
</tr>
<tr>
<td>Transport Equipment</td>
<td>37</td>
<td>105.35 (91)</td>
<td>0.172</td>
<td>5.01%</td>
<td>733.32</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>38-9</td>
<td>235.21 (70)</td>
<td>0.162</td>
<td>4.08%</td>
<td>264.60</td>
</tr>
<tr>
<td>Railroads</td>
<td>40</td>
<td>20.11 (9)</td>
<td>0.195</td>
<td>0.81%</td>
<td>658.42</td>
</tr>
<tr>
<td>Other Transportation</td>
<td>41-7</td>
<td>88.03 (51)</td>
<td>0.172</td>
<td>1.19%</td>
<td>206.83</td>
</tr>
<tr>
<td>Utilities</td>
<td>49</td>
<td>187.08 (114)</td>
<td>0.119</td>
<td>7.52%</td>
<td>642.07</td>
</tr>
<tr>
<td>Department Stores</td>
<td>53</td>
<td>54.79 (36)</td>
<td>0.161</td>
<td>2.89%</td>
<td>1,006.67</td>
</tr>
<tr>
<td>Other Retail</td>
<td>50-2, 54-9</td>
<td>377.06 (143)</td>
<td>0.172</td>
<td>3.34%</td>
<td>156.78</td>
</tr>
<tr>
<td>Finance. Real Estate</td>
<td>60-9</td>
<td>891.56 (152)</td>
<td>0.153</td>
<td>12.42%</td>
<td>244.40</td>
</tr>
<tr>
<td>Other</td>
<td>Other</td>
<td>981.18 (221)</td>
<td>0.181</td>
<td>14.90%</td>
<td>277.21</td>
</tr>
</tbody>
</table>

Average: 250.48 (93.80) 0.162 5.00% 546.34

<sup>1</sup> Minimum number of stocks in parentheses.
Table 1.2: Average Monthly Raw Excess and Abnormal Returns of Industries

Average monthly raw excess and abnormal returns of 20 industry portfolios over the July 1963 - July 1995 time period. Raw excess returns ($R_t - r_f$) are returns in excess of the 3 month Treasury Bill rate. Abnormal returns are defined relative to the FF three factor model, Carhart four factor model, and DGTW characteristic-based model. The DGTW abnormal returns correspond to the Jan. 1973 - July, 1995 time period. Standard deviations ($\sigma$) of the raw return portfolios are also provided, along with t-statistics that the abnormal returns are significantly different from zero. In addition, the Gibbons, Ross, and Shanken (1989) F-statistic, under both the normal ($W$) and elliptical ($W_e$) distributions, that abnormal returns are jointly zero, along with their p-values in parentheses, are reported at the bottom of the table, along with the average adjusted $R^2$ from the regressions and an F-statistic that industry returns are equal to each other (the first column contains the GRS F-statistic and $R^2$ under the CAPM, using a market model regression).

<table>
<thead>
<tr>
<th>Industry</th>
<th>Excess Returns $R_t - r_f$</th>
<th>$\sigma$</th>
<th>3-Factor Model FF(93)</th>
<th>4-Factor Model Carhart(97)</th>
<th>Characteristic Model DGTW(97)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\alpha}$ t-stat</td>
<td>$\hat{\beta}$ t-stat</td>
<td>$\hat{\gamma}$ t-stat</td>
<td>$\hat{\delta}$ t-stat</td>
<td>$\hat{\Delta}$ t-stat</td>
</tr>
<tr>
<td>Mining</td>
<td>0.0040 0.0569</td>
<td>0.0023  0.87</td>
<td>0.0016 0.59</td>
<td>-0.0001 0.03</td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>0.0065 0.0445</td>
<td>0.0012  1.39</td>
<td>0.0011 1.20</td>
<td>0.0029 2.29*</td>
<td></td>
</tr>
<tr>
<td>Apparel</td>
<td>0.0046 0.0619</td>
<td>-0.0040 -2.99**</td>
<td>-0.0039 -2.89**</td>
<td>-0.0007 -0.45</td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>0.0047 0.0532</td>
<td>0.0022  1.78</td>
<td>0.0022 1.75</td>
<td>0.0009 0.60</td>
<td></td>
</tr>
<tr>
<td>Chemical</td>
<td>0.0047 0.0468</td>
<td>0.0025  2.50**</td>
<td>0.0024 2.37*</td>
<td>0.0008 0.88</td>
<td></td>
</tr>
<tr>
<td>Petroleum</td>
<td>0.0055 0.0516</td>
<td>0.0025  1.23</td>
<td>0.0022 1.06</td>
<td>0.0015 1.10</td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>0.0047 0.0561</td>
<td>-0.0015 -1.27</td>
<td>-0.0019 -1.66</td>
<td>-0.0005 -0.32</td>
<td></td>
</tr>
<tr>
<td>Prim. Metals</td>
<td>0.0020 0.0618</td>
<td>-0.0013 -0.82</td>
<td>-0.0015 -0.90</td>
<td>-0.0018 -0.83</td>
<td></td>
</tr>
<tr>
<td>Fab. Metals</td>
<td>0.0054 0.0523</td>
<td>0.0001  0.11</td>
<td>-0.0002 -0.23</td>
<td>0.0016 1.62</td>
<td></td>
</tr>
<tr>
<td>Machinery</td>
<td>0.0030 0.0541</td>
<td>-0.0008 -0.77</td>
<td>-0.0008 -0.75</td>
<td>-0.0016 -1.39</td>
<td></td>
</tr>
<tr>
<td>Electrical Eq.</td>
<td>0.0049 0.0670</td>
<td>0.0003  0.24</td>
<td>-0.0001 -0.01</td>
<td>0.0017 1.39</td>
<td></td>
</tr>
<tr>
<td>Trans. Eq.</td>
<td>0.0043 0.0543</td>
<td>-0.0008 -0.78</td>
<td>-0.0010 -0.92</td>
<td>0.0011 0.71</td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.0055 0.0546</td>
<td>0.0007  0.72</td>
<td>0.0004 0.42</td>
<td>-0.0017 -1.41</td>
<td></td>
</tr>
<tr>
<td>Railroads</td>
<td>0.0055 0.0572</td>
<td>0.0015  0.80</td>
<td>0.0008 0.40</td>
<td>0.0025 0.96</td>
<td></td>
</tr>
<tr>
<td>Other Trans.</td>
<td>0.0040 0.0675</td>
<td>-0.0017 -1.07</td>
<td>-0.0016 -0.99</td>
<td>-0.0017 -0.78</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>0.0027 0.0389</td>
<td>-0.0003 -0.24</td>
<td>-0.0001 -0.12</td>
<td>-0.0011 -0.75</td>
<td></td>
</tr>
<tr>
<td>Dept. Stores</td>
<td>0.0051 0.0592</td>
<td>-0.0023 -1.38</td>
<td>-0.0025 -1.48</td>
<td>0.0006 0.31</td>
<td></td>
</tr>
<tr>
<td>Retail</td>
<td>0.0055 0.0562</td>
<td>-0.0007 -0.74</td>
<td>-0.0008 -0.89</td>
<td>0.0013 1.17</td>
<td></td>
</tr>
<tr>
<td>Finance</td>
<td>0.0045 0.0462</td>
<td>-0.0013 -1.38</td>
<td>-0.0015 -1.51</td>
<td>-0.0001 -0.09</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.0046 0.0454</td>
<td>0.0008  0.99</td>
<td>0.0007 0.80</td>
<td>0.0007 1.18</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.0043 0.0538</td>
<td>-0.00003 -0.04</td>
<td>0.0002 (-0.19)</td>
<td>0.0003 (0.31)</td>
<td></td>
</tr>
<tr>
<td>Avg. $R^2$</td>
<td>0.743</td>
<td>0.831</td>
<td>0.831</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>GRS F-stat: $W$</td>
<td>2.92 (0.000)</td>
<td>2.01 (0.007)</td>
<td>1.82 (0.018)</td>
<td>1.68 (0.035)</td>
<td></td>
</tr>
<tr>
<td>GRS F-stat: $W_e$</td>
<td>2.90 (0.000)</td>
<td>1.99 (0.007)</td>
<td>1.78 (0.021)</td>
<td>1.64 (0.042)</td>
<td></td>
</tr>
<tr>
<td>F-stat: $R^2$</td>
<td>0.829 (0.677)</td>
<td>1.94 (0.011)</td>
<td>1.94 (0.011)</td>
<td>1.43 (0.111)</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the 5% level. ** Significant at the 1% level. The Bonferroni-adjusted critical values at the 5% and 1% level are 3.10 and 3.34, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 1.3: Momentum Profits for Individual Equities, Industries, and Random Industries

Panel A reports average monthly returns of winners minus losers. Wi-Lo. (highest 30% minus lowest 30%) momentum portfolios of individual equities over July, 1963 - July, 1995 (i.e., T=383). Portfolios are formed based on $L = 6$-month lagged returns and held for $H = 6$ months. The $L$-month lagged returns are always raw returns, to be used for portfolio formation. Results are reported for holding period raw, DGTW, size and BE/ME adjusted ($r^b_{jt}$), size and BE/ME adjusted minus industry ($r^{b,\text{ind}}_{jt}$), size and BE/ME adjusted minus pure industry ($r^{b,\text{PI}}_{jt}$), and size and BE/ME adjusted minus random $^a$ industry returns ($r^{b,\text{R}}_{jt}$). Panel B reports average monthly profits of momentum strategies of industries and random industries. Raw and DGTW adjusted profits are reported for the industry momentum strategies. Panel C reports the raw profits of the industry neutral, excess industry, and high industry losers minus low industry winners portfolios described in Section 1.4.1.8.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r^b_{jt}$</td>
<td>$r^{b,\text{ind}}_{jt}$</td>
<td>$r^{b,\text{PI}}_{jt}$</td>
</tr>
<tr>
<td>(6,6)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0043</td>
<td>(4.65**)</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Size, BE/ME, and Industry</td>
<td>Pure Industry</td>
<td>Random $^a$ Industry</td>
<td></td>
</tr>
<tr>
<td>$r^{b,\text{ind}}_{jt}$</td>
<td>$r^{b,\text{PI}}_{jt}$</td>
<td>$r^{b,\text{R}}_{jt}$</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0008</td>
<td>(0.91)</td>
<td>-0.0003</td>
</tr>
<tr>
<td>(6,6)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Raw Industry</td>
<td>DGTW Industry</td>
<td>Raw Random $^a$</td>
<td></td>
</tr>
<tr>
<td>$R^*_f$</td>
<td>$R^*_{1f}$</td>
<td>Industry. $R^*_{1f}$</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0043</td>
<td>(4.24**)</td>
<td>0.0020</td>
</tr>
<tr>
<td>(6,6)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Industry</td>
<td>Excess</td>
<td>High Ind. Losers</td>
<td>Neutral</td>
</tr>
<tr>
<td>$r^b_{jt}$</td>
<td>$r^{b,\text{PI}}_{jt}$</td>
<td>$r^b_{jt}$</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0003</td>
<td>(0.20)</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

$^a$ Random industries are generated by replacing each stock return with an equal-weighted average of the stocks ranked above and below it based on their past 6-month returns. $^*$Significant at the 5% level. $^{**}$Significant at the 1% level.
Table 1.4: Industry Momentum Trading Profits

Average monthly returns of the three industry momentum trading strategies over the July, 1963 - July, 1995 time period (i.e., \(T=583\)). The industry momentum portfolios are formed based on \(L\)-month lagged returns and held for \(H\) months. Three industry momentum strategies are employed to identify the winning and losing industries: IMS1) Winners = return of highest momentum industry; IMS2) Winners = equal-weighted return of highest 3 momentum industries. Losers = equal-weighted return of lowest 3 momentum industries; IMS3) Winners = equal-weighted return of all industries above the average prior return. Losers = equal-weighted return of all industries below the average prior return. Panels A, B, and C contain the results for IMS1, IMS2, and IMS3, respectively, where the returns for the winners (Wi), losers (Lo), and winners minus losers (Wi-Lo) are reported, as well as t-statistics for the latter in parentheses. For brevity, we report only the \(L = 1, 6, 12, 24\)-month lagged and \(H = 1, 6, 12, 24, 36\) month holding period strategies.

<table>
<thead>
<tr>
<th>(L)</th>
<th>(H)</th>
<th>Panel A: IMS1</th>
<th>Panel B: IMS2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Wi</td>
<td>Lo</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0.0198</td>
<td>0.0144</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0064</td>
<td>0.0103</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>0.0135</td>
<td>0.0041</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>(4.47)</td>
<td>(2.86)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.0170</td>
<td>0.0165</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0132</td>
<td>0.0120</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0038</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>(1.22)</td>
<td>(2.97)</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0.0193</td>
<td>0.0160</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0111</td>
<td>0.0104</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0082</td>
<td>0.0056</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>(2.60)</td>
<td>(3.98)</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0.0174</td>
<td>0.0147</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0133</td>
<td>0.0114</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0041</td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>(1.26)</td>
<td>(2.20)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(L)</th>
<th>(H)</th>
<th>Panel C: IMS3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Wi</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0.0058</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>(2.61)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>(3.77)</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>(2.31)</td>
</tr>
</tbody>
</table>

The Bonferroni-adjusted critical values at the 5% and 1% significance levels are 3.24 and 3.67, respectively.
Table 1.5: Impact of Lead-Lag Effects on Industry Momentum Profits

Panel A reports the 1-month, 1-month momentum strategy raw profits on individual securities, which buys the highest 30% of past 1-month return stocks and sells the lowest 30% of past 1-month return stocks. The 30% and 70% breakpoints are set within each industry, and within each ‘random’ industry, as described in Section 1.5.2.2. Panel B reports the decomposition of the industry 1-month, 1-month momentum strategy profits (IMS2(1,1)) into various components related to size.

<table>
<thead>
<tr>
<th>Weight:</th>
<th>True Industries</th>
<th>Random Industries (Past Return)</th>
<th>Random Industries (Size)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (T-stat)</td>
<td>Equal Value</td>
<td>Equal Value</td>
<td>Equal Value</td>
</tr>
<tr>
<td>-0.0148 (-12.14)</td>
<td>-0.0088</td>
<td>-0.0116</td>
<td>-0.0044</td>
</tr>
<tr>
<td>0.0108 (8.04)</td>
<td>-0.0046</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Industry Momentum 1-Month, 1-Month Strategy

<table>
<thead>
<tr>
<th>Weight:</th>
<th>IMS2(1,1) for Largest Stocks</th>
<th>IMS2(1,1) for Smallest Stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw DGTW</td>
<td>Equal Value</td>
<td>Equal Value</td>
</tr>
<tr>
<td>0.0110 (6.54)</td>
<td>0.01099 (5.39)</td>
<td>0.0178 (9.70)</td>
</tr>
<tr>
<td>0.0053 (3.79)</td>
<td>0.0043 (3.54)</td>
<td>0.0070 (5.64)</td>
</tr>
</tbody>
</table>

Decomposition of IMS2(1,1) Profits: Value-Weighted Industries

<table>
<thead>
<tr>
<th>Size Quintiles</th>
<th>Profits %</th>
<th>( \delta ) (FF) %</th>
<th>( \delta ) (Carhart) %</th>
<th>( \delta ) (DGTW) %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00098</td>
<td>0.92%</td>
<td>0.00092</td>
<td>0.83%</td>
</tr>
<tr>
<td>2</td>
<td>0.000251</td>
<td>2.36%</td>
<td>0.000258</td>
<td>2.23%</td>
</tr>
<tr>
<td>3</td>
<td>0.000646</td>
<td>6.07%</td>
<td>0.000650</td>
<td>5.84%</td>
</tr>
<tr>
<td>4</td>
<td>0.001629</td>
<td>15.30%</td>
<td>0.001592</td>
<td>14.31%</td>
</tr>
<tr>
<td>5</td>
<td>0.008024</td>
<td>75.35%</td>
<td>0.008543</td>
<td>76.69%</td>
</tr>
<tr>
<td>Total</td>
<td>0.01051</td>
<td>0.011134</td>
<td>0.01060</td>
<td>0.004256</td>
</tr>
</tbody>
</table>

Decomposition of IMS2(1,1) Profits: Equal-Weighted Industries

<table>
<thead>
<tr>
<th>Size Quintiles</th>
<th>Profits %</th>
<th>( \delta ) (FF) %</th>
<th>( \delta ) (Carhart) %</th>
<th>( \delta ) (DGTW) %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.003488</td>
<td>22.68%</td>
<td>0.003499</td>
<td>22.36%</td>
</tr>
<tr>
<td>2</td>
<td>0.003285</td>
<td>21.36%</td>
<td>0.003289</td>
<td>21.03%</td>
</tr>
<tr>
<td>3</td>
<td>0.003360</td>
<td>21.85%</td>
<td>0.003384</td>
<td>21.63%</td>
</tr>
<tr>
<td>4</td>
<td>0.003003</td>
<td>19.53%</td>
<td>0.003070</td>
<td>19.62%</td>
</tr>
<tr>
<td>5</td>
<td>0.002242</td>
<td>14.58%</td>
<td>0.002206</td>
<td>15.37%</td>
</tr>
<tr>
<td>Total</td>
<td>0.011377</td>
<td>0.015648</td>
<td>0.015625</td>
<td>0.006756</td>
</tr>
</tbody>
</table>

* 'Past return random' industries are generated by replacing each stock return with the stock ranked above it based on the prior month return.

* 'Size random' industries are generated by replacing each stock return with the stock ranked above it based on the prior month market capitalization.
Table 1.6: Industry Momentum Independent of Individual Stock Momentum

Panel A reports average monthly DGTW-adjusted returns of the IMS2(L,H) industry momentum trading strategy over the Jan., 1973 - July, 1995 time period (i.e., T=270). Industry momentum portfolios are formed based on L-month lagged raw returns and held for H months, where holding period returns are adjusted via DGTW. T-statistics for the Wi-Lo zero-cost portfolios are reported in parentheses. In addition, the winners minus losers returns from the "random industries" (Wi-Lo)**, are also reported, as well as the difference between Wi-Lo and (Wi-Lo)**. Panel B reports results from the regression of the IMS2 DGTW-adjusted profits on the momentum factor PRIYR and on the Carhart four factor model, as a robustness check on the DGTW return adjustment.

<table>
<thead>
<tr>
<th>Strategy(L,H)</th>
<th>Panel A: IMS2 DGTW-Adjusted Profits</th>
<th>Panel B: Regressed on PRIYR and Carhart Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1.1) (6.6) (6.9) (6.12) (12.6) (12.9) (12.12)</td>
<td>α (0.0013) 0.0029 0.0022 0.0023 0.0019 0.0014</td>
</tr>
<tr>
<td>Wi</td>
<td>0.0068 0.0024 0.0026 0.0022 0.0028 0.0022 0.0017</td>
<td>(3.89**) (1.47) (4.30**) (3.81**) (2.98**) (2.97**) (2.38*)</td>
</tr>
<tr>
<td>Lo</td>
<td>0.0003 0.0004 -0.0007 -0.0003 -0.0002 -0.0002 -0.0000</td>
<td>(0.04) (-1.09) (-0.96) (-0.98) (-1.92) (-1.79) (-1.80)</td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0065 0.0020 0.0032 0.0024 0.0030 0.0023 0.0018</td>
<td>(2.96**) (5.01**) (4.31**) (4.17**) (4.14**) (3.73**)</td>
</tr>
<tr>
<td>(t-stat)</td>
<td>(3.99**) (2.27**) (4.90**) (4.33**) (3.89**) (3.70**) (3.11**)</td>
<td></td>
</tr>
<tr>
<td>(Wi-Lo)**</td>
<td>0.0000 -0.0005 -0.0004 -0.0003 -0.0009 -0.0007 -0.0006</td>
<td></td>
</tr>
<tr>
<td>(t-stat)</td>
<td>(0.04) (-1.09) (-0.96) (-0.98) (-1.92) (-1.79) (-1.80)</td>
<td></td>
</tr>
<tr>
<td>(Wi-Lo)-(Wi-Lo)**</td>
<td>0.0064 0.0025 0.0036 0.0027 0.0039 0.0031 0.0024</td>
<td></td>
</tr>
<tr>
<td>(t-stat)</td>
<td>(2.96**) (2.64**) (5.01**) (4.31**) (4.17**) (4.14**) (3.73**)</td>
<td></td>
</tr>
</tbody>
</table>

Panel B:

- \( \bar{\alpha} \)
- \( \beta_{PRIYR} \)
- \( \beta_{Mkt-rf} \)
- \( \beta_{SMB} \)
- \( \beta_{HML} \)
- \( \beta_{PRIYR} \)

\( \bar{\alpha} \) Random industries are generated by replacing each stock return with an equal-weighted average of the stocks ranked above and below it in momentum.

\*Significant at the 5% level. **Significant at the 1% level.
Table 1.7: Diversification of Industry Momentum Strategies

The table reports the out of sample ex ante minimum variance (MVP) and tangency (T) portfolios of the 20 industries using rolling 36-month (from t - 36 to t - 1) sample covariance matrix estimates and returns from the prior period. The weights of the MVP and T portfolios are formed in t - 1 and applied to the returns at time t. eff-IM is the efficient portfolio with the same return as IMS2(1.1). Mean industry returns are computed using sample mean returns from date 0 to t - 1 for the tangent and eff-IM weights. T-stats are reported in parentheses.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Mean</th>
<th>t-stat</th>
<th>Sharpe</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMS2(1,1)</td>
<td>0.0105</td>
<td>(5.63**)</td>
<td>0.305</td>
<td>(5.98**)</td>
</tr>
<tr>
<td>MVP</td>
<td>0.0129</td>
<td>(4.24**)</td>
<td>0.338</td>
<td>(2.23**)</td>
</tr>
<tr>
<td>MVP-IMS2(1,1)</td>
<td>0.0024</td>
<td>(0.87)</td>
<td>0.033</td>
<td>(0.22)</td>
</tr>
<tr>
<td>T</td>
<td>0.0194</td>
<td>(4.84**)</td>
<td>0.527</td>
<td>(3.71**)</td>
</tr>
<tr>
<td>T-IMS2(1,1)</td>
<td>0.0089</td>
<td>(2.11*)</td>
<td>0.222</td>
<td>(1.57)</td>
</tr>
<tr>
<td>eff-IM</td>
<td>0.0071</td>
<td>(1.17)</td>
<td>0.302</td>
<td>(2.08*)</td>
</tr>
<tr>
<td>eff-IM - IMS2(1,1)</td>
<td>-0.0035</td>
<td>(-0.42)</td>
<td>-0.003</td>
<td>(-0.02)</td>
</tr>
</tbody>
</table>

*Significant at the 5% level. **Significant at the 1% level.
Figure 1.1: Ex-Post Efficient Frontiers of Industry and 'Random' Industry Portfolios

We compare the ex post unconditional efficient frontier of our 20 industry portfolios to those of four sets of 'random' industry portfolios. Random industries are formed by replacing stocks with another security of similar past return (defined as the past 6 month return from t−9 to t−3, to avoid the 1-month reversal effect). Ranking stocks based on their past returns, four sets of value-weighted random industries are formed by replacing the true stock with: 1) the stock ranked above it, weighted by the market capitalization of the true stock ('true' weight); 2) the stock ranked above it, weighted by the market capitalization of the replacement stock ('replacement' weight); 3) the stock ranked below it, using the 'true' weights; and 4) the stock ranked below it, using the 'replacement' weights.
Table 1.8: Average Monthly Returns of Industry Momentum Factor Loading and Past Return Sorted Portfolios

Average monthly returns (t-stats in parentheses) for the 25 industry momentum loading and past return characteristic portfolios are reported below over the period July, 1966 to July, 1995. Estimated factor loadings are computed by regressing stock returns on $R_m - r_f, S M B, H M L,$ and IMS2(12.12) over the prior 36 months. Stocks are placed into quintiles based on their IMS2(12.12) loading and then within each of these groups placed into quintiles based on their prior 12 month returns. Stocks are weighted within each of the groups by the inverse of their residual variation from the pre-formation regression. The Wi-Lo row represents high prior 12 month return stocks minus low prior 12 month return stocks within each of the factor loading groups (i.e., rows 5+4 minus rows 1+2). The $\beta^{high} - \beta^{low}$ portfolio represents the high factor loading portfolios minus the low factor loading portfolios, which is the equal-weighted average of the last two columns minus the first two columns. Results for loadings based on IMS2(1.1) and PRIYR are reported as well.

<table>
<thead>
<tr>
<th>Past Return Portfolio</th>
<th>Industry Momentum Factor Portfolio: IMS2(12.12)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0089 (2.90)</td>
<td>0.0080 (2.11)</td>
<td>0.0090 (2.52)</td>
<td>0.0093 (2.56)</td>
<td>0.0143 (2.86)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0074 (2.11)</td>
<td>0.0100 (3.53)</td>
<td>0.0102 (3.84)</td>
<td>0.0108 (3.80)</td>
<td>0.0092 (2.55)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0074 (2.26)</td>
<td>0.0106 (4.28)</td>
<td>0.0107 (4.70)</td>
<td>0.0112 (4.45)</td>
<td>0.0124 (3.79)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0101 (3.24)</td>
<td>0.0117 (4.81)</td>
<td>0.0120 (5.30)</td>
<td>0.0124 (4.95)</td>
<td>0.0142 (4.46)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0131 (3.78)</td>
<td>0.0139 (4.85)</td>
<td>0.0145 (5.31)</td>
<td>0.0142 (4.95)</td>
<td>0.0158 (4.47)</td>
<td></td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0034 (1.68)</td>
<td>0.0038 (1.99)</td>
<td>0.0037 (2.05)</td>
<td>0.0032 (1.69)</td>
<td>0.0033 (1.44)</td>
<td></td>
</tr>
</tbody>
</table>

$\beta^{high} - \beta^{low} = 0.0023 (2.36)$

<table>
<thead>
<tr>
<th>Past Return Portfolio</th>
<th>Industry Momentum Factor Portfolio: IMS2(1.1)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0093 (2.12)</td>
<td>0.0085 (2.29)</td>
<td>0.0092 (2.56)</td>
<td>0.0102 (2.59)</td>
<td>0.0142 (2.85)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0077 (2.18)</td>
<td>0.0089 (3.17)</td>
<td>0.0099 (3.68)</td>
<td>0.0107 (3.68)</td>
<td>0.0089 (2.49)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0077 (2.41)</td>
<td>0.0102 (4.01)</td>
<td>0.0111 (4.69)</td>
<td>0.0120 (4.71)</td>
<td>0.0123 (3.75)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0099 (3.18)</td>
<td>0.0108 (4.37)</td>
<td>0.0116 (4.90)</td>
<td>0.0129 (5.12)</td>
<td>0.0142 (4.41)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0127 (3.71)</td>
<td>0.0136 (4.70)</td>
<td>0.0140 (5.03)</td>
<td>0.0153 (5.24)</td>
<td>0.0151 (4.29)</td>
<td></td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0028 (1.39)</td>
<td>0.0035 (2.94)</td>
<td>0.0033 (1.90)</td>
<td>0.0037 (1.92)</td>
<td>0.0031 (1.35)</td>
<td></td>
</tr>
</tbody>
</table>

$\beta^{high} - \beta^{low} = 0.0026 (2.18)$

<table>
<thead>
<tr>
<th>Past Return Portfolio</th>
<th>Carhart Momentum Factor Portfolio: PRIYR</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0116 (2.63)</td>
<td>0.0087 (2.31)</td>
<td>0.0079 (2.17)</td>
<td>0.0077 (2.02)</td>
<td>0.0106 (2.11)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0084 (2.40)</td>
<td>0.0096 (3.34)</td>
<td>0.0092 (3.34)</td>
<td>0.0094 (3.40)</td>
<td>0.0088 (2.52)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0094 (2.91)</td>
<td>0.0098 (3.71)</td>
<td>0.0113 (4.69)</td>
<td>0.0111 (4.63)</td>
<td>0.0108 (3.48)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0115 (3.53)</td>
<td>0.0119 (4.57)</td>
<td>0.0122 (5.03)</td>
<td>0.0123 (5.10)</td>
<td>0.0133 (4.47)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0148 (4.12)</td>
<td>0.0156 (5.11)</td>
<td>0.0152 (5.42)</td>
<td>0.0153 (5.22)</td>
<td>0.0148 (4.45)</td>
<td></td>
</tr>
<tr>
<td>Wi-Lo</td>
<td>0.0031 (1.64)</td>
<td>0.0046 (2.53)</td>
<td>0.0052 (2.85)</td>
<td>0.0053 (2.74)</td>
<td>0.0043 (1.77)</td>
<td></td>
</tr>
</tbody>
</table>

$\beta^{high} - \beta^{low} = 0.0003 (0.28)$
References for Chapter 1


Géczy, Christopher C., 1996. Some generalized tests of mean-variance efficiency and performance. Working paper. Ohio State University, Columbus, OH.


55


CHAPTER 2

An Analysis of Risk, Pricing Anomalies, and Behavior: How Risky are Size, Book-to-Market, and Momentum Strategies?

2.1 Introduction

Efforts to explain the cross-section of expected stock returns via a multifactor linear asset pricing model have generally failed due to the existence of several empirical anomalies. Consider the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965), and Black (1972). Under the CAPM, market \( \beta \)'s (or the covariation of a security's return with the return on the market) should be sufficient to describe the cross-section of expected returns. However, many empirical anomalies are found to contradict this implication. Most notably Banz (1981) finds that size (market capitalization) has additional explanatory power over market \( \beta \) for describing the cross-section of expected stock returns. In addition, other variables have been shown to explain cross-sectional differences in average returns in the presence of both market \( \beta \) and size.\(^1\) Finally, Jegadeesh and Titman (1993) document the momentum anomaly, which appears unexplained by any of these

\(^1\)Most notably Bhandari (1988) finds that leverage, Stattman (1980) and Rosenberg, Reid, and Lanstein (1985) find that the ratio of book value of equity to market value of equity, and Basu (1983) finds that earnings-price ratios all have additional explanatory power over \( \beta \) and size. For an excellent discussion of this literature, see Fama and French (1992).
other anomalies. Specifically, momentum is the fact that stocks that have performed well in the past 3-12 months (winners) continue to perform well over the next 3-12 months, and stocks that have performed poorly (losers) continue to perform poorly.

The existence of these so-called anomalies may be due to several sources, which we can broadly group into three distinct categories. The first category is risk-based explanations. In other words, the first possibility is that these anomalies arise because the model is not capturing a component of systematic risk, which these firm characteristics may be correlated with. The second set of explanations are behavioral, suggesting that these anomalies arise because investors care about certain firm attributes, perhaps more so than portfolio volatility, or that investors act irrationally to information, or have psychological biases in their interpretation of information, all of which may induce an apparent relation between average returns and these characteristics. Finally, the third set of explanations for these anomalies may be due to flawed methodology. Specifically, there may be biases in computing returns due to firm survivorship or microstructure effects such as bid-ask bounce, as well as other statistical errors. In addition, the numerous and extensive studies that have been conducted on the U.S. equity market suggests that data mining may be a serious concern.

While issues of flawed methodology are important, most of these studies take great care in avoiding these empirical pitfalls. In addition, recent research has shown that many of these anomalies are present in international equity markets.

---

2 Several recent attempts to explain these anomalies through investor behavior have had success. For examples of such work see Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), and Hong and Stein (1998).

3 For examples of studies dealing with these issues see Lo and MacKinlay (1990), Kaul and Nimalendran (1990), Conrad and Kaul (1993), and Jegadeesh and Titman (1995).

providing out of sample evidence which alleviates data mining concerns. Consequently, in this paper we focus on discerning between risk-based and behavioral explanations for these anomalies.

We begin with risk-based explanations. Fama and French (1992, 1993, 1996) find that size and book-to-market equity (BE/ME) subsume the other cross-sectional anomalies (except for momentum) and form factor-mimicking portfolios derived from size and BE/ME which seem to capture the cross-section of expected stock returns. Specifically, they form two zero-investment portfolios: a portfolio long small stocks and short large stocks, and a portfolio long high BE/ME firms and short low BE/ME firms, designed to mimic factors in an APT framework, or that represent the changing investment opportunity set as in Merton (1973). However, as Fama and French (1996) concede, while they can capture a whole host of pricing anomalies with their three factor model (the two size and BE/ME portfolios plus the excess return on the market), momentum still remains the greatest challenge to their model. To reconcile the momentum anomaly, one could form a momentum (winners minus losers) “factor-mimicking” portfolio as Carhart (1997) does and claim that a four factor model now governs the return generating process for stocks (that is, until the next anomaly violates this model). However, this process of adding variables to the regression for each new anomaly is somewhat unappealing. Furthermore, to date there is no established link between these anomalies and the covariance matrix of asset returns. Certainly Fama and French and others have demonstrated that covariation with these “factor-mimicking” portfolios is associated with higher average returns, but as Ferson, Sarkissian, and Simin (1998) point out, factor-mimicking portfolios designed from anomalies will be statistically correlated with returns by definition. Thus, to call such a relation evidence of risk is treacherous. In addition, as Daniel and Titman

and references therein for evidence of a host of anomalies in international markets.
(1997) show, firm characteristics have much stronger explanatory power than do loadings on these ‘characteristic-generated factors’ for describing the cross-section of expected returns. In fact, they demonstrate that controlling for dispersion in the characteristic, there is no relation between factor loadings and average returns. Such evidence appears to contradict these risk-based interpretations.

On the behavioral side, if these anomalies are due to investor sentiment or irrationality, then profitable trading strategies can be designed to exploit these anomalies and extract pure rents for rational investors. For instance, Jegadeesh and Titman (1993) offer a behavioral explanation for momentum, claiming that positive stock return autocorrelation is driven by “delayed price reactions to firm-specific information” (p. 67). However, if investors persistently and irrationally underreact to firm-specific information, then rational investors can profit from their irrational counterparts. In an economy with an infinite number of assets, a self-financing portfolio long high past return stocks and short low past return stocks (weighted to have a similar factor beta configuration as the winner portfolio) could be created with zero risk and positive expected return, implying that asymptotic arbitrage opportunities would exist. In a large asset economy, therefore, prices of high past return stocks should be bid up and those of low past return stocks bid down until the profitability of such strategies disappears. Unless ‘irrational’ investors exist who are willing to hold the opposite position (i.e., long losing stocks and short winning stocks), a wealth depleting strategy. Thus, behavioral explanations appear to give rise to enormous profit opportunities. Indeed, momentum strategies alone generate an average 12% premium per year, per dollar long on a zero-cost portfolio. The question, of course, is how risky are these strategies?

5 Even with short sales restrictions, the same argument holds as rational investors would tilt their portfolios toward high past return stocks and away from low past return stocks.

60
This paper attempts to address these issues by answering two important, yet related questions. The first is to what extent do these anomalies contribute to the covariance matrix of asset returns? More specifically, we know factor-mimicking portfolios formed from these anomalies will have significant power for explaining average returns, but do they also forecast risk (e.g., variances and covariances)? The second question is very similar, but is asked from a behavioral perspective: how risky are investment strategies designed to exploit these anomalies?

Since the three strongest and most cited anomalies are those of size, BE/ME, and momentum, we focus on these, and assess the plausibility of risk-based and behavioral explanations for their existence.

Examining the unconditional sample distribution of size, BE/ME, and momentum strategies, it appears unlikely that these anomalies are due to risk. However, when we allow the volatility of these strategies to change through time, risk-based explanations appear more reasonable. Using a multivariate GARCH (Generalized Autoregressive Conditional Heteroscedastic) parameterization for asset return second moments, we find that size and BE/ME appear to be spanned by systematic sources of variation (proxied by principal components analysis), but that momentum does not. Furthermore, forming minimum variance portfolios that load on each of the anomalies individually, while hedging out factors associated with the other anomalies and with other sources of variation, we find that only for momentum can we generate a strategy with ‘abnormally large’ Sharpe ratios, suggesting that momentum is unrelated to risk. However, both size and BE/ME appear to have substantial variation, even when we control for exposure to the other anomalies as well as the principal components factors. Thus, both size and BE/ME appear more likely to be driven by risk.

In addition, we evaluate the performance of various factor portfolios derived
from the anomalies in explaining the cross-section of expected returns as well as the future covariance structure of asset returns. We find that these factors have significant explanatory power and that an industry momentum factor outperforms all other factors in capturing the future dispersion in mean stock returns as well as the future covariances of stock returns, consistent with a risk-based explanation for momentum. However, the contribution to the covariance matrix of returns made by the momentum factor is much smaller than those made by size and BE/ME, suggesting that momentum is a less likely candidate as a proxy for risk. Following the same analysis, evidence is found of strong predictability by the market portfolio for future risk, and the market remains the single most important factor contributing to the covariance matrix of asset returns.

Finally, although momentum appears unexplained by risk, such conclusions appear to depend on our definition of risk. For risk defined as the second moment of returns, such claims appear true, but when we include a co-skewness and co-kurtosis factor in the model, momentum profits drop dramatically. Thus, the mean-variance tradeoff may not be a suitable performance measure for momentum strategies, which seem to induce negative skewness and excess kurtosis, higher moments investors may care about.

The rest of the paper is organized as follows. Section 2 discusses the portfolio/trading strategies designed to exploit the anomalies, and presents the distribution of the profits from these strategies over time. Section 3 presents a model for time-varying volatility, highlighting several proposed asset pricing models, and evaluates the risk of these strategies conditionally. Section 4 then evaluates how risky these strategies are, controlling for other possible sources of variation. Section 5 tests the out of sample performance of various competing factor models derived from the anomalies, for predicting future expected returns and covari-
ances, and assesses the contribution each makes to the covariance matrix of asset returns. Finally, Section 6 concludes the paper.

2.2 Profitable Trading Strategies

If these anomalies are not driven by risk, then rational investors will be able to profit from their existence. To exploit these anomalies, we form self-financing portfolios based on firm characteristics, and then test the performance of these strategies out of sample. More specifically, to take advantage of the size effect, for example, we go long the 30% smallest stocks (based on market capitalization in the previous year) and short the biggest 30% of firms. This small minus big (SMB) portfolio is the same portfolio constructed by Fama and French (1993) and used as a factor-mimicking portfolio in their model. We will also employ this portfolio as a "factor" later in the paper to determine its validity as a proxy for risk. For now, however, we will treat this portfolio as a device to exploit the size anomaly. Likewise, we form a high minus low (HML) BE/ME portfolio, a winners minus losers momentum portfolio, based on individual stocks and using a prior one year return ranking (PRIYR), and an industry momentum portfolio (IM) long the three highest past 1-month return industries and short the three lowest, where industries are defined by groups of two digit SIC codes. The construction of each of these portfolios are described in detail in Appendix 2.7. We employ two momentum portfolios since Moskowitz (1997) finds that momentum is largely an industry phenomenon. However, employing both an individual stock momentum strategy as well as an industry momentum strategy may shed further light on what drives this anomaly.
2.2.1 Distribution of Trading Strategy Profits

Figure 2.1 plots the distribution of the profits (returns per dollar long) from these strategies over time. For comparison, the excess return on the market (CRSP value-weighted index) is also reported. As shown in the table at the bottom of Figure 2.1, the momentum strategies are both more profitable and more risky. The means and medians of the momentum strategies are higher than those for \( SMB \) and \( HML \), and \( IM \) and \( PRIYR \) have a larger number of positive realizations. In addition, however, the standard deviation and range of the profits from these momentum strategies are higher than either \( SMB \) or \( HML \), suggesting that momentum strategies are more risky than size and BE/ME strategies. This table also reports the unconditional market \( \beta \) of these portfolios by regressing each of these strategies on the excess return of the market. Only \( SMB \) loads positively on the market, and all four strategies have very small \( \beta \)'s. In fact, \( IM \)'s market \( \beta \) is statistically indistinguishable from zero. In addition, we compute the Gibbons, Ross, and Shanken (1989) test statistic that the abnormal returns (relative to the market) from these strategies are jointly zero. The F-statistic from this test is 9.25, which strongly rejects the null hypothesis that abnormal returns are zero. Thus, these anomalies do not appear to be captured by the market.

However, MacKinlay (1995) shows that the zero-intercept F-test (such as Gibbons, Ross, and Shanken (1989)) can be useful for distinguishing between risk-based and non-risk alternatives. Under the risk-based alternative, an upper bound exists on the maximum squared Sharpe ratio in the economy \( S^2 \), but under a non-risk alternative, the maximum Sharpe measure is not bounded as the number of assets in the economy, \( N \), increases. MacKinlay (1995) shows that the distribution of the F-statistic under the null hypothesis can be compared to
those under the risk and non-risk alternatives. If rejection of the null pricing model is due to missing risk factors, then an optimal orthogonal portfolio (i.e., orthogonal to the risk factors or factor-mimicking portfolios) can be designed to capture the abnormal returns leading to rejection. Furthermore, under the risk-based alternative, the optimal orthogonal portfolio, \( h \), will have a squared Sharpe ratio bounded by the maximum squared Sharpe ratio in the economy (i.e., \( S_h^2 \leq S_q^2 \)). The bound on the maximum Sharpe measure translates into a bound on the noncentrality parameter of the risk-based alternative distribution, which follows a non-central \( F \). Thus, under the risk-based alternative, the \( F \)-statistic follows.

\[
F_1 \equiv \frac{T - N - K}{N} \left[ \frac{\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}}{1 + S_h^2} \right] \sim F_{N,N - K}(\phi) \tag{2.1}
\]

\[
\phi = \frac{T}{1 + S_h^2} \left[ S_h^2 \right] \tag{2.2}
\]

where \( F_1 \) is the zero-intercept statistic, \( T \) is the number of time periods, \( N \) is the number of assets, \( K \) is the number of factor portfolios, \( \hat{\alpha} \) is the vector of abnormal returns (relative to the \( K \) factor portfolios), \( \hat{\Sigma} \) is the residual covariance matrix, \( S_h^2 \) is the maximum squared Sharpe measure of the \( K \) factor portfolios, and \( \phi \) is the noncentrality parameter where \( S_h^2 \) represents the squared Sharpe measure of the optimal orthogonal portfolio, which is a portfolio orthogonal to the \( K \) factors designed to capture the maximal amount of abnormal returns leading to rejection of the null.

Under the non-risk alternative, \( F_1 \) still follows a non-central \( F \) distribution, but no bound exists on the non-centrality parameter, \( \phi \), since there is no link between pricing deviations (\( \hat{\alpha} \)) and the residual covariance matrix (\( \hat{\Sigma} \)) under a non-risk based explanation. Therefore, \( \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} \) replaces \( S_h^2 \) in equation (2.2). The distribution of the zero intercept \( F \)-statistic (\( F_1 \)) under the null hypothesis of no abnormal performance (which follows a central \( F \) distribution, i.e., \( \phi = 0 \))
can therefore be compared to those under the risk and non-risk alternatives, to
determine the likelihood of these alternatives when the null is rejected.

We can examine the likelihood of risk and non-risk alternative hypotheses by
estimating the parameters of these alternative distributions. We estimate the ex
post (monthly) maximum squared Sharpe ratio of the market to be $S^2 = 0.0156$,
and, subtracting this from the maximum squared Sharpe measure of the economy,
$S^2_q$ (which we define to be 0.031 as in MacKinlay (1995)), provides an estimate for
$S^2_h$.\footnote{Since portfolio $h$ is orthogonal to the factors, and since it is
designed to capture abnormal returns, the maximum squared Sharpe ratio in the
 economy will be the sum of the squared Sharpe measures of $h$ and the factor
portfolios (i.e., $S^2 = S^2_h + S^2_q$). See MacKinlay (1995) for
further details.} Under the non-risk alternative, no bound exists on $\phi$, thus $\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$ replaces
$S^2_h$ in (2.2), where $\hat{\alpha}$ and $\hat{\Sigma}$ are the intercepts and residual
covariance matrix from the regression of each of the four portfolios on the market. Based on these
parameters, the distribution of the F-statistic under the three hypotheses is.

\begin{align*}
\text{Null: } F_1 & \sim F_{4,378}(0) \\
\text{Risk-Based: } F_1 & \sim F_{4,378}(5.81) \\
\text{Non-Risk: } F_1 & \sim F_{4,378}(57.35)
\end{align*}

which are plotted in Figure 2.2. with the $F_1$ statistic of 9.25 highlighted by
the vertical line. From the figure, it is apparent that we can reject the risk-
based alternative, as the $F_1$ statistic is not covered by the risk-based alternative
distribution.

However, the Gibbons, Ross, and Shanken (1989) test assumes returns are
distributed under a normal distribution,\footnote{Even when we recompute the Gibbons, Ross, and Shanken (1989) test statistic assuming
returns are distributed elliptically, following Géczy (1997), $F_1 = 9.10$, which still strongly rejects
the null.} and the MacKinlay (1995) analysis assumes investors only care about the first two moments of returns. If investors
care about higher moments when considering the risk of a trading strategy, then Sharpe ratios will provide an inadequate measure of performance. Furthermore, these strategies are dynamic, particularly momentum strategies, which condition on past returns, and therefore, will likely follow non-Gaussian distributions. Examining the skewness and kurtosis of these strategies, we find that IM and SMB are slightly positively skewed, HML is slightly negatively skewed, and PR1YR is strongly negatively skewed. Furthermore, both IM and PR1YR have fatter tails than SMB and HML. and PR1YR has enormous fat tails (as evidenced by its kurtosis measure). Thus, an individual stock momentum strategy has very high probability of experiencing large negative returns. This is illustrated by the graph of these profits, as several times the PR1YR strategy experiences huge losses. When compared to the industry momentum strategy, IM dominates PR1YR on each of the first four moments, consistent with Moskowitz’s (1997) claim that industry momentum is the better (true) strategy.\(^9\) Finally, we reject soundly the Chi-squared test that the returns from these strategies are distributed normally. Thus, our analysis that risk-based explanations for these anomalies are unlikely may depend a lot on the assumptions underlying that analysis. Assumptions that are empirically false. Finally, our analysis examined these strategies unconditionally, assuming that the $\beta$'s (and risks) do not change over time. As we will see in the next section, this is a poor assumption.

\(^8\)As Harvey and Siddique (1996) suggest they do, evidenced by the existence of a premium on a conditional co-skewness variable.

\(^9\)The only statistic for which PR1YR dominates IM is the median, but the risk of PR1YR appears much higher.
2.3 Time-Varying Risk

Since the risks of these strategies are unlikely to be constant over time, in order to evaluate how risky size, BE/ME, and momentum strategies are, and whether the profits from such strategies represent fair compensation for risk, we need to examine the conditional volatility of such strategies. We model time-varying variances and covariances via a generalized autoregressive conditional heteroscedastic (GARCH) model. Two multivariate GARCH models are employed to obtain estimates of the conditional covariance matrix of asset returns, $H_t$. A typical GARCH($p, q$) parameterization for $H_t$ is

$$H_t = C'C + \sum_{i=1}^{p} A_i'\varepsilon_{t-i}\varepsilon_{t-i}'A_i + \sum_{j=1}^{q} B_jH_{t-j}B_j$$

(2.3)

where $C$, $A_i$ and $B_j$ are $(N \times N)$ matrices of constant coefficients, and $\varepsilon_t$'s are deviations of realized returns from their expected values, conditioned on information observed through time $t-1$,

$$\varepsilon_t = R_t - E_{t-1}[R_t], \quad \varepsilon_t|\Phi_{t-1} \sim N(0, H_t)$$

(2.4)

where $R_t$ is the $(N \times 1)$ vector of returns, and $E_{t-1}[\cdot] = E[\cdot|\Phi_{t-1}]$ is the expectation conditional on information revealed through time $t-1$.\(^{10}\)

Our first approach is to estimate a GARCH(1,1) model and restrict both $A$ and $B$ to be diagonal matrices. Thus, the variances in $H_t$ depend only on past squared residuals and an autoregressive component, while the covariances depend upon past cross-products of residuals and an autoregressive component.\(^{11}\) Thus,

---

\(^{10}\)In terms of implementation, we use the time-series sample average return from time 0 to time $t-1$ for $E_{t-1}[R_t]$.

\(^{11}\)This model corresponds to those used by Bollerslev, Engle, and Wooldridge (1998) and DeSantis and Gerard (1997), for example. Other GARCH models assume constant correlation in order to reduce the number of parameter estimates. We take the former approach because it allows for correlation among the assets to change over time.
equation (2.3) becomes.

\[ H_t = C'C + aa' \cdot \epsilon_t \epsilon_{t-1}' + bb' \cdot H_{t-1} \]  
(2.5)

where \(a\) and \(b\) are the \((N \times 1)\) vectors of the diagonal elements of \(A\) and \(B\), and \(\cdot\) denotes the Hadamard element-by-element matrix product. In addition, Ding and Engle (1994) show that if the system is covariance stationary, the unconditional covariance matrix is.

\[ H_0 = C'C \cdot (\mu' - aa' - bb')^{-1}, \]  
(2.6)

where \(\mu\) is an \((N \times 1)\) vector of ones. Thus, they suggest replacing \(C'C\) with \(H_0 \cdot (\mu' - aa' - bb')\), where the sample covariance matrix can be used as a consistent estimator for \(H_0\), and the conditional covariance matrix becomes.

\[ H_t = H_0 \cdot (\mu' - aa' - bb') + aa' \cdot \epsilon_{t-1} \epsilon_{t-1}' + bb' \cdot H_{t-1}. \]  
(2.7)

This is our first model for conditional covariances, requiring only \(2N\) parameters to be estimated.

Our second GARCH model uses the methodology of Ledoit and Santa-Clara (1998), who develop a unique multivariate GARCH(1,1) technique that allows them to impose less structure on the parameters. Their GARCH(1,1) parameterization is the following:

\[ H_t = C + A \cdot \epsilon_{t-1} \epsilon_{t-1}' + B \cdot H_{t-1} \]  
(2.8)

Briefly, Ledoit and Santa-Clara (1998) estimate the GARCH(1,1) parameters for each pair of assets \((N = 2)\) independently, as a starting point. They then minimize the distance between each of these bivariate estimates and the space of potential values for \(C, A,\) and \(B\) that satisfy conditions ensuring positive definite and covariance stationary matrices. The advantage of this approach is that it does
not impose the restriction that $A$ and $B$ are diagonal matrices. and estimates $C$ simultaneously with the other parameters. Thus, $3N + \frac{N(N+1)}{2}$ parameters are estimated, providing a less restrictive model, and more robust estimates of conditional volatility.$^{12}$ These parameters are then used to reconstruct a time-series of covariance matrices, where the unconditional sample covariance matrix is used as a starting point for $H_0$.

Both models are estimated via maximum likelihood, where, under conditional normality, the log-likelihood function is given by,

$$\ln(L(\theta)) = -\frac{T \cdot N}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln|H_t(\theta)| - \frac{1}{2} \sum_{t=1}^{T} \epsilon_t(\theta)^\top H_t(\theta)^{-1} \epsilon_t(\theta)$$

(2.9)

where $\theta$ is the set of parameters $\{A, B, C\}$.

The conditional covariance matrices are estimated for 32 assets: 20 industry portfolios, six size and BE/ME portfolios, the CRSP value-weighted index, and five 12-month past return sorted portfolios. The description of these portfolios can be found in Appendix 2.7. We only report the results using the Ledoit and Santa-Clara (1998) estimates for much of the analysis in the paper, as the results using the Ding and Engle (1994) approach were very similar.

\[2.3.1\] Competing Factor Models

Throughout the paper, we employ several asset pricing models which incorporate “factor-mimicking” portfolios derived from the anomalies. This section briefly describes the various pricing models employed in the paper.

$^{12}$For a detailed description of their algorithm, see Ledoit and Santa-Clara (1998).
2.3.1.1 Fama and French (1993) Three Factor Model

The first model is a conditional version of Fama and French’s (1993) three factor model, which represents a three factor version of the APT or where the factors represent the three portfolios that best capture state variables that determine the changing investment opportunity set as in Merton (1973). The three factors are, respectively, the market, a size factor portfolio \( SMB \), and a “distress” factor proxied by \( HML \). The model can be represented as follows.

\[
R_t - r_{ft} = \beta_{1t}[R_{mt} - r_{ft}] + \beta_{2t}SMB_t + \beta_{3t}HML_t + \epsilon_t, \quad (2.10)
\]

\[
\epsilon_t|\Phi_{t-1} \sim N(0, H_t) \quad (2.11)
\]

\[
\beta_{iF} \equiv [\beta_{1t} : \beta_{2t} : \beta_{3t}] = H_t w(w'H_t w)^{-1}, \quad \forall t \quad (2.12)
\]

where \( w \) is an \((N \times 3)\) matrix of zeros with a 1 in the 27th row of the first column; 1/3 in the 21st, 22nd, and 23rd rows and -1/3 in the 24th, 25th, and 26th rows of the second column; and 1/2 in the 21st and 24th rows and -1/2 in the 23rd and 26th rows of the third column.

2.3.1.2 Carhart (1997) Four Factor Model

The second model is a conditional version of Carhart’s (1997) four factor model, which adds the momentum factor, \( PR1YR \), to the FF model. The model is represented as follows.

\[
R_t - r_{ft} = \beta_{1t}[R_{mt} - r_{ft}] + \beta_{2t}SMB_t + \beta_{3t}HML_t + \beta_{4t}PR1YR_t + \epsilon_t \quad (2.13)
\]

\[
\epsilon_t|\Phi_{t-1} \sim N(0, H_t) \quad (2.14)
\]

\[
\beta_{iC} \equiv [\beta_{1t} : \beta_{2t} : \beta_{3t} : \beta_{4t}] = H_t w(w'H_t w)^{-1}, \quad \forall t \quad (2.15)
\]

where \( w \) is an \((N \times 4)\) matrix whose first three columns are as described above for the FF model, and whose fourth column is a vector of zeros with a 1 in the
32\textsuperscript{nd} row and a \(-1\) in the 28\textsuperscript{th} row.

2.3.1.3 Industry Momentum Four Factor Model

The third model replaces the Carhart (1997) momentum factor $PR1YR$ with an industry momentum factor, $IM$. The model is represented as follows.

$$R_t - r_{ft} = \beta_1 [R_{mt} - r_{ft}] + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 IM_t + \epsilon_t \text{, \hspace{1cm} (2.16)}$$

$$\epsilon_t | \Phi_{t-1} \sim N(0, H_t) \text{ \hspace{1cm} (2.17)}$$

$$\beta_t^{Mo} = [\beta_{1t} : \beta_{2t} : \beta_{3t} : \beta_{4t}] = H_t w_t (w'_t H_t w_t)^{-1}, \forall t \text{ \hspace{1cm} (2.18)}$$

where $w_t$ is an $(N \times 4)$ matrix whose first three columns are as described above for the FF model, and whose fourth column is a vector of zeros with $1/3$ in each of the rows corresponding to the three industries which exhibited the highest returns last period (time $t - 2$) and $-1/3$ in the rows corresponding to the three industries which exhibited the lowest returns last period.\textsuperscript{13}

2.3.1.4 Out of Sample Principal Components Model (PC)

This model extracts the first four principal components from $H_{t-1}$, rescales the eigenvectors to sum to one (to form portfolio weights), and multiplies the rescaled weights by the excess return vector at time $t$. That is, we compute the principal components pricing factor weights at time $t - 1$, and apply them to time $t$ returns. In other words, we want to test how well the principal components perform out of sample, since, by construction, they will perform the best of any four factors in sample. Specifically, the model is represented as follows,

$$R_t - r_{ft} = \beta_t PC_t + \epsilon_t, \epsilon_t | \Phi_{t-1} \sim N(0, H_t) \text{ \hspace{1cm} (2.19)}$$

\textsuperscript{13}Since the covariance matrix at time $t$, $H_t$, is a forecast based on information observed through time $t - 1$, it is necessary to rank industries on their $t - 2$ returns, since the $t - 1$ returns will be the ‘realized’ (within sample) returns corresponding to $H_t$. 

72
\[ \beta_{t}^{PC} = H_{t}w_{t-1}(w'_{t-1}H_{t}w_{t-1})^{-1}. \forall t \] (2.20)

where \( w_{t-1} \) is an \((N \times 4)\) matrix whose columns are the scaled (to sum to one) eigenvectors associated with the four largest eigenvalues of \( H_{t-1} \), and \( PC_{t} \) is the \((4 \times 1)\) vector of principal components factor realizations at time \( t \) (i.e., \( PC_{t} = w'_{t-1}[R_{t} - r_{ft}] \)).

### 2.3.1.5 Unconditional Principal Components Model (UPC)

The final model employed in the paper takes the first four principal components from the unconditional sample covariance matrix of the 32 portfolio returns. The eigenvectors associated with the four largest eigenvalues are rescaled to sum to one, and these portfolio weights are then held constant over time and applied to the excess returns vector at each time \( t \). Obviously, this model will not be tested out of sample, but may provide an interesting benchmark from which to compare the other models. The purpose of employing this model is to test how well the unconditional factors perform conditionally. That is, we know these factors fit the unconditional covariance matrix very well by design, but they may perform poorly at specific points in time, and may not fit the conditional covariance matrix. The model is presented as follows.

\[ R_{t} - r_{ft} = \beta_{t}^{UPC} + \epsilon_{t}, \quad \epsilon_{t}|\Phi_{t-1} \sim N(0, H_{t}) \] (2.21)

\[ \beta_{t}^{UPC} = H_{t}w(w'H_{t}w)^{-1}, \quad \forall t \] (2.22)

where \( w \) is an \((N \times 4)\) matrix whose columns are the scaled (to sum to one) eigenvectors associated with the four largest eigenvalues of the unconditional sample covariance matrix, and \( UPC_{t} \) is the \((4 \times 1)\) vector of principal components factor realizations at time \( t \) (i.e., \( UPC_{t} = w'[R_{t} - r_{ft}] \)).
2.3.2 Conditional Tests

So far, we have demonstrated that it is highly unlikely that these anomalies can be explained by risk when we define risk to be the second moment of returns, and when we assume risk is constant over time. Likewise, other studies, including Jegadeesh and Titman (1993) and MacKinlay (1995), come to a similar conclusion under the same assumptions. In this section, we relax these assumptions to see if these anomalies are better captured by risk when we allow for time-variation in the covariance matrix of asset returns, and when we include higher moments in our analysis.

This paper is not the first to allow for changing risk over time, however. Fama and French (1993, 1996, 1997), Daniel and Titman (1997), Brennan, Chordia, and Subrahmanyam (1998), and Grundy and Martin (1997) for instance all allow for changing factor loadings over time. However, estimates of these time-varying loadings are found via rolling regressions, which employ the prior 3-5 years of data. We take a different approach, utilizing our GARCH-estimated covariance matrices, which we feel provide a more accurate picture of conditional volatility. In addition, most of the studies mentioned (except Brennan, Chordia, and Subrahmanyam (1998)) focus on the CAPM and the three factor Fama and French (1993) model as the pricing benchmark. We employ two additional models based on principal components: the unconditional principal components model, which is similar to the Connor and Korajczyk (1988) principal components factors, which Brennan, Chordia, and Subrahmanyam (1998) employ in their study, and the conditional principal components model, which has never been tested. The choice of model may be significant. For instance, in assessing momentum, Grundy and Martin (1997) find that high past return securities load higher on the Fama and French (1993) factors than do low past return securities, suggesting
that momentum strategies are tilted towards systematically riskier stocks. But, they find that the realizations of the Fama and French (1993) factors do not correspond to the exposure of the momentum strategy to the factors. In other words, factor realizations are negative precisely when high past return stocks load higher on the factors than do low past return stocks. Thus, things seem to move in the wrong direction for momentum to be explained by the model. In addition, both Grundy and Martin (1997) and Moskowitz (1997) find that there is no positive autocorrelation among the Fama and French (1993) factors (at the 6-month level), and Brennan, Chordia, and Subrahmanyam (1998) find negative autocorrelation among the Connor and Korajczyk (1988) factors, suggesting that the momentum anomaly is unlikely to be explained by these models. Examining the factors from the conditional principal components model, however, we find significant positive autocorrelation among the factors (at 1-12 month intervals). Thus, the choice of model may be important, and to rule out momentum as being explained by risk may be premature.

To test whether these models can capture the various anomalies once we allow for time-varying loadings on the factors, we compute the residual returns of our 32 portfolios under each of the factor models presented in Section 2.3.1, where we compute factor loadings using the GARCH estimated covariance matrices.\textsuperscript{14} We

\textsuperscript{14}Figure 2.3 plots the GARCH-estimated market $\beta$s against rolling regression $\beta$ estimates which use the prior 36 months of returns. Specifically, we extract the $\beta$ estimates from each conditional covariance matrix $H_t$ for each of the four strategies (i.e., $SMB$, $HML$, $IM$, and $PR1YR$), and plot them over time. We also regress each of the last 36 months of portfolio returns on the past 36 months of market realizations in excess of the risk-free rate. The estimated coefficients from these rolling regressions are also plotted over time. As the figure shows, the GARCH-estimated $\beta$s are significantly different from the rolling regression estimates. Correlations between the estimates for each strategy are surprisingly low. Only for $SMB$ is there significant positive correlation between GARCH and regression $\beta$s. In addition, the regression $\beta$s are generally much more smooth than the GARCH estimates, suggesting that the regression may be missing (smoothing) important information about the changing risks of these strategies over time. Thus, the GARCH estimates may provide a more accurate picture of true risk exposure, and at the very least provide a unique set of estimates to test the profitability of these strategies.
then re-form each of the strategies designed to exploit the anomalies using the residual returns under each factor model. For example, the industry momentum strategy entails ranking the 20 industries on their past 1-month raw returns, and forming a portfolio long the highest three past return industries, and short the lowest three, where the residual returns under each factor model are computed for this strategy over the next month. For instance, for the CAPM, $\beta$'s are extracted from $H_t$ for $IM$, and then applied to the time $t$ return of the market. Specifically, the residual return for $IM$ at time $t$ is,

$$IM_t = \hat{\beta}_t[R_{mt} - r_{ft}]$$

(2.23)

where $\hat{\beta}_t$ is the factor loading of $IM$ on the market based on the GARCH-estimated covariance matrix $H_t$ (which is a forecasted covariance matrix based on realizations up to time $t - 1$).

The results are reported in Table 2.1. As the Table demonstrates, only about 10% of the 1-month industry momentum profits are accounted for by any of the CAPM, FF, Carhart, or out of sample principal components models, and the unconditional principal components actually exaggerate the profits. More interestingly, however, is the standard deviation of the residual profits exhibited in the table. The size of the residual momentum profits are inversely related to the standard deviation of the profits (as evidenced by the t-stats), suggesting that the models may not be accounting for a component of systematic risk. This suggests that an industry momentum factor may be a reasonable proxy for a missing risk factor, a conjecture we will test throughout the paper.\footnote{This is consistent with the findings in Brennan, Chordia, and Subrahmanyam (1998), who find negative autocorrelation among the Connor and Korajczyk (1988) factors.}

\footnote{It is interesting to note that the Carhart model, with its individual stock momentum factor, $PRIYR$, does not account for industry momentum very well, consistent with the findings of Moskowitz (1997).}
2.3.2.1 Co-skewness and Kurtosis

However, since the industry momentum strategy exhibited a strong departure from normality, and since Kraus and Litzenberger (1976) and Harvey and Siddique (1996) document a premium for co-skewness, we add a co-skewness variable to each of our models in order to see if this is the missing component leading to apparently abnormal profits.

We estimate the premium for co-skewness by first regressing individual stock returns on the excess return of the market plus the squared excess market return,

\[ \bar{r}_{jt} - r_{ft} = \alpha + \beta_t[\bar{R}_{mt} - r_{ft}] + \gamma_t[\bar{R}_{mt} - r_{ft}]^2 + \varepsilon_t. \]  

(2.24)

We run this regression for each stock at each point in time using the prior 36 months of returns to obtain estimates of \( \gamma_t \), the co-skewness coefficient. We then form a portfolio long the 30% of stocks with the highest co-skewness (\( \hat{\gamma}_t \)) and short the lowest 30% of firms, where stocks are value-weighted within these quintiles, and stocks with less than 24 observations are excluded. This procedure is repeated at every time \( t \), obtaining a time-series of returns on a portfolio designed to capture the premium associated with co-skewness with the market. We call this portfolio \( Skew \), and add it to our factor models. For example, the Fama and French (1993) three factor model now becomes.

\[ R_t - r_{ft} = \beta_{1t}[R_{mt} - r_{ft}] + \beta_{2t}SMB_t + \beta_{3t}HML_t + \delta_tSkew_t + \xi_t. \]  

(2.25)

We then estimate the loadings on these factors of our industry momentum portfolio, \( IM \), using rolling regressions with the past 36 months of returns (since we do not have the \( Skew \) portfolio in our GARCH matrices). Estimates of these factor loadings are then applied to the time \( t \) realizations of the factors to adjust \( IM \) for the market, \( SMB \), \( HML \), and \( Skew \). As Table 2.1 shows, adding the
co-skewness variable reduces the profitability of IM under all factor models, and seems to eliminate most of the profits in combination with the Fama and French, Carhart, and principal components factors.

Finally, we also add a co-kurtosis variable to (2.25), by forming a self-financing portfolio long stocks with high coefficients on $[R_{mt} - r_{ft}]^3$ (controlling for co-skewness) and short stocks with low co-kurtosis in the same manner. The resulting portfolio, $Kurt$, is then added to (2.25) as an additional factor, and the same analysis is conducted. Kurtosis may be important to investors because it increases the probability of large returns (both positive and negative). For instance, if investors dislike losses more than they prefer gains (as Kahnemann and Tversky’s Prospect theory suggests), then investors will be averse to excess kurtosis. We add the co-skewness and co-kurtosis factors to the model to account for these preferences (if they exist). As Table 2.1 shows, adding the co-kurtosis variable also reduces the profitability of IM significantly, making the average profits from the strategy statistically indifferent from zero.

Table 2.1 also reports the residual profits for SMB, HML, and PRIYR. As the table indicates, PRIYR, the individual stock momentum profits, are exaggerated by all of the factor models except the industry momentum model, which seems to capture the bulk of these profits (consistent with Moskowitz (1997)). Adding the co-skewness and co-kurtosis variables to each model, PRIYR profits seem to decline dramatically under all models, similar to the IM profits. SMB, which is not very strong anyway, appears to be subsumed by UPC, the unconditional principal components factors, and adding the co-skewness and co-kurtosis variables seem to eliminate these profits entirely. Finally, HML profits are exaggerated by both principal components models, and are not explained by the CAPM. Adding the higher moment factors to the models actually exaggerates
profits under the CAPM. However, the profits are reduced under the principal components models when the co-skewness and co-kurtosis variables are added.

Thus, while we cannot explain the profitability of these trading strategies by our factor models, even when we account for the time-variation in factor loadings, we are able to capture a significant portion of these profits when we include higher moment factors. This suggests that these strategies induce exposure to higher moments that investors care about. Hence, when we account for the premium associated with these higher moments, the apparent profitability of these strategies are reduced. However, while co-skewness and co-kurtosis factors seem to add explanatory power for the anomalies, the economic significance of the trading strategy profits (other than size) are still non-trivial.

2.3.3 Common Volatility (ARCH) Tests

If these anomalies are driven by systematic risk, then strategies designed to exploit them should be spanned by systematic risk factors. Thus far, we have only examined the first moments of these strategies in relation to systematic risk. If these anomalies are related to risk, however, then we also need to determine if their second moments can be explained by risk. Therefore, we conduct a test for whether these strategies follow a common volatility process with proxies for systematic risk factors. Specifically, since these strategies exhibit significant time-varying volatility (evidenced by the rejection of the ARCH(12) tests in Table 2.2),

\[ r_{it} = \alpha_i + \sum_{j=1}^{q} r_{i,t-j} + \epsilon_{it} \]  

\[ (2.26) \]

\[^{17}\text{Testing for ARCH effects in the time-series of returns requires estimating the following regression.}\]
to determine if the second moments of these strategies are spanned by sources of systematic variation.

The idea related to cointegration is that a feature will be common to two time series if a linear combination of the two series fails to have the feature, even though each of the individual series exhibits the feature. That is, two series that display time-varying variances (ARCH) individually, can be combined into a portfolio that has constant variance, if the two series share a common volatility process. The economically appealing aspect of this framework is that the common feature can be thought of as a common risk factor in a multi-factor linear pricing model.

For example, suppose two series of returns $\tilde{x}_t$ and $\tilde{y}_t$ can be expressed as,

$$\tilde{x}_t = \mu_x + \beta_x \tilde{F}_{1t} + \tilde{\varepsilon}_{xt} \quad (2.28)$$
$$\tilde{y}_t = \mu_y + \beta_y \tilde{F}_{2t} + \tilde{\varepsilon}_{yt} \quad (2.29)$$

where $\mu_x$ and $\mu_y$ are the unconditional means of $x$ and $y$, respectively, $\tilde{F}_{1t}$ and $\tilde{F}_{2t}$ are time-varying risk factors, and $\tilde{\varepsilon}_{xt}$ and $\tilde{\varepsilon}_{yt}$ are unexpected noise. In addition, assume the unexpected components of returns and factors follow an unknown distribution $G(\cdot)$ with zero mean and finite variance, and that the disturbances have constant variance, while the factors have time-varying variances that follow an ARCH process:

$$\tilde{\varepsilon}_{xt}|\Phi_{t-1} \sim G(0, \sigma_x^2), \quad \tilde{\varepsilon}_{yt}|\Phi_{t-1} \sim G(0, \sigma_y^2) \quad (2.30)$$

$$\tilde{\varepsilon}_{it}^2 = a_i + \sum_{j=1}^{p} \tilde{\varepsilon}_{i,t-j}^2 + u_{it} \quad (2.27)$$

where $r_{it}$ are the returns of asset $i$ at time $t$, and $\tilde{\varepsilon}_{it}$ are the estimated residuals from the first regression. Engle (1982) shows that the test statistic for no ARCH in the time series is obtained by multiplying the $R^2$ from the second regression by the sample size $T$, which has an asymptotic chi-squared distribution with degrees of freedom equal to the number of lagged squared residuals, $p$ (i.e., $T \times R^2 \sim \chi^2(p)$). The results in Table 2.2 are for $q = 24$ and $p = 12$, referred to as ARCH(12).
\[ \tilde{F}_{1t} | \Phi_{t-1} \sim G(0, H_{1,t}), \quad \tilde{F}_{2t} | \Phi_{t-1} \sim G(0, H_{2,t}) \]  
\[ H_{1,t} = \alpha_1 + B_1 \tilde{F}_{1t-1}^2, \quad H_{2,t} = \alpha_2 + B_2 \tilde{F}_{2t-1}^2 \]  
\[ E[\tilde{F}_{1t} \tilde{\varepsilon}_{zt} | \Phi_{t-1}] = 0, \quad E[\tilde{F}_{2t} \tilde{\varepsilon}_{yt} | \Phi_{t-1}] = 0. \]

where \( \Phi_{t-1} \) is the information set available to investors at time \( t \), including all past realizations of \( \tilde{x} \) and \( \tilde{y} \).

The variance of the two return series are, respectively, \( \text{Var}(\tilde{x}_t) = \beta_x^2 H_{1,t} + \sigma_x^2 \) and \( \text{Var}(\tilde{y}_t) = \beta_y^2 H_{2,t} + \sigma_y^2 \) both of which are time-varying. A common ARCH feature between the two series is present if there exists a non-zero linear combination of \( \tilde{x}_t \) and \( \tilde{y}_t \) that does not exhibit an ARCH effect. That is, a series \( \tilde{z}_t = \tilde{x}_t - \omega \tilde{y}_t \) which has constant variance through time. In this example, when \( \tilde{F}_1 = \tilde{F}_2 \) and \( \omega = \beta_x/\beta_y \), \( \tilde{z}_t \) will have constant variance (i.e., \( \text{Var}(\tilde{z}_t) = \sigma_x^2 + \omega^2 \sigma_y^2 \)). In other words, when the two series of returns are generated by the same risk factor, we can create a portfolio which exhibits constant variation.

In general, \( \tilde{y}_t \) can be a \((T \times K)\) matrix of returns, where a test for a common feature is a test of whether the second moments of the \( K \) assets in \( \tilde{y}_t \) span the second moments of \( \tilde{x}_t \). The common ARCH feature test is an estimate of the \( T \times R^2 \) statistic obtained by solving the following minimization problem:

\[ \min_w \quad T \times R^2(w), \]  
\[ z_t^2(w) = (x_t - y_t w)^2, \]  
\[ R^2 = z_t^2(w)' \Phi_{t-1} (\Phi_{t-1}' \Phi_{t-1})^{-1} \Phi_{t-1}' z_t^2(w) / \sigma^2 \]

where \( x_t \) is a \((T \times 1)\) vector of portfolio returns, \( y_t \) is a \((T \times K)\) matrix or factor realizations (or the returns on factor-mimicking portfolios), \( w \) is the \((K \times 1)\) vector of cofeature parameters (or portfolio weights), \( \sigma^2 \) is a consistent estimator of the variance of \( z_t^2(w) \), and \( \Phi_{t-1} \) is the information set composed of \( x_{t-1}^2, \ldots, x_{t-p}^2, y_{t-1}^2, \ldots, y_{t-p}^2 \), and lagged cross-products \( x_{t-1} y_{t-1}, \ldots, x_{t-p} y_{t-p} \). This
is a general method of moments type test, the objective of which is to identify a linear combination of \( x_t \) and \( y_t \) whose squared returns are uncorrelated with any information in \( \Phi_{t-1} \). That is, the objective is to find a portfolio of \( x_t \) and the \( K \) assets in \( y_t \) that exhibits no ARCH, providing evidence that a common volatility process exists between them.

Engle and Kozicki (1993) show that, under the assumptions of Hansen (1982), \( \min_w T \times R^2(w) \) follows a \( \chi^2 \) distribution with degrees of freedom given by the number of overidentifying restrictions (i.e., number of instruments included in \( \Phi_{t-1} \) minus the number of variables used plus 1). Thus, if \( p \) is the number of lags used in the information set and \( K \) is the number of factors in \( y_t \), then \( T \times R^2 \) is distributed as \( \chi^2_{\left(\frac{(K+2)(K+1)}{2} p - K\right)} \). We refer to this common ARCH test as COARCH(p).

Specifically, we test whether the set of within sample principal components span the second moments of the trading strategies SMB, HML, IM, and PR1YR. If they do not, then these strategies may follow their own volatility process, not explained by common sources of risk.\(^\text{18}\) Thus, a rejection of the null hypothesis of no ARCH is evidence in favor of these strategies having their own volatility process, independent from the factor proxies. Table 2.2 reports the common ARCH tests using both 1 and 6 lags in the information set (i.e., COARCH(1) and COARCH(6)). We also include common ARCH tests for the excess return on the market as a benchmark.

Using 1 lag in the information set, the tests fail to reject the null of no ARCH at the 5% significance level for the market, SMB, and HML, indicating that these portfolio return second moments are spanned by the proxies for systematic

\(^{18}\)Since the within sample principal components capture over 95% of the cross-sectional variation in returns of our 32 assets at a point in time, we employ them as proxies for the true systematic factors in the economy.
risk. The null is rejected for $IM$ and $PR1YR$, however, suggesting that these strategies are not fully captured by systematic risk. Rejection of the null may be due to our use of only 1 lag in the information set, however. When we run the tests employing 6 lags in the information set, we fail to reject the null for $IM$ as well. However, $PR1YR$ still appears to follow a volatility process significantly different from the principal components factors. Thus, the market, $SMB$, and $HML$ seem to follow the same volatility process as the proxies for systematic risk, while the momentum strategies appear to have their own variability, particularly the individual stock momentum strategy.\(^\text{19}\)

2.4 How Risky are Size, BE/ME, and Momentum Strategies?

Thus far, the evidence for risk-based explanations appears relatively weak, particularly for momentum. In this section, therefore, we analyze the risk of these anomalies from a different perspective. Specifically, we want to determine how risky strategies designed to exploit these anomalies are. If size, $BE/ME$, and momentum are related to risk, then we should not be able to create strategies that have high size, $BE/ME$, and momentum characteristics or exposure without incurring substantial variability. Furthermore, when we hedge out exposure to other characteristics or factors, such strategies should remain volatile if these anomalies proxy for risk. In this section, we design portfolios that are highly exposed to one of the anomalies, while having no exposure to the others and to other proxies for risk. Such analysis will determine how risky these investment strategies truly are. If the strategies are highly volatile, then we conclude that

\(^{19}\)This is consistent with Moskowitz (1997) who finds that an industry momentum strategy is more efficient than an individual stock momentum strategy.
the anomaly is likely correlated with a source of (non-diversifiable) systematic variation. If, however, the Sharpe ratio of the strategy is very high, then we conclude that the anomaly is unlikely due to risk, and have in the process identified a profitable trading strategy to exploit it.

Using our GARCH generated conditional covariance matrices for all 32 assets described in Section 2.3, we wish to create portfolios that minimize variability subject to having large exposure to each of size, BE/ME, and momentum. Our first strategy is to track the market portfolio by minimizing the squared deviation from the market (tracking error) subject to having a high exposure to one of the ‘anomalous’ characteristics. For ease of presentation, we will illustrate our strategies for momentum, but also conduct the same analysis for size and BE/ME. The first strategy minimizes market tracking error subject to having a past 1-month return (momentum characteristic) one standard deviation above the cross-sectional average past return of our 32 assets.

\[
\min_w \left[ w' R_t - R_{mt} \right]^2
\]

s.t. \( w' R_{t-1} = \mu_{t-1} + \sigma_{t-1} \)

where \( w \) is a \((32 \times 1)\) vector of portfolio weights. \( R_{t-1} \) is the vector of prior month excess returns, and \( \mu_{t-1} \) and \( \sigma_{t-1} \) are the cross-sectional mean and standard deviation of past 1-month returns.

In addition, we minimize variance subject to having a high momentum attribute and hedging out exposure to other factors. Specifically, the other four strategies are:

\[
\min_w w' H_t w
\]

s.t. \( w' R_{t-1} = \mu_{t-1} + \sigma_{t-1} \)

\( w' \beta_t = 0 \)
\[
\begin{align*}
\begin{align*}
w'\beta_{\text{smb},t} &= 0 & (2.42) \\
w'\beta_{\text{hml},t} &= 0 & (2.43) \\
(3) \quad \min & \quad w'H_tw \\
\text{s.t.} \quad w'\beta_{\text{IM},t} &= \bar{\beta}_t + \sigma_{\beta,t} & (2.45) \\
w'\beta_t &= 0 & (2.46) \\
w'\beta_{\text{smb},t} &= 0 & (2.47) \\
w'\beta_{\text{hml},t} &= 0 & (2.48) \\
(4) \quad \min & \quad w'H_tw \\
\text{s.t.} \quad w'R_{t-1} &= \mu_{t-1} + \sigma_{t-1} & (2.50) \\
w'\beta_{\text{PC},t} &= 0 & (2.51) \\
(5) \quad \min & \quad w'H_tw \\
\text{s.t.} \quad w'\beta_{\text{IM},t} &= \bar{\beta}_t + \sigma_{\beta,t} & (2.53) \\
w'\beta_{\text{PC},t} &= 0 & (2.54)
\end{align*}
\end{align*}
\]

where \( \beta_t, \beta_{\text{smb},t}, \beta_{\text{hml},t}, \beta_{\text{IM},t}, \) and \( \beta_{\text{PC},t} \) are the market, \( \text{SMB}, \text{HML}, \text{IM}, \) and \( \text{PC} \) factor loadings, respectively, and where we allow the strategies to have both a high momentum characteristic as well as a high momentum factor loading.

In other words, given that we load high on momentum, what is the smallest risk we can expect these portfolios to have? If momentum is related to systematic risk, then the risk of these portfolios will be substantial, if not, then these portfolios will exhibit little variation. To measure the improvement in efficiency of these portfolios, we examine their risk/return tradeoff via the Sharpe ratio. Since we employ excess returns in the optimization, we do not need the constraint that the weights, \( w \), sum to one. All we are interested in is obtaining the Sharpe measure, giving us the slope of the capital market line, which is invariant to scale. The
weights, \( w \), are computed at time \( t \), and applied out of sample, to time \( t+1 \) returns and the time \( t+1 \) covariance matrix, \( H_{t+1} \). Hence, the Sharpe ratio at time \( t+1 \) is computed as:

\[
\text{Sharpe}_{t+1} \equiv w_t' R_{t+1} (w_t' H_{t+1} w_t)^{-1/2}
\]  

(2.55)

The time-series average of these Sharpe ratios are reported in Table 2.3. In addition, we re-calculate these strategies for size and BE/ME. As the table shows, the Sharpe ratios on momentum are quite high, suggesting that momentum is unlikely driven by systematic variation. In fact, loading on the characteristic subject to hedging out the Fama and French factors produces higher Sharpe ratios than loading on the momentum factor \( IM \). However, the reverse is true when hedging out the principal components factors. The evidence on size and BE/ME is much more supportive of risk stories, since the Sharpe ratios associated with these strategies are much lower, particularly for BE/ME. Again, the evidence on factor loadings versus characteristics is ambiguous.

Earlier we claimed that 0.031 was a reasonable bound on the maximum squared Sharpe ratio in the economy. This translates into a 0.176 monthly Sharpe ratio. Examining the standard errors of the time-series of Sharpe ratios for our strategies, only momentum is not within a 95% confidence interval of 0.176. Thus, both size and BE/ME appear less profitable and more risky than momentum, as the Sharpe ratios on both size and BE/ME strategies seem reasonably small to be justified as being related to risk, while momentum strategies appear to violate what we might consider a reasonable bound on the maximum Sharpe ratio in the economy. Again, though, the dynamic nature of a momentum strategy suggests that the Sharpe ratio may be an improper measure of its performance.
2.5 Forecasting Expected Returns and Risk

Irrespective of whether these anomalies can reasonably be explained by risk, many researchers have employed trading strategies based on these anomalies as proxies for systematic risk factors in the economy. The Fama and French (1993) three factor model, for instance, employs $SMB$ and $HML$ as factor-mimicking portfolios in an APT-type setting or as proxies for state variables that represent the changing investment opportunity set as in Merton (1973). Likewise, both Carhart (1997) and Moskowitz (1997) add momentum 'factors' to the FF model. Evidence so far suggests that size and BE/ME can reasonably be proxies for risk, but that momentum seems unlikely to be risk-based. when investors only care about the first two moments of returns. Although the premium on momentum appears too high to be explained by its volatility, a momentum factor may still be useful for predicting future expected returns and covariances. In this section, we test how well these 'factor-mimicking' portfolios predict future expected returns and covariances. Specifically, we examine the performance of the various factor models specified in Section 2.3.1 to predict expected returns and risk out of sample.

Numerous studies have evaluated the explanatory power of various asset pricing models for predicting the cross-section of expected returns, but none have documented the explanatory power for predicting risk. More precisely, we wish to determine in this section, how much of the true covariance matrix of asset returns each model captures. To be a successful asset pricing model, the factors must both describe expected returns and account for a substantial portion of the covariance matrix. In this section, we test our various factor models on these two accounts. We start with the cross-section of expected returns.
2.5.1 Non-Nested Cross-Sectional Regressions

If a pre-specified factor model explains returns, then asset loadings on these factors should be positive predictors of expected returns. Therefore, we can test the explanatory power of a model by regressing the cross-section of returns at time $t$ on the estimated factor loadings using information up to time $t - 1$.

$$ R_t - r_{ft} = a_t + \hat{\beta}_{t-1} b_t + e_t $$

(2.56)

where $\hat{\beta}_{t-1}$ is the $(N \times K)$ matrix of estimated factor loadings of the $N$ assets formed in the prior period, and $b_t$ is the $(K \times 1)$ vector of 'rewards' (risk premia) for factor exposure. If the model is true, then $a_t$ should be zero and the elements of $b_t$ should be positive. Since we want the factor loadings to predict expected or average returns, we form a time-series of the cross-sectional regression coefficients in the style of Fama and MacBeth (1973) and test whether $a = 0$ and the $b$'s $\neq 0$. However, Fama-MacBeth regressions do not allow us to compare the relative performance of the various factor models. That is, tests of $a = 0$ and $b \neq 0$ are only informative about whether a particular model can explain average returns, but they tell us nothing about the fit of one model versus another. Therefore, in order to determine which factor model best predicts future expected returns, we employ the $J$-test of Davidson and MacKinnon (1981) using a non-nested regression approach in combination with the Fama-MacBeth cross-sectional regression technology. Specifically, we wish to test which of two possible sets of regressors is better at capturing the cross-section of expected returns; out of sample. For instance, suppose we have two competing factor models which imply the following expected return equations:

$$ H_0 : E[R_t] = r_{ft} + \beta_t \lambda_t $$

(2.57)

$$ H_1 : E[R_t] = r_{ft} + \Gamma_t \gamma_t $$

(2.58)
where $\beta_t$ is the $(N \times K_1)$ matrix of factor loadings and $\lambda_t$ the $(K_1 \times 1)$ vector of factor risk premia at time $t$ for the first model. Similarly, $\Gamma_t$ is the $(N \times K_2)$ matrix of factor loadings and $\gamma_t$ the $(K_2 \times 1)$ vector of factor risk premia at time $t$ for the second factor model, where $K_2$ does not necessarily equal $K_1$ (i.e., the number of factors of each model may differ).

These expected return equations can be translated into Fama-MacBeth cross-sectional regression equations,

$$H_0: R_t - r_{ft} = a_t + \hat{\beta}_{t-1}b_t + e_t \quad (2.59)$$

$$H_1: R_t - r_{ft} = c_t + \hat{\Gamma}_{t-1}g_t + u_t \quad (2.60)$$

where $a_t$ and $c_t$ are constants, $b_t$ and $g_t$ are rewards for various factor exposure, $e_t$ and $u_t$ are error terms, and $\hat{\beta}_{t-1}$ and $\hat{\Gamma}_{t-1}$ are estimates of factor loadings formed in the prior period (which will be described shortly). Since $H_1$ cannot necessarily be written as a restriction on $H_0$, we treat the two models as non-nested and employ the Davidson and MacKinnon (1981) $J$-test on these competing models. Specifically, this test computes the fitted values from the $H_1$ regression, and then adds them to the $H_0$ regression as an additional explanatory variable. That is, we test the following model.

$$R_t - r_{ft} = a_t + \hat{\beta}_{t-1}b_t + \psi(\hat{c}_t + \hat{\Gamma}_{t-1}\hat{g}_t) + e_t \quad (2.61)$$

where $\hat{c}_t$ and $\hat{g}_t$ are the fitted values from the $H_1$ regression. An asymptotically valid test of $H_0$ is a test that $\psi = 0$, where $\psi$ is distributed as standard normal. If $\psi = 0$, then we favor $H_0$ in lieu of $H_1$. If $\psi = 1$, then $H_1$ is preferred over $H_0$ as a model of returns. Running these non-nested cross-sectional regression tests at each point in time produces a time-series of $\psi$'s, which we compute a

$^{20}$Unfortunately, in testing $H_0$ versus $H_1$ and vice versa, all four possibilities: reject both, neither, or either one of the two models, could occur. This, however, is only a finite sample problem.
mean and test statistic (z-stat) for. This procedure is repeated for various pairs of competing factor models.

2.5.1.1 Portfolios

We first examine the return predictability of the various competing factor models outlined in Section 2.3.1. by applying the non-nested cross-sectional regression tests on the 32 portfolios described in Section 2.3. At time \(t\), we compute the factor loadings via linear regression using the Ledoit and Santa-Clara (1998) estimated covariance matrix. \(H_t\), as described in Section 2.3.1. and test whether these loadings forecast future average returns. For example, suppose the null \((H_0)\) is the FF model, and the alternative \((H_1)\) the four factor industry momentum model. At every time \(t\), we compute \(\hat{\beta}_t\) as in equation (2.12) and \(\hat{\Gamma}_t\) as in equation (2.18) (i.e., \(\hat{\Gamma}_t = \beta_t^{Mo}\)). We then regress time \(t + 1\) excess returns on a constant and \(\hat{\Gamma}_t\), and place the fitted values in equation (2.61). From this, \(\psi\) is computed.

We employ both ordinary least squares (OLS) and generalized least squares (GLS) estimates, where

\[
\hat{\gamma}_{ols,t} = (\hat{\Gamma}'_{t-1} \hat{\Gamma}_{t-1})^{-1} \hat{\Gamma}'_{t-1} [R_t - r_{ft,t}]
\]

\[
\hat{\gamma}_{gls,t} = (\hat{\Gamma}'_{t-1} H_t^{-1} \hat{\Gamma}_{t-1})^{-1} \hat{\Gamma}'_{t-1} H_t^{-1} [R_t - r_{ft,t}]
\]

\[
\omega_{ols,t} = \lambda (X'X)^{-1} X' [R_t - r_{ft,t}]
\]

\[
\omega_{gls,t} = (X'H_t^{-1} X)^{-1} X'H_t^{-1} [R_t - r_{ft,t}]
\]

\[
X \equiv [1 : \hat{\beta}_{t-1} : (\hat{c}_t + \hat{\Gamma}_{t-1} \hat{\gamma}_{ols,t})]
\]

\[
X \equiv [1 : \hat{\beta}_{t-1} : (\hat{c}_t + \hat{\Gamma}_{t-1} \hat{\gamma}_{gls,t})]
\]

(2.62)

The time-series average of \(\psi\) is reported in Table 2.4. where it is apparent that the industry momentum model is preferred to the FF model, based on future return predictability. This process is repeated for each of the other various factor models relative to the industry momentum model. In addition, we reverse the order of the cross-sectional regressions by using the loadings from the industry momentum model for \(\hat{\beta}\) and using the fitted values from the other factor models.
in (2.61). As Table 2.4 demonstrates, the industry momentum model appears to outperform the FF and Carhart factor models, and performs as well as the principal components.\textsuperscript{21} and unconditional principal components factor models. This suggests that an industry momentum factor is an important predictor of future expected returns.

2.5.1.2 Individual Equities

Although an \( IM \) factor seems to explain the expected returns of our 32 portfolios, this result may not be too surprising since 20 of the 32 portfolios are industry portfolios from which the \( IM \) factor is derived. Therefore, we employ the same non-nested cross-sectional regression procedure on the cross-section of individual stock returns at each point in time. Since the number of individual equities ranges from approximately 1,000 to more than 7,000 at any point in time, we can no longer feasibly employ a multivariate GARCH model to estimate conditional covariance matrices, from which to form factor loadings. Therefore, we estimate factor loadings for each individual security by regressing the prior 36 months of returns of the stock (from \( t - 36 \) to \( t - 1 \)) on the past 36 months of factor realizations.\textsuperscript{22} These loadings are then employed in (2.61), where the OLS estimates are calculated as before, but the GLS estimates are calculated using a feasible generalized least squares (FGLS) technique that computes the weighted least squares estimates of \( g \) and then of \( \psi \), where the weights are the fitted values from the OLS regression.\textsuperscript{23} The results are reported in Table 2.4, where the

\textsuperscript{21}Here, of course, the principal components factors are formed out of sample, since the within sample principal components will ‘overfit’ the cross-sectional dispersion in average returns. That is, the principal components factor portfolio weights (i.e., scaled eigenvectors) are formed in the prior period (from \( H_{t-1} \)) and applied to current returns, \( R_t \).

\textsuperscript{22}Securities with less than 24 months of prior available return data are excluded.

\textsuperscript{23}This procedure assumes that the variance of the dependent variable is proportional to the square of the mean. See Greene (1993) and references therein for further details.
industry momentum model still appears to dominate the other factor models, confirming that an \(IM\) factor is an important component for future expected return predictability.

### 2.5.2 Predicting Future Volatility

So far, we have shown that an industry momentum factor explains future expected returns quite well, and that sensitivity to this factor is associated with higher average returns. However, if the \(IM\) factor is a proxy for systematic risk, then it should aid in predicting future covariances as well. More generally, assessing the performance of our competing factor models from Section 2.3.1 entails evaluating the out of sample predictability of future covariances generated from these competing models.

#### 2.5.2.1 Portfolios

We begin simply by examining how much each factor model contributes to conditional volatility, as estimated by our Ledoit and Santa-Clara (1998) GARCH(1,1) covariance matrices. Here, we estimate the \(H_t\) covariance matrices using the sample mean up to time \(t-1\) as an approximation for the mean of the asset returns, following equations (2.4) and (2.8) as before. In other words, we are not incorporating any information about mean returns from the factor models. We can decompose the \(H_t\) matrices into a systematic and non-systematic (or residual) component under each factor model. For instance, under the Fama and French three factor model,

\[
H_t = \beta_t \Sigma_t \beta_t' + \Omega_t \tag{2.63}
\]

where \(\Sigma_t\) is the covariance matrix of the FF factors, and \(\Omega_t\) is the residual covariance matrix at time \(t\) implied by the FF model. We then test how closely the
systematic portion tracks the true covariance matrix: out of sample.

We define the ‘true’ covariance matrix at time $t$ as $V_t$, and use the product moment matrix $R_t R'_t$ as a proxy for $V_t$. We then estimate the covariance matrix implied by the FF model as,

$$\hat{V}_t = \hat{\beta}_t \hat{\Sigma}_t \hat{\beta}'_t$$

(2.64)

where $\hat{\beta}$ and $\hat{\Sigma}$ are extracted from $H_t$ (which only employs realizations up to time $t - 1$). We then measure how far away our various covariance matrix estimates are from $V_t$, by employing several measures of tracking error. We measure the distance from how far $\hat{V}_t$ is from $V_t$, the ‘true’ covariance matrix, by employing the following three metrics commonly used in the statistical literature:\(^{24}\)

$$D_{1t} \equiv \text{trace}\{(V_t - \hat{V}_t)^2\}$$

(2.65)

$$D_{2t} \equiv \text{trace}\{||(V_t - \hat{V}_t)V_t^\dagger||\}$$

(2.66)

$$D_{3t} \equiv \text{trace}\{\hat{V}_t V_t^\dagger\} - \log(\det(\hat{V}_t V_t^\dagger)) - \text{rows}(V_t)$$

(2.67)

where $\dagger$ represents the generalized inverse of a matrix.\(^ {25}\)

The first distance metric, $D_1$, measures the aggregate squared deviation from $V$. If $\hat{V}$ tracks $V$ closely, then $D_1$ will be close to zero. In other words, think of $V - \hat{V}$ as the matrix of covariance tracking errors. Since the trace of a matrix equals the sum of its eigenvalues, we can view $D_1$ as the sum of the eigenvalues of the matrix of squared tracking errors. In other words, $D_1$ represents the sum of squared tracking errors if we were to transform the matrix of squared tracking errors into $N$ orthogonal components. $D_2$ has a similar interpretation except that each deviation from $V$ is now weighted by the inverse of each element of $V$, giving more weight (penalty) to those elements of $\hat{V}$ that are further away from the corresponding element in $V$. The eigenvalues of $(V - \hat{V})V^\dagger$ are termed the

\(^{24}\)See Muirhead (1982) and references therein.

\(^{25}\)The generalized inverse is employed since $R_t R'_t$ is not guaranteed to be positive definite.
"generalized eigenvalues of $V - \hat{V}$ with respect to $V$." If $\hat{V}$ tracks $V$ closely (i.e., if the model is good at predicting future covariances), then $D2$ will be close to zero. Finally, $D3$ represents another version of a distance metric which applies less extreme weights to elements of $\hat{V}$ which deviate significantly from $V$.

Table 2.5 reports the time-series means of the three distance metrics for each of the five factor models relative to the product moment matrix. $V_t = R_t R_t'$. The magnitudes of the metrics provide a ranking for how well each of the models predicts future covariances. As the table indicates, the industry momentum model appears to be a strong predictor of future covariances, consistent with our earlier finding that this model also explains the cross-section of average stock returns well. Thus, an $IM$ factor seems to aid in forecasting future volatility, and thus may provide a good proxy for risk. The FF and unconditional principal components models also perform quite well, while the out of sample principal components model are not very accurate at forecasting risk.

In addition, we employ single factor models that analyze the covariance contribution of each of our 'anomaly-generated' factor portfolios: $IM$, $SMB$, $HML$, and $PR1YR$. We also include the excess return on the market for comparison (i.e., the CAPM). As Table 2.5 indicates, $IM$ has stronger predictive power than any other individual factor for predicting risk.

Finally, we incorporate information about the mean returns of our assets implied by the factor models, by computing conditional covariance matrices for the 32 portfolios, using the Ledoit and Santa-Clara (1998) algorithm, under each of the five competing factor models. Specifically, we compute a time-series of conditional covariance matrices for each factor model in Section 2.3.1 using equation (2.3), where the residuals, $\epsilon_t$, are defined according to the return generating model. For example, the FF-implied covariance estimates are computed using
equation (2.3) in combination with equation (2.12), where the parameters $A$, $B$, and $C$ are first estimated and then used to rebuild a time-series of matrices. This procedure is then conducted for the Carhart model, using equation (2.3) with equation (2.15), where, again, $A$, $B$, and $C$ are re-estimated, and so on for each of the various factor models in Section 2.3.1. Thus, five sets of time-varying covariance matrices are obtained (one for each factor model), which can be compared to the ‘true’ covariance matrix at each point in time, to determine which model best fits the covariance structure of asset returns. We also compute covariance matrices under each of the single factor models in the same manner. The results are reported in Table 2.5. Again, an $IM$ factor appears to capture future volatility and covolatility quite well, although all of the factors now perform relatively well.

2.5.2.2 Individual Equities

For robustness, we conduct the same ‘horse race’ using individual stocks as our basis assets. However, the GARCH technology is infeasible for such a large number of assets. Therefore, using the prior 36 month regressions described in Section 2.5.1.2, we estimate conditional covariance matrices using the estimated coefficients (i.e., factor loadings) and residuals from the regression. For example, for the FF model, we regress the monthly returns of stock $i$ on a constant plus the excess return of the market, $SMB$, and $HML$ over the past 36 months (from $t - 36$ to $t - 1$). The coefficients from this regression are defined as the factor loadings of stock $i$ at time $t - 1$, $\hat{\beta}_{i,t-1}$. Doing this for all $N$ stocks ($\forall i \in N$), produces an $(N \times 3)$ loading matrix $\hat{\beta}_{t-1}$. We then estimate the covariance matrix implied by the FF model as,

$$\hat{\Sigma} = \hat{\beta}_{t-1} \hat{\Sigma}_{t-1} \hat{\beta}'_{t-1}$$  \hspace{1cm} (2.68)
where \( \hat{\Sigma}_{t-1} \) is the estimated sample covariance matrix of the three FF factor-mimicking portfolios over the past 36 months. This procedure is repeated for all of the competing factor models, where \( \hat{\beta}_{t-1} \) becomes an \((N \times K)\) matrix of loadings and \( \hat{\Sigma}_{t-1} \) is the \((K \times K)\) sample covariance matrix of the factors over the last 36 months. Again, we define the ‘true’ covariance matrix at time \( t \) as the product moment matrix \( (V_t = R_t R'_t) \), and compute our distance measures, \( D1, D2, \) and \( D3 \) relative to \( V_t \).

However, because of the large number of individual stocks, computing these distance measures is problematic. Therefore, we select 50 stocks randomly at time \( t \) and compute the distance metrics for each model. We then resample another 50 individual securities (with replacement) and compute the metrics again, repeating this process 50 times. Thus, at each point in time we compute 50 sets of distance metrics for each factor model, using 50 sets of 50 randomly selected securities at time \( t \). The average values of \( D1, D2, \) and \( D3 \) across the 50 simulations are calculated at each point in time, averaged over time, and reported in Table 2.5. As the table shows, an industry momentum factor still predicts future covariances quite well, even for individual equities. Finally, the market has strong predictive power for individual equity future volatility, providing strong evidence of a market factor, and possibly renewed support for the CAPM.

2.5.3 Residual Volatility

As a final test for the models, rather than examining the systematic portion of the covariance matrix implied by each factor model, we examine the residual covariance matrix implied by each model. In other words, once we hedge out the factors associated with an asset pricing model, we wish to see what remains. In particular, we want to determine how much volatility remains once we account
for the factors under the various models. Using equation (2.63), we compute the residual covariance matrix, $\Omega_t$, under each factor model at a point in time. We then form minimum variance portfolio weights using $\Omega_t$ and apply these weights to residual returns out of sample, where residual returns are also defined according to the factor model. Specifically, the residual return on this portfolio is defined as,

$$
\epsilon_t = \Omega_t^{-1} \Omega_t^{1/2} [R_t - \hat{\beta}_t F_t]
$$

(2.69)

where $\iota$ is a $(32 \times 1)$ vector of ones, $R_t$ is the vector of asset returns, $\hat{\beta}_t$ is the $(32 \times K)$ matrix of factor loadings, estimated from the conditional covariance matrix, $H_t$, and $F_t$ is the $(K \times 1)$ vector of factor realizations. Again, this is an out of sample test. since both $\Omega_t$ and $\hat{\beta}_t$ are forecasts based on observations through time $t - 1$. If the factor model captures a substantial portion of the covariance matrix of asset returns, then the residual return of the minimum variance portfolio formed from the residual covariance matrix implied by the model, $\epsilon_t$, should be close to zero. We compute these residuals for each factor model and report their sums of squares, as well as their sum of squared deviations from their mean in Table 2.6. As the table demonstrates, the principal components models capture the most variation, and the CAPM does surprisingly well at explaining the covariance matrix. For robustness, we also compute portfolio weights using only the diagonal elements of $\Omega_t$ (e.g., just the variances), and equal-weighting the residual returns. For the diagonal/variance-only portfolio weights, the previous results are confirmed: the principal components models and the CAPM capture the most significant amount of variation. With equal portfolio weights, the CAPM is less successful and the FF model seems to perform quite well.

Finally, Panel B of Table 2.6 reports the means and t-stats of these residual portfolios. As shown in the table, all of the models seem to capture most of the
variation in asset returns, as the residual returns on the minimum variance portfolios are statistically indistinguishable from zero. particularly for the principal components models. Examine the single factor models, it is again apparent that the market captures substantial variation across assets, as the minimum variance portfolio produces zero average residual profits. Both \textit{SMB} and \textit{HML} also contribute significantly to covariance risk, while the momentum portfolios contribute the least. Thus, although a momentum factor is useful for forecasting future risk, its contribution to total volatility is quite small. Thus, the variability and covariability of a momentum strategy appears too small to justify the premium for momentum, while the contribution by \textit{SMB} and \textit{HML} to the covariance matrix appears substantial, suggesting that the premium on size and \textit{BE/ME} may be compensation for risk.

\subsection*{2.6 Conclusion}

Overall, our findings on the risk of size and \textit{BE/ME} are promising for rational asset pricing theorists, as size and \textit{BE/ME} 'factors' appear to be spanned by sources of systematic variation, appear to have substantial variation even when controlling for other possible sources of variability, have Sharpe ratios that appear within the bounds of 'reasonable' magnitudes, and are significant in describing the cross-section of expected returns as well as the covariance matrix of returns, and comprise a significant portion of the covariance matrix. The evidence for a risk-based explanation for momentum is much weaker, despite the fact that an industry momentum factor forecasts expected returns and covariances very well. Momentum factor portfolios do not appear to be spanned by systematic risk, have Sharpe ratios that appear to violate 'reasonable' bounds for rational asset pricing, and comprise a smaller portion of the covariance matrix than either
size or BE/ME. Thus, size and BE/ME seem more likely to proxy for risk, while momentum seems to be driven by other phenomena, perhaps related to investor behavior. There is a possible rational alternative explanation for momentum discussed in Appendix 2.8, which also seems to fail.

These conclusions are, of course, based on our definition of risk as the second moments of returns. When higher moments are taken into account, all these anomalies seem to be fairly priced, as momentum exhibits negative skewness and high kurtosis, making such strategies seemingly more risky. Adjusting for co-skewness and co-kurtosis reduces the profitability of momentum strategies immensely, suggesting that an asset pricing model containing higher moments may be useful.

Finally, our tests also conclude that the market portfolio is significant in explaining future risk and comprises a significant portion of the covariance matrix of asset returns, providing new support for the CAPM.
2.7 Appendix A: Portfolio Construction

SMB and HML are constructed in the same manner as Fama and French (1993). After ranking all NYSE firms by their BE/ME\textsuperscript{26} ratios at the end of year \( t - 1 \) and their market capitalization at the end of June of year \( t \), we form 30 percent and 70 percent breakpoints for BE/ME and a 50 percent breakpoint for market capitalization. Beginning in July of year \( t \),\textsuperscript{27} all NYSE, AMEX, and NASDAQ stocks are placed into the three book-to-market groups (High, Medium, and Low) and two size groups (Small and Big) based on the breakpoints. Six value-weighted portfolios are then formed based on the intersection between the size and book-to-market groups. These six portfolios are used in the GARCH estimation to represent the size and BE/ME factors in the economy. The HML returns are defined as: \((r_{HB} + r_{LS} - r_{LB} - r_{LS})/2\), and the SMB returns are defined as: \((r_{HS} + r_{MS} + r_{LS} - r_{HB} - r_{MB} - r_{LB})/3\). Only firms that have been listed on COMPUSTAT for at least two years prior to portfolio formation, and have prices on CRSP in December of year \( t - 1 \) and June of year \( t \) are included.

PR1YR is constructed similar to Carhart (1997) by forming a value-weighted portfolio of firms with the highest 20 percent prior 11-month returns, lagged one month (i.e. from \( t - 12 \) to \( t - 2 \)), and subtracting from this a value-weighted portfolio of firms with the lowest 20 percent prior 11-month returns.\textsuperscript{28} The resulting zero-cost portfolio is re-formed monthly. In addition, value-weighted quintiles

\textsuperscript{26}Book value of equity is defined as stockholder's equity plus deferred taxes and investment tax credits. These numbers are obtained from COMPUSTAT. BE/ME is calculated by dividing the most recent book value in year \( t - 1 \) by the number of shares outstanding times price at the end of year \( t - 1 \), obtained from CRSP. All firms with BE/ME values less than zero are excluded.

\textsuperscript{27}As in Fama and French (1993), the end of June is used as the portfolio formation date to ensure that book values of equity are publicly available from annual reports at the time portfolios are formed.

\textsuperscript{28}Carhart (1997) forms equal-weighted portfolios and uses 30% and 70% breakpoints to form his factor.
based on past return rankings each month are included in the GARCH estimation, and \( PR1YR \) is defined as the highest quintile return minus the lowest quintile return.

Using the CRSP and COMPUSTAT data files, 20 value-weighted industry portfolios are formed every month from July, 1963 to July, 1995. Two digit Standard Industrial Classification (SIC) codes are used to form industry portfolios in order to maximize coverage of NYSE, AMEX, and NASDAQ stocks, while maintaining a manageable number of industries, and insuring each industry comprises a fairly significant portion of the economy. The average number of stocks per industry is 230, and the fewest number of stocks in any industry (except for Railroads) at any time is over 25. Therefore, virtually all portfolios are well-diversified (aside from industry effects). Summary statistics for these industries can be found in Table 1 of Moskowitz (1997).

\( IM \) is constructed in the same manner as the IMS2(1,1) zero-cost portfolio from Moskowitz (1997), by forming an equally-weighted portfolio of the three industries which exhibited the largest returns in the prior month, and subtracting from this the equal-weighted return of the three industries with the lowest prior month performance.

2.8 Appendix B: Alternative Hypothesis for Momentum: Conditional Industry Risk Premia

The fact that industry components seem to drive momentum, as evidenced by Moskowitz (1997), may be consistent with both behavioral and rational theories of valuation. Under behavioral theories, investors may persistently and irrationally underreact to industry information, thereby providing rational investors with a
profit opportunity. Under rational theories, time-varying industry risk premia may exist, and the persistence in stock return performance may be driven by the stock's covariation with these economy-wide factors. Thus, apparent profits from momentum investment strategies may be compensation for risk. The question, therefore, is whether industry factors represent compensation for risk, or whether another apparent abnormality in capital markets exists. At issue is whether industry effects have systematic influence on asset prices.

Consider the following multifactor linear model of returns, which may be interpreted as a specification of the Arbitrage Pricing Theory (APT),

$$\bar{r}_{jt} = E[\bar{r}_{jt}] + \sum_{k=1}^{K} \beta_{jk,t} \bar{F}_{kt} + \sum_{m=1}^{M} \theta_{jm,t} \bar{\delta}_{mt} + \bar{\epsilon}_{jt}$$

(2.70)

where $\bar{r}_{jt}$ is the return of stock $j$ at time $t$, $\bar{F}_{kt}$ are unconditionally priced factors or the returns on factor-mimicking portfolios, $\beta_{jk,t}$ are the factor sensitivities, $\bar{\delta}_{mt}$ are correlated components of returns across assets or 'approximate factors' which we can think of as industry components, $\theta_{jm,t}$ are the 'approximate factor' sensitivities, and $\bar{\epsilon}_{jt}$ are idiosyncratic components which are uncorrelated across assets. At this point, we make no claims as to whether $\bar{\delta}$ risk is diversifiable. That is, we do not know whether the $\bar{\delta}$'s carry a risk premium or whether they are a component of diversifiable risk.

Unconditionally, industries do not appear to exhibit risk premia, but conditionally, (i.e., at certain times) some industries may have risk premia associated with them. At other times, these industries may not have risk premia, and other industries might. If these risk premia average out to zero over time, then each industry will appear to have an insignificant influence on prices, even though it may have substantial pricing impact at certain times. We can therefore express

---

See Fama and French (1997) and Moskowitz (1997) for evidence supporting this claim.
expected returns as,

$$E[ar{r}_{jt}] = r_{ft} + \sum_{k=1}^{K} \beta_{jk,t} \lambda_{kt} + \sum_{m=1}^{M} \theta_{jm,t} \eta_{mt}$$  \hspace{1cm} (2.71)$$

where $\lambda_{kt}$ are the risk premia on the $K$ unconditional factors, and $\eta_{mt}$ are the risk premia on the $M$ industry factors at time $t$. Under the behavioral hypothesis, industries are diversifiable, and $\sum_{m=1}^{M} \eta_{mt} = 0, \forall t$. But, under the rational hypothesis, $\sum_{m=1}^{M} \eta_{mt} \neq 0$. Since, at any point in time, only certain industries will carry a premium, we can effectively capture the conditional influence of industries by employing an industry momentum factor in place of $\sum_{m=1}^{M} \eta_{mt}$. In other words, if conditional industry risk premia are driving momentum, then an industry momentum portfolio will provide a factor-mimicking approximation for the conditional pricing influence of these industries.

If momentum profits are compensation for risk, then high past return industries must be systematically more risky on average, and therefore will command a premium, and low past return industries must be less risky on average, and therefore will exhibit lower expected returns. In this section, we test the validity of this argument.

If momentum is related to risk, then the conditional volatility of industries experiencing high past returns should be higher than those with low past returns, to be consistent with a risk-based explanation for the higher average returns of high past return industries in the subsequent period. That is, since past performance is associated with future performance, past performance must also be associated with future risk (volatility), if a rational pricing theory explains momentum. Using the Ledoit and Santa-Clara (1998) GARCH estimation to form a series of conditional covariance matrices of our 32 portfolios, we test whether high past return industries are riskier than low past return industries.
However, it is not only the volatility of the industries which matters, but also how other assets in the economy covary with the industry. In other words, the conditional ‘covolatility’ of the industry is the important measure of risk, which should be positively related to past performance in order to have a rational explanation for momentum. Using the Ledoit and Santa-Clara (1998) GARCH-estimated covariance matrices, we conduct the following experiment. We first rank our 20 industries based on their prior 1-month returns, skipping a month (time \( t - 2 \)).\(^{30}\) Then, taking the conditional covariance matrix at time \( t \), \( H_t \), of our 20 industry returns, we delete the row and column corresponding to the industry with the highest past 12-month return. The resulting \((31 \times 31)\) reduced covariance matrix, \( H_t^R \), represents the economy when the variance of and, more importantly, all covariances with, the highest past performing industry are removed. As a measure of the percentage decline in aggregate volatility or ‘covolatility’ from removing the highest past return industry, we compute the following statistic,

\[
\%V_t \equiv 1 - \sqrt{\frac{||H_t^R||}{||H_t||}}
\]

(2.72)

where \( || \cdot || \) is a metric of aggregate volatility for the particular matrix. Three metrics are employed: 1) the sum of the squared eigenvalues of the matrix;\(^ {31}\) 2) the maximum squared eigenvalue of the matrix, and 3) the sum of the largest four squared eigenvalues of the matrix. Therefore, \( \%V_t \) represents the amount (in percentage terms) of total volatility or ‘covolatility’ captured or contributed by the highest past return industry.

\(^{30}\)Since \( H_t \) is constructed from realized returns from month \( t - 1 \), it is necessary to skip a month for the sorting procedure in order to obtain unbiased estimates. In fact, since GARCH uses all past observations, we compute our covariance matrices only using odd months, and sort/rank industries based on even month returns, to ensure unbiased estimates. We then reverse the odd-even GARCH-sorting procedure and average the two results.

\(^{31}\)Note that the sum of squared eigenvalues of a matrix equals the sum of the squared entries of the matrix. Hence, this metric is a nice measure of the aggregate volatility or covolatility represented by the matrix.
Of course, this measure is meaningless unless compared to a proper benchmark. Therefore, we compute $\%V_t$ for all industries by removing each industry by order of its past return ranking. That is, at time $t$, we rank the industries based on their return at time $t-2$, and then remove the highest past performing industry from $H_t$ and compute $\%V_t$ as described above. We then remove the second highest industry (replacing the highest industry) and repeat the calculations, doing this for every industry in order of its past return ranking. This procedure is repeated at every time $t$, where the time-series average of the $\%V_t$ measures are reported in Table 2.7. Note that the ranking of industries changes every period, thus the $\%V_t$ measure for the highest past return industry measures the contribution to economy-wide covolatility of the highest past return industry, no matter which particular industry that happens to be. In other words, the identity of the industry itself changes from period to period, but it is always the highest past return industry. The same holds true for the lowest past return industry, as well as all other rankings. As Table 2.7 and Figure 2.4 demonstrate, there is little relation between past performance and future covolatility. In fact, the relation is somewhat hump-shaped, suggesting that conditional industry risk premia are an unlikely explanation for momentum.
Table 2.1: Size, BE/ME, and Momentum Profits Attributed to Risk

Using the various factor models presented in Section 2.3.1, we adjust the 32 portfolios used in the GARCH framework for risk and compute size, BE/ME, individual stock momentum, and industry momentum profits using the residual returns. The factor loadings of each asset under each factor model are estimated using the GARCH time-varying covariance matrices. In addition, we add a co-skewness and co-kurtosis measure to each model, and estimate factor loadings using 36-month prior return rolling regressions. The mean and t-stats of the risk-adjusted profits are reported.

<table>
<thead>
<tr>
<th>Factor Model</th>
<th>(IM) mean (t-stat)</th>
<th>(PR1) mean (t-stat)</th>
<th>(SMB) mean (t-stat)</th>
<th>(HML) mean (t-stat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.00719 (3.77)</td>
<td>0.01123 (4.23)</td>
<td>0.00257 (1.83)</td>
<td>0.00423 (3.20)</td>
</tr>
<tr>
<td>CAPM</td>
<td>0.00662 (3.58)</td>
<td>0.01493 (5.57)</td>
<td>0.00274 (1.96)</td>
<td>0.00396 (3.10)</td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00607 (2.75)</td>
<td>0.00711 (3.00)</td>
<td>0.00060 (0.43)</td>
<td>0.00528 (3.54)</td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00535 (2.37)</td>
<td>0.00629 (2.64)</td>
<td>0.00210 (1.67)</td>
<td>0.00609 (3.73)</td>
</tr>
<tr>
<td>FF</td>
<td>0.00669 (3.58)</td>
<td>0.01232 (4.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00418 (1.79)</td>
<td>0.00664 (2.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00264 (1.12)</td>
<td>0.00436 (1.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carhart</td>
<td>0.00653 (3.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00381 (1.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00268 (1.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF+IM</td>
<td></td>
<td>0.00342 (1.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00652 (2.58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00400 (1.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>0.00691 (3.65)</td>
<td>0.01255 (4.64)</td>
<td>0.00207 (1.58)</td>
<td>0.00513 (3.58)</td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00446 (1.89)</td>
<td>0.00521 (2.11)</td>
<td>-0.00090 (-1.39)</td>
<td>0.00252 (1.86)</td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00366 (1.52)</td>
<td>0.00407 (1.63)</td>
<td>-0.00050 (-0.52)</td>
<td>0.00179 (1.63)</td>
</tr>
<tr>
<td>UPC</td>
<td>0.01254 (4.51)</td>
<td>0.01385 (2.94)</td>
<td>0.00142 (0.82)</td>
<td>0.01161 (6.17)</td>
</tr>
<tr>
<td>co-skew</td>
<td>0.00296 (1.25)</td>
<td>0.00629 (2.48)</td>
<td>0.00041 (0.46)</td>
<td>0.00603 (4.79)</td>
</tr>
<tr>
<td>co-kurt</td>
<td>0.00303 (1.25)</td>
<td>0.00532 (2.08)</td>
<td>0.00101 (1.08)</td>
<td>0.00540 (4.31)</td>
</tr>
</tbody>
</table>

106
Table 2.2: Common Volatility Tests of Size, BE/ME, and Momentum Portfolios

Common ARCH tests using 1 and 6 lags in the information set (COARCH(1) and COARCH(6)) for each of SMB, HML, PRIYR, IM, and the market relative to the within sample principal components factors are conducted, along with an ARCH(12) test for each of the portfolios.

<table>
<thead>
<tr>
<th></th>
<th>Common ARCH Tests&lt;sup&gt;a&lt;/sup&gt;</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARCH(12)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>COARCH(1)</td>
<td>COARCH(6)</td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>$14.79$</td>
<td>$4.90$</td>
<td>$10.68$</td>
<td></td>
</tr>
<tr>
<td>(p-value)</td>
<td>$(0.253)$</td>
<td>$(0.936)$</td>
<td>$(1.00)$</td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>$15.53$</td>
<td>$8.05$</td>
<td>$24.03$</td>
<td></td>
</tr>
<tr>
<td>(p-value)</td>
<td>$(0.213)$</td>
<td>$(0.709)$</td>
<td>$(1.00)$</td>
<td></td>
</tr>
<tr>
<td>PRIYR</td>
<td>$53.28$</td>
<td>$40.89$</td>
<td>$162.55$</td>
<td></td>
</tr>
<tr>
<td>(p-value)</td>
<td>$(0.000)$</td>
<td>$(0.000)$</td>
<td>$(0.000)$</td>
<td></td>
</tr>
<tr>
<td>IM</td>
<td>$22.97$</td>
<td>$20.76$</td>
<td>$98.94$</td>
<td></td>
</tr>
<tr>
<td>(p-value)</td>
<td>$(0.010)$</td>
<td>$(0.036)$</td>
<td>$(0.161)$</td>
<td></td>
</tr>
<tr>
<td>$R_m - r_f$</td>
<td>$19.23$</td>
<td>$0.86$</td>
<td>$1.89$</td>
<td></td>
</tr>
<tr>
<td>(p-value)</td>
<td>$(0.083)$</td>
<td>$(0.999)$</td>
<td>$(1.00)$</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> The COARCH(1) and COARCH(6) test statistics are distributed as $\chi^2_{df}$ with degrees of freedom $df = \left(\frac{(1+2)/(K+1)}{2} - K\right)$, and $df = \left(\frac{(1+2)/(K+1)}{2} \cdot 0 - K\right)$, respectively, where $K$ is the number of factors or basis assets used to explain return second moments.

<sup>b</sup> The ARCH(12) $T \times R^2$ statistic is distributed as $\chi^2_{12}$. 

107
Table 2.3: Minimizing Risk of Size, BE/ME, and Momentum Strategies

Portfolios are designed to load on one of the characteristics or factor-mimicking portfolios derived from the characteristic, while at the same time minimizing variance and hedging out exposure to other characteristics. The optimization of five portfolios are conducted for each of size, BE/ME, and momentum strategies, utilizing the conditional covariance estimates of 32 portfolios via the Ledoit and Santa-Clara (1998) GARCH(1,1) process. The five portfolio optimization problems are:

\[
\begin{align*}
(1) \quad & \min \ [w' R_t - R_{ml}]^2 \\
& \text{s.t. } w' R_{t-1} = \mu_{t-1} + \sigma_{t-1} \\
(2) \quad & \min w' H_t w \\
& \text{s.t. } w' R_{t-1} = \mu_{t-1} + \sigma_{t-1} \\
& w' \beta_t = 0 \\
& w' \beta_{smt} = 0 \\
& w' \beta_{hml} = 0 \\
(3) \quad & \min w' H_t w \\
& \text{s.t. } w' J_{M,t} = J_t + \sigma_{J,t} \\
(4) \quad & \min w' H_t w \\
& \text{s.t. } w' R_{t-1} = \mu_{t-1} + \sigma_{t-1} \\
& w' \beta_{PC} = 0 \\
(5) \quad & \min w' H_t w \\
& \text{s.t. } w' J_{M,t} = J_t + \sigma_{J,t} \\
\end{align*}
\]

The time-series (monthly) average of the Sharpe ratios of each of these strategies are reported below, along with the time-series standard errors of the Sharpe measures in parentheses.

<table>
<thead>
<tr>
<th>Optimization Problem</th>
<th>Momentum</th>
<th>Size</th>
<th>BE/ME</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>0.149</td>
<td>0.124</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>(0.0513)</td>
<td>(0.0523)</td>
<td>(0.0491)</td>
</tr>
<tr>
<td>(2)</td>
<td>0.306</td>
<td>0.235</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td>(0.0836)</td>
<td>(0.0588)</td>
<td>(0.0468)</td>
</tr>
<tr>
<td>(3)</td>
<td>0.292</td>
<td>0.222</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>(0.0668)</td>
<td>(0.0568)</td>
<td>(0.0436)</td>
</tr>
<tr>
<td>(4)</td>
<td>0.336</td>
<td>0.219</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(0.0892)</td>
<td>(0.0741)</td>
<td>(0.0669)</td>
</tr>
<tr>
<td>(5)</td>
<td>0.428</td>
<td>0.228</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>(0.0706)</td>
<td>(0.0741)</td>
<td>(0.0712)</td>
</tr>
</tbody>
</table>

108
Table 2.4: Non-Nested Cross-Sectional Regression Tests of Factor Models

Using the factor models outlined in Section 2.3.1, we run non-nested cross-sectional regression J-tests of Davidson and MacKinnon (1981) to compare which factor model best predicts future expected returns. Specifically, we run a cross-sectional regression of returns on estimated factor loadings in the style of Fama and MacBeth (1973) for one of the models, and then place the fitted values from this regression into the second model as an additional explanatory variable. We then regress returns cross-sectionally on estimated factor loadings for the second model plus the fitted values from the first regression. The time-series of coefficients on the fitted values ($\psi$) are averaged in the style of Fama and MacBeth (1973), and reported below. A $\psi$ coefficient close to 1 rejects the second model ($H_0$) in favor of the first ($H_1$). The test is conducted on the 32 portfolios used to estimate the GARCH conditional covariance matrices, and on the set of all (CRSP-listed) individual securities. The test is described in more detail in Section 2.5.1, and both OLS and GLS estimates are provided, where a feasible GLS procedure is used for the individual security regressions. T-stats are in parentheses.

<table>
<thead>
<tr>
<th>Factor Models</th>
<th>32 Portfolios</th>
<th>Individual Stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>$H_1$ (fitted values)</td>
<td></td>
</tr>
<tr>
<td>$PC^{(2)}$</td>
<td>FF+IM</td>
<td>$\psi_{OLS}$ (t-stat)</td>
</tr>
<tr>
<td>FF+IM</td>
<td>$PC^{(2)}$</td>
<td>0.826 (39.76)</td>
</tr>
<tr>
<td>FF</td>
<td>FF+IM</td>
<td>0.894 (49.63)</td>
</tr>
<tr>
<td>FF+IM</td>
<td>FF</td>
<td>-0.744 (-0.95)</td>
</tr>
<tr>
<td>Carhart</td>
<td>FF+IM</td>
<td>0.816 (11.24)</td>
</tr>
<tr>
<td>FF+IM</td>
<td>Carhart</td>
<td>0.873 (5.46)</td>
</tr>
<tr>
<td>$UPC$</td>
<td>FF+IM</td>
<td>0.607 (19.08)</td>
</tr>
<tr>
<td>FF+IM</td>
<td>$UPC$</td>
<td>1.023 (44.28)</td>
</tr>
</tbody>
</table>
Table 2.5: Predicting Future Volatility

Using the product moment matrix as a proxy for the true conditional covariance matrix at time $t$, $V_t = R_t R_t'$, we run a 'horse race' between various factor models for predicting the 'true' covariance matrix, $\hat{V}_t$, out of sample. The Ledoit and Santa-Clara GARCH technique is employed for each of the factor models on the 32 asset portfolios described in Section 2.3, and the estimated covariance matrix at time $t$, $\hat{V}_t$, is compared to $V_t$. Three distance metrics, described in Section 2.5.2.1, are employed to measure how far $\hat{V}_t$ deviates from $V_t$ on average. In addition, we run the same 'horse race' for predicting the future volatility of 50 individual stocks, selected randomly. Here, rather than using a GARCH technique, we estimate $\hat{V}_t$ for each factor model using equation (2.68) from Section 2.5.2.2.

<table>
<thead>
<tr>
<th>Competing Models</th>
<th>Same Mean Estimate</th>
<th>Different Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22 Portfolios</td>
<td>Individual Stocks</td>
</tr>
<tr>
<td>$D_1$</td>
<td>$D_2$</td>
<td>$D_3$</td>
</tr>
<tr>
<td>FF</td>
<td>155.12</td>
<td>123.42</td>
</tr>
<tr>
<td>FF+LM</td>
<td>81.56</td>
<td>49.52</td>
</tr>
<tr>
<td>Carhart</td>
<td>2171.71</td>
<td>2139.90</td>
</tr>
<tr>
<td>PC</td>
<td>2066.70</td>
<td>2035.68</td>
</tr>
<tr>
<td>UPC</td>
<td>351.55</td>
<td>320.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Competing Models</th>
<th>32 Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>$D_2$</td>
</tr>
<tr>
<td>CAPM</td>
<td>871962.84</td>
</tr>
<tr>
<td></td>
<td>9.24</td>
</tr>
<tr>
<td>M</td>
<td>35.59</td>
</tr>
<tr>
<td>SMB</td>
<td>180.84</td>
</tr>
<tr>
<td>HML</td>
<td>28676.45</td>
</tr>
<tr>
<td>PR1YR</td>
<td>13008.97</td>
</tr>
</tbody>
</table>

110
Table 2.6: Remaining Volatility

We examine the residual covariance matrix implied by each model by hedging out the factors associated with the model, and computing how much volatility remains. Using equation (2.63), we compute the residual covariance matrix, $\Omega_t$, under each factor model at a point in time. We then form minimum variance portfolio weights using $\Omega_t$ and apply these weights to residual returns out of sample, where residual returns are also defined according to the factor model. Specifically, the residual return on this portfolio is defined as:

$$\epsilon_t = \frac{\Omega_t^{-1} \Gamma_t}{\epsilon\Omega_t^{-1} \epsilon}$$

where $\epsilon$ is a $(32 \times 1)$ vector of ones, $R_t$ is the vector of asset returns, $\hat{\beta}_t$ is the $(32 \times K)$ matrix of factor loadings, estimated from the conditional covariance matrix, $H_t$, and $F_t$ is the $(K \times 1)$ vector of factor realizations. We compute these residuals for each factor model and report their sums of squares, as well as their sum of squared deviations from their mean. In addition, we also compute portfolio weights using only the diagonal elements of $\Omega_t$ (e.g., just the variances), and equal-weighting the residual returns.

### Panel A: Metrics of Residuals

<table>
<thead>
<tr>
<th>Factor</th>
<th>MVP</th>
<th>Variance Only</th>
<th>Equal-Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$\sum_{t=1}^{T} \epsilon_t^2$</td>
<td>$\sum_{t=1}^{T} (\epsilon_t - \bar{\epsilon}_t)^2$</td>
<td>$\sum_{t=1}^{T} \epsilon_t^2$</td>
</tr>
<tr>
<td>FF</td>
<td>0.3164</td>
<td>0.3152</td>
<td>0.0941</td>
</tr>
<tr>
<td>FF+IM</td>
<td>0.3704</td>
<td>0.3690</td>
<td>0.0953</td>
</tr>
<tr>
<td>Carhart</td>
<td>0.5134</td>
<td>0.5106</td>
<td>0.1732</td>
</tr>
<tr>
<td>PC</td>
<td>0.0119</td>
<td>0.0111</td>
<td>0.0476</td>
</tr>
<tr>
<td>UPC</td>
<td>0.2409</td>
<td>0.2406</td>
<td>0.0419</td>
</tr>
</tbody>
</table>

### Panel B: Portfolio Residual Returns

<table>
<thead>
<tr>
<th>Factor</th>
<th>MVP</th>
<th>Variance Only</th>
<th>Equal-Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>mean (t-stat)</td>
<td>mean (t-stat)</td>
<td>mean (t-stat)</td>
</tr>
<tr>
<td>FF</td>
<td>0.0018 (1.22)</td>
<td>-0.0018 (-2.24)</td>
<td>-0.0045 (-9.60)</td>
</tr>
<tr>
<td>FF+IM</td>
<td>0.0019 (1.20)</td>
<td>-0.0014 (-1.75)</td>
<td>-0.0014 (-0.57)</td>
</tr>
<tr>
<td>Carhart</td>
<td>0.0027 (1.44)</td>
<td>-0.0011 (-1.00)</td>
<td>-0.0036 (-6.02)</td>
</tr>
<tr>
<td>PC</td>
<td>0.0003 (1.17)</td>
<td>-0.0009 (-1.63)</td>
<td>-0.0029 (-5.69)</td>
</tr>
<tr>
<td>UPC</td>
<td>-0.0008 (-0.63)</td>
<td>-0.0077 (-21.15)</td>
<td>-0.0091 (-50.01)</td>
</tr>
</tbody>
</table>

### Single Factor Models

<table>
<thead>
<tr>
<th>Factor</th>
<th>MVP</th>
<th>Variance Only</th>
<th>Equal-Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>0.0007 (0.66)</td>
<td>-0.0018 (-2.33)</td>
<td>-0.0056 (-5.30)</td>
</tr>
<tr>
<td>IM</td>
<td>0.0056 (2.09)</td>
<td>0.0045 (2.05)</td>
<td>0.0072 (2.39)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.0029 (1.08)</td>
<td>0.0041 (1.98)</td>
<td>-0.0015 (-0.44)</td>
</tr>
<tr>
<td>HML</td>
<td>0.0036 (1.67)</td>
<td>0.0054 (2.52)</td>
<td>0.0086 (3.74)</td>
</tr>
<tr>
<td>PRIYR</td>
<td>0.0053 (2.26)</td>
<td>0.0053 (2.39)</td>
<td>0.0055 (2.34)</td>
</tr>
</tbody>
</table>
Table 2.7: Risk of High and Low Past Return Industries

We remove each industry, based on its past return ranking, from the covariance matrix at each point in time, and compute the contribution to overall covolatility contributed by each ranked industry. The rankings change each period, and 1 corresponds to the highest past return industry, and 20 to the lowest at every point in time. We compute both the contribution to total covolatility, \( \%V_t \), contributed by each ranked industry, as well as the change in contribution to covolatility \( \Delta \%V_t \). The time-series averages of \( \%V_t \) (see equation (2.72)) and \( \Delta \%V_t \) are reported below for the three metrics described in Section 2.8. The values below are percentages, and thus are multiplied by 100. All t-stats are reported in parentheses.

<table>
<thead>
<tr>
<th>Past Return</th>
<th>( \sum_{n=1}^{N} \text{eig}_n )</th>
<th>max(eig)</th>
<th>( \sum_{n=1}^{4} \text{eig}_n )</th>
<th>( \sum_{n=1}^{N} \text{eig}_n )</th>
<th>max(eig)</th>
<th>( \sum_{n=1}^{4} \text{eig}_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ranking</td>
<td>%( V_t )</td>
<td></td>
<td>%( V_t )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>highest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>4.04</td>
<td>3.94</td>
<td>3.97</td>
<td>0.03 (0.50)</td>
<td>0.01 (0.12)</td>
<td>0.03 (0.38)</td>
</tr>
<tr>
<td>2.</td>
<td>4.13</td>
<td>4.18</td>
<td>4.11</td>
<td>0.17 (3.06)</td>
<td>0.15 (2.45)</td>
<td>0.17 (2.98)</td>
</tr>
<tr>
<td>3.</td>
<td>4.60</td>
<td>4.84</td>
<td>4.62</td>
<td>0.17 (2.90)</td>
<td>0.15 (2.53)</td>
<td>0.16 (2.71)</td>
</tr>
<tr>
<td>4.</td>
<td>4.49</td>
<td>4.76</td>
<td>4.52</td>
<td>0.10 (2.15)</td>
<td>0.11 (2.19)</td>
<td>0.10 (2.16)</td>
</tr>
<tr>
<td>5.</td>
<td>4.41</td>
<td>4.65</td>
<td>4.43</td>
<td>0.05 (1.08)</td>
<td>0.04 (0.95)</td>
<td>0.05 (1.09)</td>
</tr>
<tr>
<td>6.</td>
<td>4.49</td>
<td>4.78</td>
<td>4.51</td>
<td>0.09 (1.90)</td>
<td>0.11 (2.15)</td>
<td>0.09 (1.85)</td>
</tr>
<tr>
<td>7.</td>
<td>4.33</td>
<td>4.65</td>
<td>4.37</td>
<td>0.03 (0.80)</td>
<td>0.05 (1.16)</td>
<td>0.04 (0.89)</td>
</tr>
<tr>
<td>8.</td>
<td>4.35</td>
<td>4.69</td>
<td>4.39</td>
<td>0.03 (0.70)</td>
<td>0.05 (1.19)</td>
<td>0.04 (0.86)</td>
</tr>
<tr>
<td>9.</td>
<td>4.38</td>
<td>4.67</td>
<td>4.41</td>
<td>-0.00 (-0.08)</td>
<td>0.00 (0.10)</td>
<td>-0.00 (-0.03)</td>
</tr>
<tr>
<td>10.</td>
<td>4.36</td>
<td>4.65</td>
<td>4.39</td>
<td>-0.03 (-0.54)</td>
<td>-0.03 (-0.54)</td>
<td>-0.02 (-0.41)</td>
</tr>
<tr>
<td>11.</td>
<td>4.16</td>
<td>4.44</td>
<td>4.19</td>
<td>-0.09 (-2.10)</td>
<td>-0.10 (-2.07)</td>
<td>-0.09 (-1.97)</td>
</tr>
<tr>
<td>12.</td>
<td>4.41</td>
<td>4.75</td>
<td>4.45</td>
<td>-0.04 (-0.87)</td>
<td>-0.03 (-0.62)</td>
<td>-0.04 (-0.78)</td>
</tr>
<tr>
<td>13.</td>
<td>4.47</td>
<td>4.80</td>
<td>4.52</td>
<td>-0.06 (-1.29)</td>
<td>-0.06 (-1.25)</td>
<td>-0.06 (-1.21)</td>
</tr>
<tr>
<td>14.</td>
<td>4.47</td>
<td>4.78</td>
<td>4.51</td>
<td>-0.10 (-2.22)</td>
<td>-0.12 (-2.43)</td>
<td>-0.11 (-2.27)</td>
</tr>
<tr>
<td>15.</td>
<td>4.49</td>
<td>4.79</td>
<td>4.52</td>
<td>-0.01 (-0.24)</td>
<td>-0.00 (-0.09)</td>
<td>-0.01 (-0.18)</td>
</tr>
<tr>
<td>16.</td>
<td>4.29</td>
<td>4.53</td>
<td>4.30</td>
<td>-0.05 (-1.12)</td>
<td>-0.04 (-0.69)</td>
<td>-0.05 (-1.03)</td>
</tr>
<tr>
<td>17.</td>
<td>4.72</td>
<td>4.90</td>
<td>4.72</td>
<td>-0.02 (-0.30)</td>
<td>-0.01 (-0.24)</td>
<td>-0.02 (-0.30)</td>
</tr>
<tr>
<td>18.</td>
<td>4.19</td>
<td>4.33</td>
<td>4.17</td>
<td>-0.13 (-2.37)</td>
<td>-0.12 (-2.11)</td>
<td>-0.12 (-2.24)</td>
</tr>
<tr>
<td>19.</td>
<td>4.23</td>
<td>4.22</td>
<td>4.18</td>
<td>-0.17 (-3.26)</td>
<td>-0.14 (-2.48)</td>
<td>-0.17 (-3.05)</td>
</tr>
<tr>
<td>lowest 20.</td>
<td>4.33</td>
<td>4.23</td>
<td>4.26</td>
<td>-0.17 (-2.78)</td>
<td>-0.17 (-2.48)</td>
<td>-0.18 (-2.77)</td>
</tr>
<tr>
<td>High - Low</td>
<td>-0.30</td>
<td>-0.29</td>
<td>-0.29</td>
<td>0.21</td>
<td>0.17</td>
<td>0.21</td>
</tr>
<tr>
<td>(t-stat)</td>
<td>(-1.89)</td>
<td>(-1.60)</td>
<td>(-1.78)</td>
<td>(1.97)</td>
<td>(1.55)</td>
<td>(1.90)</td>
</tr>
</tbody>
</table>
Figure 2.1: Distribution of Trading Strategy Profits

Plots of the trading strategy profits over time are presented below, along with summary statistics on the distribution of each of the strategies.

The W-test is a Wald test for normality given by \( W = T\left[\frac{\text{skew}^2}{6} + \frac{(\text{kurt} - 3)^2}{24}\right] \sim \chi^2(2) \).
Figure 2.2: Distributions for Zero-Intercept F-test Statistic Under the Null (CAPM), Risk-Based, and Non-Risk Alternatives

The distributions for the null (the CAPM), risk-based alternative, and non-risk alternative are $F_{4,378}(0)$, $F_{4,378}(5.81)$, and $F_{4,378}(57.35)$, respectively. The probabilities are calculated using an interval width of 0.02. The test statistic is 9.25, represented by the vertical line.
Figure 2.3: GARCH vs. Rolling Regression-Estimated Market $\beta$'s

The GARCH-estimated (Ledoit and Santa-Clara (1998), GARCH(1,1) process) market factor loadings are plotted against rolling regression (past 36 months of returns) estimates, for each of the four strategies designed to exploit the anomalies.
Figure 2.4: Relation Between Conditional Risk and Past Performance

The relation between covolatility at time $t$, $\%V_t$, and the ranking of industries based on past performance from time $t - 12$ to $t - 2$ is plotted below. The metric used for the matrices is the sum of squares of all of the eigenvalues of the matrix. The black line is the best-fit curve (two degree polynomial) of the data points.
References for Chapter 2


Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A theory of overconfidence, self-attribution, and security market under- and over- reac-


Karolyi, A. and B. Kho. 1996, Time-varying risk premia and the returns to buying winners and selling losers: Caveat emptor et venditor, Working paper. The Ohio State University, Columbus, OH.


CHAPTER 3

The Geography of Investment: Are There Gains to Investing Locally?

3.1 Introduction

Until recently, the role of geography in domestic equity markets has been largely overlooked in the academic literature. This is particularly surprising given its recognition by practitioners and the national media. For instance, a recent article states that fund managers of state and regional funds believe they can obtain abnormal performance by investing in “their own back yard,” with the belief that “geography gives them a competitive advantage.”¹ This belief in a geographic investment advantage is perhaps best exemplified by Conrad Herrmann, manager of Franklin California Growth Fund, who claims.

“We have a unique advantage over someone investing from over 3,000 miles away. We read the local newspapers, socialize with people that work for these companies, and we can get a sense for how the region’s doing.”

While a large number of studies examine the relation between geography and investment in the international setting, only two recent studies consider this


121
relation within a domestic setting. Coval and Moskowitz (1997a) examine the domestic holdings of U.S. mutual fund managers in relation to their location. They find a strong preference for geographically-proximate equities across a wide range of fund managers, and document a positive relation between firm leverage and degree of local bias exhibited by a fund manager, and a negative relation between firm size and local bias. They contend that these results support an information-based explanation for local equity preference. Huberman (1997) also examines the existence of a local equity preference, but at the household rather than institutional level. He finds that individual investors are more likely to invest in their local U.S. Regional Bell Operating Companies (RBOCs) than any other RBOC, even though the companies are listed on the same exchange. In contrast to Coval and Moskowitz (1997a), however, he does not attribute such local preferences to information asymmetries. Rather, he contends that investors have psychological biases toward the familiar, and thus prefer local investments. While a preference for the familiar seems reasonable for individual investors, such biases should be unlikely among professional money managers who compete for funds. If biasing toward the familiar results in a significant loss of diversification, without additional economic benefit, then such investment behavior is unlikely to be sustainable in a competitive environment, where mutual fund managers must compete for funds, and have strong incentives to outperform their constituents. Moreover, if psychology drives local preference, then returns to investing locally should be no different than those from non-local investments.

Both the Coval and Moskowitz (1997a) and Huberman (1997) studies, however, rely on individual cross-sections of data, and are therefore limited in identifying the underlying sources of the preference for local investments. This paper undertakes a time-series analysis of investor preferences to yield additional insights into the local bias investment phenomenon. In particular, we examine the
performance of fund manager holdings in relation to geography to help resolve whether an investor's local bias arises from superior local information or from a psychological preference for the familiar. If investors invest in local firms in which they have superior information, their performance in these firms should exceed that of their distant holdings and exceed that of passive benchmarks with similar risk characteristics.²

In a broader context, examining the performance of local investments made by mutual funds will also address whether fund managers have any stock selection or timing ability. Numerous studies have documented only slight ability of fund managers to select or time their investments, as risk-adjusted portfolio returns among fund managers are close to zero.³ However, fund managers may exhibit significant performance in their local investments, where they may have an informational advantage. Conversely, if they do not exhibit any performance, even in their local holdings, then not only does their propensity to invest locally appear to be inefficient (in a mean-variance sense) and possibly due to sources other than information, but their lack of ability will be even more apparent, as local firms are those in which fund managers are expected to have the greatest advantage. Therefore, examining the performance of mutual fund managers' local holdings will provide insights into the ability (or lack thereof) of these professional money managers, as well as shed light on the reason these managers choose to bias their

²As Huberman (1997) points out, ceteris paribus an investor will not hold all firms in which he has an information advantage - only those that are expected to outperform. Yet, if uninformed, distant investors recognize the adverse selection problem they face when trading with informed locals in these firms, in equilibrium, local investors will end up holding a disproportionate share of all firms in which they have an information advantage - and getting appropriately compensated for doing so.

³See for example, Grinblatt and Titman (1989, 1993, 1994), Daniel, Grinblatt, Titman, and Wermers (1997), Carhart (1997), and a whole host of others. Daniel, Grinblatt, Titman, and Wermers (1997) for example, find that the average fund manager exhibits about 50 basis points per year in excess of similar size, book-to-market equity (BE/ME), and momentum passive benchmark portfolios. However, all of this performance is concentrated in the first half of the sample (1974-1985), and slight underperformance is documented in the latter half.
holdings locally.

We examine the domestic portfolio holdings of U.S. mutual fund managers from September 30, 1980 to December 31, 1994, using a different dataset than Coval and Moskowitz (1997a). We find a strong preference for geographically local equities that has declined over time, and strong abnormal performance in local holdings relative to distant holdings. In addition, the degree of local preference and significance of abnormal performance are more acute among mutual funds exhibiting active management styles, while weak among more passive managers. These findings point to an information-based explanation for the preference for local equities. Specifically, if local fund managers receive information about local stocks through both their own research and possible relationships with local corporate executives, we would expect such a bias in their portfolio holdings. Being in the same geographic area may allow managers to visit a firm's operations more frequently, follow the local media, and talk to suppliers and customers of the firm. Non-local fund managers may also engage in this research, but perhaps at a greater cost. As information dissemination improves over time, this cost decreases, and hence so should the degree of local bias. On the other hand, information received through close relationships between fund managers and local corporate executives cannot be obtained by non-local fund managers, and thus should be unaffected over time. This pattern is consistent with our finding of a slight decline in the intensity of local bias over time.

However, the most compelling evidence for an information-based explanation of geographic investment preferences is our finding that the locally held portion of a portfolio significantly outperforms the non-local portion by approximately 2.7% per year on a risk-adjusted basis. These findings are largely contained in the first half of the sample period, however. Furthermore, the abnormal local investment
performance of active funds is strong, while that of passive funds is weak, and funds that exhibit the largest degree of local bias also exhibit the largest abnormal performance in their local holdings, providing strong support for our information hypothesis. In addition, fund managers from small, remote cities, who exhibit the strongest preference for local stocks, also exhibit the strongest performance in their local investments. Since small, remote markets arguably have the tightest investment community ties, and have the least amount of competition among fund managers to compete away rents from superior information, these results are consistent with an information story.

Finally, there may be other explanations for the local preference in mutual fund holdings. For example, agency problems between fund managers and their clients could be driving the results. For instance, mutual fund managers may simply find it more convenient to hold proximate investments. In other words, fund managers may exhibit a ‘convenience’ bias or travel aversion. Therefore, locally-biased portfolios may represent the smaller cost (travel time, effort, convenience) of investing locally. In addition, fund managers may feel more comfortable with local investments, perceiving them to be less risky, or they may derive some benefit from investing with the ‘home team.’ If close ties exist between local fund managers and local corporate executives, an agency problem may result with mutual benefits derived from keeping fund money in local companies. However, the abnormal performance of local investments contradicts such hypotheses and supports information advantages as the plausible explanation for local equity preference. Therefore, based on evidence in this paper, we suggest that the local bias represents superior information availability for local fund managers, and hence may be rational.
In the final section of the paper, we shift the focus of the analysis from a fund manager to a stock perspective. Coval and Moskowitz (1997a) find that local equity preference is inversely related to firm size and positively related to firm leverage. Furthermore, these and other firm characteristics have been found to be related to abnormal returns and aid in explaining the cross-section of expected stock returns. Consequently, a link may exist between the degree of local investment (a proxy for asymmetric information) and these cross-sectional asset pricing anomalies. We investigate this possibility in Section 6 and find that the degree to which a stock is held by local fund managers is a positive predictor of expected returns, even when controlling for size, book-to-market equity (BE/ME), and momentum.

The remainder of the paper is organized as follows. Section 2 describes the data and methodology employed in the paper. Section 3 presents the results of tests for local bias over time. Section 4 briefly examines turnover rates of local and non-local investments. Section 5 evaluates the relative performance of local and distant holdings from a fund manager perspective, and section 6 focuses the analysis from a stock perspective, addressing the relation between expected returns and degree of local investment in a stock, with a possible link to cross-sectional asset pricing anomalies. Section 7 concludes the paper.

---

4The size effect was first documented by Banz (1981), Keim (1983), Reinganum (1983), and Roll (1983); the effect of leverage was first documented by Bhandari (1988); the book-to-market effect was first documented by Stattman (1980), and Rosenberg, Reid, and Lanstein (1985); the momentum effect by Jegadeesh and Titman (1993).
3.2 Data and Methodology

3.2.1 Data

Our data sample consists of merging the Investment Company Common Stock Holdings and Transactions tapes from CDA Investment Technologies, Inc. of Rockville, Maryland, with latitude and longitude data obtained from Geographic Names Information System Digital Gazetteer, published by the U.S. Geological Survey. We examine only those funds primarily devoted to equity investments, reducing the sample to approximately 2,600 funds. We then match each fund with its management company from Nelson's Directory of Investment Managers, and obtain the corresponding location of the fund manager (city and state), which is then translated into latitude and longitude coordinates.

We examine quarterly domestic equity holdings of mutual funds from September 30, 1980 to December 31, 1994, with the restriction that a stock be listed on the Center for Research in Securities Prices (CRSP) tapes at the time it is held. The location of each fund's headquarters is obtained from Disclosure and translated into latitude and longitude coordinates. Our final sample consists of

---

5These tapes list the quarterly equity holdings of over 5,000 mutual funds from December 31, 1974 to December 31, 1994, along with a self-declared investment objective beginning June 30, 1980. This database does not suffer from survivorship bias. Further details on the construction of this database and summary statistics can be found in Wermers (1997).

6Headquarters location is used as opposed to state of incorporation, for the simple reason that firms tend to incorporate in a state with favorable tax laws, bankruptcy laws, etc., rather than for operational reasons, and typically do not have the majority of their operations in their state of incorporation. In fact, few firms in our sample were headquartered in the same state they were incorporated.

7International funds and firms and funds located in Alaska, Hawaii, and Puerto Rico, are eliminated from the analysis. International funds are excluded because their domestic equity holdings are likely to be a small percentage of total fund assets, and because their objectives for investment may differ from other equity funds. In section 3, however, we examine these funds specifically. Alaskan, Hawaiian, and Puerto Rican funds and firms are eliminated to avoid the potential problem of Hawaiian based funds and stocks, for instance, exaggerating the effect of distance on portfolio choice. The number of Hawaiian, Alaskan, and Puerto Rican funds and stocks is very small, however.
between 118 (1980:3) and 1.278 (1994:4) mutual funds holding between 499 and 4.617 different companies over the entire sample period.

3.2.2 Specification Test

To assess whether our sample of mutual funds exhibit a preference for local stocks, we compute two distance measures for each fund. Indexing funds by $i$ and stocks by $j$, we compute the actual distance, $d_{ij}$, between fund $i$ and the headquarters of each firm $j$ it holds.\footnote{Since locations are identified by latitude and longitude, we calculate the arclength, $d_{ij}$, between each location as.}

$$d_i = \sum_{j=1}^{N} \theta_{ij} \cdot d_{ij}, \quad (3.2)$$

$$\theta_{ij} = \frac{H_{ij} \cdot P_j}{\sum_{j=1}^{N} H_{ij} \cdot P_j} \quad (3.3)$$

where fund $i$'s portfolio weight on firm $j$ is denoted by $\theta$. $H_{ij}$ is the number of shares held of stock $j$ by fund $i$ at the beginning of the quarter, $P_j$ is the price of stock $j$, and $N$ is the total number of stocks available for investment. Since fund $i$ holds a subset of the $N$ assets, if stock $j$ is not held by fund $i$, then $\theta_{ij} = 0$.\footnote{Since only CRSP-listed equities are examined, hypothetical rather than actual portfolio weights are employed, where we recompute the weights on each holding as if the true portfolio consisted of CRSP-listed equities only.}

$$d_{ij} = \arccos \left\{ \cos(lat_i) \cos(lon_i) \cos(lat_j) \cos(lon_j) + \cos(lat_i) \sin(lon_i) \cos(lat_j) \sin(lon_j) + \sin(lat_i) \sin(lat_j) \right\} (2\pi / 3.1)$$

where lat and lon are fund and company latitudes and longitudes and $r$ is the radius of the earth ($\approx 6.378$ km).
We then determine the average distance fund $i$ is from the universe of domestic assets it could invest in, weighting the distance between fund $i$ and every stock in the economy by the share of aggregate fund investment that particular stock comprises,

$$
\overline{d_i} = \sum_{j=1}^{N} \Psi_j \cdot d_{ij}, \quad (3.4)
$$

$$
\Psi_j = \frac{\sum_{i=1}^{F} H_{ij} \cdot P_j}{\sum_{j=1}^{N} \sum_{i=1}^{F} H_{ij} \cdot P_j}, \quad (3.5)
$$

where stock $j$’s share of aggregate fund investment is given by $\Psi_j$ and $F$ is the total number of funds. This ‘market’ weighted distance measure, using the aggregate actual holdings of all mutual funds, is exactly equal to $d_i$ (i.e., no local bias) if all funds hold the same portfolio. The local bias measure for fund $i$ is the difference between these two measures: $\overline{d_{i,t}} - d_{i,t}$.

Due to serial correlation in fund holdings, we first average local bias measures over time for each fund, and then average across funds. Thus, the aggregate degree of local bias in the economy is measured as:

$$
\bar{x} = \frac{1}{FT} \sum_{i=1}^{F} \sum_{t=1}^{T} \left( \overline{d_{i,t}} - d_{i,t} \right), \quad (3.6)
$$

where $T$ is the number of time periods (58 quarters).13

---

10 Only those firms being held by at least one fund are considered as the universe of assets available for investment, rather than all firms in the U.S. for which data is available. We also ran tests using all available stocks (regardless of whether they were held by at least one of our funds) as the set of equities available for investment, and found very similar results.

11 We ran tests weighting the distance between a fund and each stock in the economy by the share of market capitalization each stock comprises, and equal weighting every distance between a fund and stock. We report only the results for the measure in equation (3.4) for brevity and because they are the most conservative.

12 For example, a fund with a strong local bias in one period will likely appear to have a local bias in future periods, even if its holdings move away from local stocks over time.

13 We also compute a value-weighted average difference measure with virtually identical results.
Our null hypothesis, that \( \bar{d} \) is not significantly different from zero, states that funds should hold stocks that are the same distance from them as the average stock being held by funds. That is, deviations from the average stock being held should be unrelated to distance. Furthermore, weighting the market distance by \( \Psi \) (the aggregate actual holdings of all funds) ensures that if funds hold the same portfolio, \( \bar{d} \) will be zero, regardless of where funds are located. Thus, another interpretation of the null hypothesis is that distance does not have an effect on portfolio choice.\(^{14}\)

Since the sum of holdings of all funds equals the universe of assets being held, distance measures of different funds may be correlated. However, with our large sample size (1,278 funds), this correlation should be very small. Furthermore, even though we might expect mutual fund holdings to be somewhat correlated due to manager herding, the use of momentum strategies, and the practice of benchmarking for performance evaluation,\(^{15}\) the average distance measures of funds are not expected to be more than negligibly correlated. Thus, we assume cross-sectional independence between distance measures of funds.

As another measure of geographic influence, however, we also compute the covariance between a fund's holdings deviation and distance deviations. Fund \( i \)'s holdings deviation is the difference between the portfolio weight it places on stock \( j \) and the average portfolio weight all other funds place on stock \( j \). Fund \( i \)'s distance deviation is the difference between the distance it is from stock \( j \)

\(^{14}\)There is the possibility that mutual fund managers should not be holding a geographically diverse portfolio in the first place. For instance, if clients hold several mutual funds across a wide geographic area, then the need for fund managers to diversify their holdings is diminished. Thus, managers may simply minimize their information gathering and monitoring costs by specializing in local stocks. Such an outcome is reached if clients are able to diversify their own portfolios more efficiently than fund managers could do for them. However, Coval and Moskowitz (1997b) find that the majority of clients tend to be local, providing little support for this conjecture.

\(^{15}\)As well as, perhaps, for cultural, psychological, or agency reasons.
and the average distance all funds are from stock \( j \). If fund \( i \) biases its holdings toward local investments, then the covariance between these two measures will be positive. Again, averaging over time for each fund and then cross-sectionally,

\[
\bar{x}_{cov} = \frac{1}{FT} \sum_{i=1}^{F} \sum_{t=1}^{T} \sum_{j=1}^{N} (\theta_{i,j,t} - \bar{\theta}_{j,t})(\bar{d}_{j,t} - \bar{d}_{i,j,t}),
\]

(3.7)

where \( \bar{\theta}_{j,t} = \frac{1}{F} \sum_{i=1}^{F} \theta_{i,j,t} \), and \( \bar{d}_{j,t} = \frac{1}{F} \sum_{i=1}^{F} d_{i,j,t} \).\(^{16}\) If a fund manager places more (less) weight on a stock than the average fund manager, and that stock is closer to (farther from) him on average, then the covariance measure will be positive.

The results of these tests are presented in Section 3.

### 3.3 Time Series Analysis of Geographic Equity Preference

#### 3.3.1 Tests on U.S. Mutual Fund Domestic Equity Holdings

Table 3.1 presents the results of tests for local bias based on the \( \bar{x} \), and \( \bar{x}_{cov} \) measures, and reports the value of these measures at the end of each year. As shown in the table, fund managers bias portfolios by more than 70 km, a preference for local equities that appears throughout the time period. However, the mean degree of local bias per year is over 100 km. Test statistics for these measures are highly significant, even for the non-parametric Wilcoxon signed rank tests, providing strong evidence of a geographic preference for local equities among U.S. mutual fund managers. Furthermore, because a large number of funds and companies are located in the New York area,\(^{17}\) this could be driving the results. Therefore, we compute our distance measures excluding all New York based fund managers and find that the degree of local preference is actually more pronounced. Thus,

\(^{16}\)We also compute value-weighted (by fund total asset value) averages with virtually identical results.

\(^{17}\)Approximately 6.1% of stocks held are located in New York, and approximately 20.7% of our funds are based in New York.
consistent with the findings in Coval and Moskowitz (1997a), we find a significant preference for local equities among mutual fund managers, using a different data set.

In addition, the degree of local bias slightly diminishes over time. Table 3.1 and Figure 1, which plots the average local bias measure, $\bar{z}$, over time, demonstrate this clearly. Furthermore, the degree of local bias is stronger in the first half of the sample period, where fund managers invested in firms approximately 150 km closer to them (1980-1986), while only biasing their portfolios by 40 km in the latter half of the sample. The decreasing intensity of local bias may be due to increased investor sophistication and improved dissemination of information over time.

3.3.2 Cross-Section of Funds and Geographic Equity Preference

To further analyze the local bias phenomenon, we examine funds most highly correlated with the preference for local investment. For ease of comparison and brevity, we report only the $\bar{z}$ local bias measure.

3.3.2.1 Investment Objective and Local Bias

Table 3.2 reports the $\bar{z}$ local bias measure for funds grouped by their self-declared investment objective.\(^\text{18}\) Since the set of securities available for investment by aggressive growth fund managers may be vastly different from that of a balanced fund manager, for instance, we redefine the null hypothesis for each fund objective

\(^{18}\)CDA Spectrum contains nine possible fund objectives: Aggressive Growth, Growth, Growth and Income, Balanced, International, Municipal Bond, Bond and Preferred, Metals, and Unclassified funds. We grouped funds into one of six investment objective categories by combining municipal bond, bond and preferred, metals, and unclassified funds into our ‘Other’ fund category, because of their small numbers and small percentage of equity holdings.
to include only those securities being held by funds with the same investment objective. For example, if fund $i$ is an aggressive growth fund, we compute $d_i$ as before in equation (3.2), but now compute our null distance measure, $\tilde{d}_i^v$, from equation (3.4), by replacing $\Psi_j$ with $\Psi_j^{ag} = \frac{\sum_{i=1}^{F_{ag}} H_{ij} \cdot P_j}{\sum_{j=1}^{N} \sum_{i=1}^{F_{ag}} H_{ij} \cdot P_j}$, where $F_{ag}$ is the total number of aggressive growth funds ($= 112$). Thus, the null distance measure for aggressive growth fund managers is the average distance all aggressive growth funds are from their holdings. Therefore, deviations from this measure for fund $i$ represent deviations from the average aggressive growth fund that are related to distance.

The same procedure of redefining the benchmark distance measure, $\tilde{d}_i^v$, is repeated for each of the other five investment objective categories. The results of these tests presented in Table 3.2 show that only the Growth and Other fund objective categories exhibit a significant preference at the 5% level for local equities, while Balanced funds exhibit a marginal preference for local stocks. The negative measure or dispreference for local stocks exhibited by aggressive growth fund managers is an anomaly we do not have an explanation for. However, these funds have the highest turnover, which may be consistent with turnover being greatest in distant assets. an empirical anomaly documented in the international literature.

To study international fund manager motivation, we now include international funds in the analysis for the first time. Since international fund managers presumably seek diversification benefits as their primary objective, such funds should be less likely to exhibit a geographic preference, even in their domestic holdings. As Table 3.2 demonstrates, the 90 international funds in our database exhibit no significant local bias in their holdings. Therefore, the presence of local
bias in other funds may indicate a desire to forgo additional diversification for some local investment benefit. Typically, international funds track global markets for diversification purposes, and likely are concerned with country exposure rather than individual stock selection. In addition, the greater number of securities one has to follow and the problems of dealing with national idiosyncracies (such as accounting standards, tax laws, and language barriers) make the pursuit of firm-specific information more difficult. Consequently, these funds tend to be more passive in nature. If information is the motive behind local investing, then passive funds, such as international funds, should exhibit very little local bias, a pattern consistent with our findings.

3.3.2.2 Fund Size and Local Bias

In addition to reporting the $\bar{\kappa}$ measure for funds grouped by size (total asset value), number of holdings (taken at the beginning of each quarter), age, and metropolitan area, we also calculate local bias measures and test statistics every quarter, and report their time-series averages in Table 3.3. Thus, two sets of values are reported. The first calculates local bias for each fund over time, grouping funds into categories based on their characteristics at that time (i.e., end of 1994), and then computes the cross-sectional average of each group. These statistics reflect our 'cross-sectional averages of the time-series' measures for funds, reported in Panel A. The second measure groups funds at each point in time (i.e., every quarter), computing the cross-sectional local bias measures and associated test statistics for that quarter. These values are then averaged over time. These are the 'time-series averages of the cross-section' of quarterly local bias measures and test statistics, reported in Panel B. The second procedure enables funds to move across fund categories over time and places relatively greater weight on earlier
time periods (by weighting each quarter equally).\textsuperscript{19}

Local bias appears to be inversely related to fund size, a result consistent with our information hypothesis. Large funds have greater access to information through better resources, larger number of analysts employed, etc., and have a greater ability to spread out the costs of information acquisition. In addition, large funds typically belong to large management families (such as Fidelity or Schwab), who have branch offices and money managers dispersed across the country, and who can consolidate and disseminate local information. Therefore, frictions associated with distance should have a greater impact on small fund managers if information is responsible for local preferences. In addition, since passive funds are typically larger on average, the relatively small preference for local assets exhibited by these funds is also consistent with an information motive for local bias.

Furthermore, the preference for local equities decreases almost monotonically with the number of stocks held, consistent with the information hypothesis. Funds holding a large number of stocks are probably passive portfolios, which do not seek information, and therefore do not exhibit local bias,\textsuperscript{20} while active funds, with typically fewer holdings, seek and monitor information, and thus restrict themselves to a manageable number of stocks they can follow closely. Therefore, small holdings categories are probably active funds that strongly bias their holdings locally as a consequence of seeking information.

\textsuperscript{19}It should be noted that the serial correlation problem discussed in section 2 will not apply here, as we are not computing time-series test statistics, but merely reporting the average cross-sectional test statistics of each quarter.

\textsuperscript{20}In addition, large funds may simply run out of local stocks, forcing them to hold distant assets.
3.3.2.3 Age of Fund and Local Bias

Table 3.3 shows little relation between a fund’s age (how long it has been on the CDA Spectrum tapes) and its degree of local bias according to the cross-sectional averages. However, the time-series averages demonstrate a much stronger local bias for younger funds. This pattern may represent an attempt by young fund managers to differentiate themselves from established funds by seeking a competitive advantage in local stocks. However, older funds tend to be larger on average, which may confound the results.

3.3.2.4 Metropolitan Location and Local Bias

Finally, we divide our sample of funds into managers based in small and large cities.\textsuperscript{21} As Table 3.3 demonstrates, funds located in small cities exhibit almost twice the local bias as those in large cities. Furthermore, this result may be understated. since a fund manager from just outside of Chicago, for instance, may be actively involved with the Chicago investment community. Thus, we redefine our small city sample by excluding funds located within 200 km of any of the large cities. This sample of ‘very small cities’ exhibits an even stronger local bias. This finding is consistent with an information theory of local bias, since ties between fund managers and the investment community are likely to be closer and more important in small, remote cities.

The results are particularly interesting given that large cities contain most of the equities available for investment. In other words, a New York fund man-

\textsuperscript{21}We chose 20 of the largest cities in the U.S. for our sample of large cities, and grouped funds from any other city besides these into our ‘small cities’ category. The large cities are: New York, Los Angeles, Chicago, Dallas, Atlanta, Minnesota, Washington D.C., Boston, Philadelphia, San Francisco, Seattle, Denver, Miami, St. Louis, Baltimore, Indianapolis, Detroit, Phoenix, San Diego, and Houston.
ager can likely replicate the market portfolio by investing solely in New York stocks. Thus, from a diversification standpoint, the degree of local bias should be strongest in large cities. Fund managers from small cities, on the other hand, severely limit their investment choices by investing locally, as a portfolio of local stocks from these areas is unlikely to comprise a significant or representative portion of the economy. Therefore, fund managers in small cities forgo substantial diversification benefits by investing close to home. In Section 5, we show that this loss of diversification is tempered by the higher returns from investing locally.

3.4 Relation Between Distance and Fund Turnover

If fund managers receive better information about local firms, then turnover (i.e., the frequency or amount of trade) in these investments may differ significantly from non-local investments. For instance, empirical results from the international literature find that foreign holdings are turned over at a significantly higher rate than domestic holdings (see for example, Tesar and Werner (1995), Cooper and Kaplanis (1994), and Kang and Stulz (1995)), which may be consistent with information asymmetries as shown by Brennan and Cao (1996) and Coval (1996).

To test the relation between fund turnover and distance, we compute a similar covariance measure to that of equation (3.7) by replacing \( \theta_{ij} \) with \( \tau_{ij} = \frac{1}{H_{ij,t}} - \frac{H_{ij,t-1}}{H_{ij,t-1}} \), where \( H_{ij,t} \) is fund \( i \)'s holding of stock \( j \) at time \( t \).

\[
\bar{\bar{\tau}}_{\text{turn}} = \frac{1}{F_T} \sum_{i=1}^{F} \sum_{t=1}^{T} \sum_{j=1}^{N} (\tau_{ij,t} - \bar{\tau}_{j,t})(\bar{d}_{j,t} - d_{ij,t}).
\]  

(3.8)

This statistic measures the covariance between a fund's absolute turnover (amount of trade) and distance deviations.\(^{22}\) As shown in Table 3.4, although

\(^{22}\)A fund's absolute turnover deviation is the difference between the absolute amount the fund trades in security \( j \) and the average absolute amount all funds trade in security \( j \) at each
the covariance measures are positive with and without New York funds and in the pre and post sample periods. none of the measures are statistically significant. Thus, there appears to be no significant relation between turnover and distance. This result is contrary to those found in the international investment literature, and warrants further investigation that is beyond the scope of this paper.

However, the international empirical literature does not control for factors unrelated to geography that might influence trade. For instance, if distant investors hold large stocks on average (as documented by Kang and Stulz (1995) and Coval and Moskowitz (1997a)), and large stocks can be traded more frequently and cheaply than small stocks, then it is not surprising that turnover in distant investments is larger than in local ones. While this may not explain the entire difference between domestic and foreign turnover rates, such confounding influences may have a more pronounced effect in a domestic setting. Thus, higher turnover in local stocks could exist due to managers actively trading in these stocks as they receive information, while remaining passive (buy and hold) in their distant holdings. However, this may be mitigated by the fact that uninformed/distant investors are less willing to trade with local/informed investors, due to the adverse selection problem they face. increasing the bid-ask spreads for these stocks and reducing their liquidity. This may not only cause local investors to hold a disproportionate share of local firms, but also may make trading such stocks more difficult. As a result, investors may end up holding small, illiquid, local stocks, and large, liquid, non-local stocks. Therefore, local stocks may be traded less frequently than distant ones for non-geographic related reasons, thus counteracting any information effects.\textsuperscript{23}

\textsuperscript{23} Although not reported for brevity, when we control for non-geographic factors affecting turnover (i.e., bid-ask spreads, dollar trading volume, and price), we are left with the geographic
3.5 Effect of Local Bias on Fund Performance

3.5.1 Performance Results of Local versus Distant Holdings

To test whether local holdings outperform distant holdings, we divide each fund into a local portion (holdings located over 95% closer than the market ($< \bar{d}_i/20$), and a distant portion (holdings located beyond this distance). Monthly returns on each of these two portions are then computed using returns from the daily CRSP files. The return differential represents the relative performance of these two sets of holdings for each fund. After computing the return differential, we average the difference cross-sectionally over all funds to obtain an average measure of the degree to which local holdings outperform distant holdings at each point in time. The average performance measure is calculated monthly from September 30, 1980 to December 31, 1994, and then averaged over the 15-year period. Since portfolio weights are reported quarterly, we maintain the weights from the end of each quarter over the subsequent three months. but allow the returns to change monthly. Formally, we compute this measure as:

$$P^* = \frac{1}{3T} \sum_{z=1}^{3} \sum_{t=1}^{T} \sum_{i=1}^{F_t} \frac{T_A V_i}{T_A V_i} \left[ \sum_{j=1}^{L_{i,t}} \omega_{ij,t} \cdot R_{j,t+z} - \sum_{j=1}^{D_{i,t}} \delta_{ij,t} \cdot R_{j,t+z} \right], \quad (3.9)$$

where $T = 58$ is the number of quarters, $F_t$ is the number of funds at time $t$.

component of trade, which is consistently smaller for local holdings, in accordance with the international empirical evidence.

24We assume portfolio weights for funds are constant over intermediate months between quarterly reports. For example, the weights at the end of September 1980 (the first observation) are multiplied by monthly returns for October, November, and December, 1980. The portfolio weights are then updated based on reports at the end of December 1980, and multiplied by returns for January, February, and March 1981, and so on.

25We value weight both the funds and holdings in computing our performance measure to avoid assigning too much weight to funds holding only one or two stocks locally, and to avoid biasing toward small stocks.
\( \omega_{ij,t} \) is the weight fund \( i \) places on stock \( j \) at time \( t \) in the local portion of fund \( i \). \( \delta_{ij,t} \) is the weight fund \( i \) places on stock \( j \) in the distant portion of fund \( i \). and \( L_i \) and \( D_i \) represent the number of local holdings and distant holdings of fund \( i \), respectively. Thus,

\[
\sum_{j=1}^{L_{i,t}} \omega_{ij,t} = 1, \text{ and } \sum_{j=1}^{D_{i,t}} \delta_{ij,t} = 1. \forall t.
\]

This performance measure represents a feasible investment strategy that is long all of the local holdings of every fund and short all of the distant holdings. Because 'local' is relative to a particular fund manager, local and distant portfolios will be geographically diversified and contain many of the same stocks. The average raw return difference of local and distant holdings over the 15-year period is reported in Table 3.5.

However, local and distant stocks may have very different characteristics and very different risks. Indeed, Coval and Moskowitz (1997a) find that local holdings tend to be in small and highly-levered firms. Therefore, we adjust for the relative risk of each portion by recomputing \( P^* \) using returns adjusted for size, book-to-market, and momentum investment strategies obtained from Daniel, Grinblatt, Titman, and Wermers (1997) (hereafter DGTW).\(^{26}\) The results, as well as the test statistic that \( P^* > 0 \), are reported in Table 3.5. The local portfolio significantly outperforms the distant portfolio over the entire sample period by 2.7% per year on a risk-adjusted basis.\(^{27}\) Furthermore, this performance is contained largely in

\(^{26}\)DGTW match a stock with a diversified portfolio of securities with similar size, book-to-market, and momentum characteristics. These three characteristics have been shown to be the best predictors of future stock returns (see Fama and French (1992, 1996), Jegadeesh and Titman (1993), and Daniel and Titman (1997)), and represent feasible investment strategies investors could pursue with no knowledge of firm-specific information. Thus, we use these strategies as a benchmark for which fund managers are compared, and leave risk interpretations to the reader. The adjusted return \( \bar{R}_{jt} \) is the return on the stock minus the return on the matched portfolio (i.e., \( \bar{R}_{jt} - \bar{R}_{jt-1}^{L} \), where the latter is the month \( t \) return of the matched characteristic-based portfolio for stock \( j \)). For a description of the methodology and construction of their characteristic-based performance measure, see DGTW (1997).

\(^{27}\)An equal weighted \( P^* \) measure across funds produces a performance difference of 3.0% per
the first half of our sample, a finding consistent with the fact that the majority of the local bias is also concentrated in the first half of the data. If agency or psychology are at the core of the local bias phenomenon, then we would not expect local investments to outperform non-local investments on average. Such a finding is consistent, however, with information as the motivation for local investment. Thus, it appears that mutual fund managers receive superior information about firms located near them, and exploit their informational advantage by actively trading in these stocks.

Finally, for an alternate measure of performance, we regress the DGTW-adjusted quarterly returns of every mutual fund holding on the degree of local bias of that holding. Specifically, we identify the degree of local bias for stock $j$ held by fund $i$ as the distance between them, subtracted from the average distance between the fund and all of its holdings, divided by the average distance fund $i$ is from its holdings, or $\frac{d_i - d_{ij}}{d_i}$, where $d_i$ is defined as in equation (3.2). Since the regression pools all funds and their holdings, a stock's return may appear as a dependent variable several times. To account for cross-sectional dependence, we run the regression at each point in time, and average the coefficients over time in the style of Fama and MacBeth (1973). To account for serial dependence in fund holdings when computing standard errors of our time-series of coefficient estimates, we employ a Newey and West (1987) adjustment on the time-series of regression coefficient standard errors using 4, 8, and 12 quarter lags. The results are presented in Table 3.6. As shown in the table, there is a significant positive relation between the abnormal return on a holding and the proximity of that holding. Thus, fund managers seem to reap higher abnormal returns from their local investments, consistent with our conjecture that they have an informational

year with a T-stat of 2.14. and the $P^*$ for local defined within 75 km is 2.9% per year with a T-stat of 2.12.
advantage in local stocks.

3.5.2 Fund Characteristics and Local Performance

Grouping funds by investment objective, Table 3.5 reports that both Growth and Growth & Income funds exhibit strong positive performance in their local holdings (over their distant holdings), while Aggressive Growth and Balanced funds do not. This finding provides additional evidence for our information hypothesis, since Growth funds exhibit a significant local equity preference, while Aggressive Growth funds do not.

Grouping funds by size (total asset value) the results indicate that large funds do not have a significant local advantage, consistent with the earlier result that these funds exhibited no preference for local assets. and supporting our claim that large funds are more passive in nature. However, for the second smallest size category, which exhibited a significant local bias, the local investment gain is substantial.\(^{28}\)

Using number of holdings as another measure for fund size demonstrates that small funds do indeed reap significant gains from their local investments. This result offers the most compelling support for our information theory of local bias. Funds following few stocks likely follow them closely, attempting to obtain an informational advantage. Conversely, funds following a large number of stocks employ passive strategies. As demonstrated in Table 3.3, the number of stocks a fund holds is inversely related to the degree of local investment the fund is pursuing. Thus, the strong performance difference of local stocks over distant

\(^{28}\) Although the smallest funds bias their holdings strongly toward local stocks, the insignificance of their local performance indicates that either the information they receive is noisy or the local bias is not information-based. However, the smallest total asset value group may contain funds in poor financial health or persistent underperformers, which may confound performance results.
stocks for the three smallest number of holdings quintiles reaffirms the idea that
these fund managers are active investors who pursue and obtain information from
local firms, enabling them to reap higher returns in local holdings. In fact, the
gains from their local investments over their distant ones is astounding, ranging
from 4.2% to 7.2% greater abnormal returns per year. The two largest number of
holdings quintiles do not exhibit any performance differential, consistent with the
idea that these fund managers are passive investors, and consistent with previous
results that they exhibit little or no local bias. Overall, these results indicate
strongly that information asymmetries drive the local bias phenomenon.

Finally, the gains to investing locally do not seem to be prevalent among fund
managers from large cities. This fact may be due to increased competition in
large cities for informational rents, or looser investment community ties within
large cities. On the other hand, fund managers from small cities obtain significant
positive abnormal returns from investing locally. These results are consistent with
the greater degree of local investment exhibited by small city fund managers in
Table 3.3, and thus strengthens the validity of an information-based explanation
for local equity preference.

3.6 Relation Between Local Bias, Information, and the
Cross-Section of Expected Returns

Since Coval and Moskowitz (1997a) document a relation between degree of lo-
cal bias and firm characteristics (i.e., size and leverage) known to be related to
returns, we shift the focus of the paper from the fund manager perspective to
that of an individual security. Specifically, we examine local bias from an indi-
vidual stock perspective in order to identify a possible link between the local bias
phenomenon and several well-documented asset pricing anomalies.

3.6.1 Zero-Cost Portfolio Comparisons

We begin by identifying stocks held predominantly by local fund managers, defining the local ownership of a stock as the difference in the distance between the firm's headquarters and all mutual fund managers, and the distance between its headquarters and those fund managers investing substantially in the firm. This difference is scaled by the standard deviation of the distance of the firm from every mutual fund, and then divided by the square root of the number of funds investing in the stock. Thus, the local ownership measure for a stock is similar to a test statistic for the null hypothesis that funds investing in a particular stock are no closer to that stock than the average fund that could invest in the stock. Specifically,

\[ LO_j = \frac{d_j^m - d_j}{\sigma_j^m / \sqrt{n_j}} \]  

(3.10)

where \( d_j^m \) is the average distance a stock is from all mutual funds, \( \sigma_j^m \) is the standard deviation of that distance, and \( d_j \) is the average distance a stock is from the \( n_j \) funds that hold it, where each distance is weighted by the dollar amount a fund invests in stock \( j \).\(^{29}\) Thus, if a stock is held by the market of funds, it will have a LO measure of zero, and if it is held predominantly by local funds, then LO will be significantly positive.

A simple investment strategy is to buy stocks which are significantly held locally, and short those which are not. Since our measure is similar to a test statistic, we create a portfolio that is long stocks with a local ownership measure (LO) that meets at least a 10% significance level (i.e., \( LO > 1.65 \)), and short

\(^{29}\)Both equal-weighted and value-weighted (by dollar amount of holdings) LO measures were computed, both with very similar results. We report the results for the value-weighted measure.
those which do not meet this criterion.\textsuperscript{30} This portfolio exploits the informational content of mutual fund managers without acquiring or monitoring that information, yet obtaining a geographically-diversified portfolio. Figure 3.3 plots the cumulative returns/profits of this self-financing portfolio (which we call Local Minus Distant - LMD) over time. It also compares these returns to those of three other zero-cost portfolios: a High Minus Low book-to-market equity (HML) portfolio, a Small Minus Big market value of equity (SMB) portfolio, a high minus low prior year momentum (PR1YR) portfolio, and a portfolio that is long the value-weighted market funded by the risk-free rate (MKT-Rf).\textsuperscript{31} Also reported is the test statistic that the average return over the period is positive for each of the zero-cost portfolios. Having a smooth, upward trend, the LMD portfolio appears to be less risky than the other portfolios, and has the second largest test statistic (and therefore Sharpe ratio) over the sample period (second only to the PR1YR momentum portfolio).

Creating the same LMD portfolio adjusted for size, book-to-market, and momentum attributes using DGTW-adjusted returns, we see in Figure 3.4 that abnormal returns are still present. In fact, the cumulative return on this portfolio is striking, given that we are adjusting for factors known to explain a sizeable portion of the cross-sectional variation in returns. As expected, cumulative DGTW-adjusted returns of the other portfolios are near zero, since we are adjusting for returns related to the specific characteristics these portfolios are formed on. Thus, the abnormal return on our portfolio suggests either a substantial profit opportunity or an unexplained component of risk.

\textsuperscript{30}Using a cut-off level of 1.96, corresponding to a 5\% significance level, we get even stronger results. Using a cut-off level of less than 10\% significance, produces slightly less positive results, and the lower the significance level, the smaller the abnormal returns generated from the portfolio. This suggests that returns are related to degree of local investment.

\textsuperscript{31}The market is the CRSP value-weighted portfolio and the risk-free rate is the three month Treasury Bill rate.
3.6.2 Expected Return Prediction of Local Ownership

To test whether the degree of local investment in a stock captures cross-sectional variation in asset returns, we examine the risk-adjusted returns of five portfolios, sorted by degree of local ownership. The average characteristics of these portfolios are reported in Table 3.8, which shows very little difference between high and low \textit{LO} portfolios based on size and momentum, and a negative relation between local ownership and book-to-market equity, suggesting, if anything, that returns should be higher in low \textit{LO} stocks.

We use two return adjustment procedures. The first is the DGTW matched portfolio method, and the second is similar to Carhart (1997), where we regress each portfolio’s excess unadjusted return on the HML, SMB, MKT-Rf, and PR1YR returns, using the intercept from these regressions as our measure of performance. Specifically, we employ the following measure of abnormal return:

\[ \tilde{R}_{pt} - R_{ft} = \alpha_p + \beta_{p,1}\tilde{R}_{HML,t} + \beta_{p,2}\tilde{R}_{SMB,t} + \beta_{p,3}\tilde{R}_{MKT-Rf,t} + \beta_{p,4}\tilde{R}_{PR1YR,t} + \epsilon_{p,t} \]

where \( \tilde{R}_{pt} \) is the return on each of the local investment sorted portfolios.

Table 3.7 displays the equal and value-weighted averages of monthly raw and DGTW-adjusted returns of the five portfolios as well as the intercept from a regression of the portfolio unadjusted returns on the HML, SMB, market excess, and PR1YR returns. In all cases, the portfolio of highest local investment performs significantly well, with an average annual risk-adjusted return of approximately 5%. In addition, both the returns and intercept value increase with the degree of local investment, exemplified by the significance in returns of the highest minus lowest locally owned portfolios (LO5-LO1).

As a check on the DGTW return adjustment method, we regress the value-weighted DGTW-adjusted LO5-LO1 portfolio on the Carhart factors. The results
in Table 3.9 show clearly that the significance of this zero-cost portfolio is not due to improper return adjustment from the DGTW method. Indeed, none of the Carhart factors are significant in explaining the variation in the LO5-LO1 returns, yet the intercept remains highly significant. Thus, local ownership appears to explain cross-sectional variation in returns, even after controlling for factors known to explain a significant portion of this variation.

Furthermore, the results in Table 3.9 indicate that the standard deviations of the portfolio returns increase with the degree of local investment. For raw and DGTW-adjusted returns this makes sense if locally biased stocks are more risky (i.e. smaller, higher B/M, higher leverage, etc.), since the DGTW procedure would entail subtracting increasingly riskier portfolio returns from riskier raw returns. However, for the Carhart method of risk-adjustment, the increased riskiness of the LO sorted portfolios is interesting. With diversified portfolios, the regression residuals from the Carhart regressions should have approximately the same standard deviation. However, the results indicate that degree of local ownership is positively related to risk. Thus, either a missing component of risk is not being accounted for, or grouping by the LO measure creates idiosyncratic risk groupings. The former explanation supports the hypothesis that local ownership proxies for a missing risk component. The latter explanation suggests fund managers receive local information, and exploit that informational advantage in stocks with the greatest idiosyncratic risk. Further analysis of these two components of risk is beyond the scope of this paper, but is an interesting avenue for further research.

---

32 These portfolios contain as few as 14 stocks and as many as several hundred.
3.6.2.1 Individual Security Cross-Sectional Regressions

Finally, we employ individual security Fama-MacBeth (1973) regressions of returns on the characteristics of size, book-to-market equity, momentum, and local ownership. Each month the raw returns of individual securities are regressed on the non-risk stock characteristics of size, book-to-market, LO, and previous 12 month cumulative returns (from \( t - 12 \) to \( t - 1 \)). We define size as the natural logarithm of market value of equity at the end of the previous year (we also use the logarithm of the number of employees as a non-market based measure for size, following Berk’s (1995) critique of market-based proxies for size). Additionally, book-to-market is defined as the natural logarithm of the book-to-market equity ratio at the end of the previous year (or most recent value before that). LO is computed as in equation (3.10), where any negative LO values are set to zero, and momentum is the previous one year return on the stock. Following Brennan, Chordia, and Subrahmanyam (1997), we express all firm characteristics as deviations from their cross-sectional means each month. Thus our regression equation is:

\[
\tilde{R}_{jt} - R_{ft} = \alpha_t + \sum_{k=1}^{4} \gamma_{kt} Z_{kt} + \tilde{e}_{jt},
\]

where \( \tilde{R}_{jt} \) is the return on security \( j \) at time \( t \), \( R_{ft} \) is the three-month Treasury bill rate at time \( t \), \( Z_{kt} \) are the four firm characteristics, and \( \gamma_{kt} \) are the coefficients of the characteristics from the regression. Coefficients are averaged over time in the style of Fama and MacBeth (1973) and reported in Table 3.10 along with their T-stats. As shown in the table, the local ownership measure appears to have

\[33\text{Once a stock is predominantly held by non-local investors, it should not matter how 'non-local' those investors are. For instance, an investor 1000 km away is probably equally as uninformed as an investor 1500 or 2000 km away. In fact, the number of negative LO stocks was small, and most of them are only being held by one or two funds, making us weary of the reliability of the measures for these stocks.}\]
marginal positive predictive power, providing further evidence of the relation between local ownership and returns.

Thus, adjusting stock returns via two methods: a characteristic-based approach (following DGTW and Daniel and Titman (1997)), and a factor-mimicking pricing equation (following Carhart (1997) and Fama and French (1993)), we find that stocks with a high degree of local ownership outperform stocks with a low degree of local ownership. In addition, regressing excess returns on the non-risk characteristics of size, book-to-market equity, momentum (prior one year return), and local ownership, we still find a strong positive relation between expected returns and local ownership. Therefore, whether we control for the loadings on factor-mimicking portfolios or for characteristics known to be related to returns, we cannot capture the difference in returns between locally owned and non-locally owned stocks. Therefore, our results are not due to other return anomalies, and strongly reaffirm the notion that the local bias phenomenon is the result of information asymmetries. In addition, these findings may provide insights into the existence of cross-sectional return anomalies discussed in the next section.

3.6.3 A Link between Local Ownership and Asset Pricing Anomalies?

Much research has focused on the relation between firm characteristics and the cross-section of expected returns.\textsuperscript{34} Furthermore, Coval and Moskowitz (1997a) document a strong relation between firm characteristics and the degree of local bias in a firm. Thus, the relation between local ownership and firm characteristics known to be related to returns, suggests an intriguing link between information asymmetries and asset pricing anomalies.

\textsuperscript{34}See Fama and French (1992) and Daniel and Titman (1997) for an excellent discussion of this literature.
The link between local ownership, a proxy for information asymmetries, and firm characteristics related to expected returns has the following economically appealing intuition. Since small firms have little analyst following, a larger holding of local and inside investors, and a more concentrated customer base, size may proxy for information asymmetries. Additionally, book-to-market may proxy for firm distress, as argued by Fama and French (1992), which makes the firm more informationally sensitive. Finally, the momentum effect of Jegadeesh and Titman (1993) may also be related to information in that high momentum stocks, with recent price surged, typically engage in value-signalling financing activities (such as stock splits and dividend increases) that are associated with positive returns. In other words, firms with the greatest degree of information asymmetry will have the greatest price reaction to these signals, and the greatest subsequent return. Thus, information asymmetries and the associated adverse selection problem may drive cross-sectional return differentials among assets. In this case, firm characteristics simply help signify the degree of asymmetry.

However, adverse selection due to information asymmetry may have both a priced and unpriced component. Typically, investors are informed about firms they are already tied to, either geographically or through employment. Therefore, to hedge human capital and regional-specific risks, they may trade off an informational advantage for diversification, incurring adverse selection costs. These costs can be minimized by investing in large, healthy firms, whose future cash flows are less sensitive to information. But why can’t both objectives be accomplished by buying futures or a market index? To the extent that indexes and futures contracts may be poor substitutes when information arrives, this strategy may be suboptimal. Certainly, further study on the systematic and non-systematic
components of adverse selection is warranted.

Finally, there are several caveats with the analysis worth noting. First, although CDA Spectrum constructs its database with no survivorship requirement, our analysis may be subject to survivorship bias due to matching ex post latitude and longitude data with prior years’ fund managers and firms. However, tests to measure the importance of survivorship indicated little, if any, bias on our sample’s returns. Secondly, since we use cartesain distance rather than an economic measure of distance such as average travel costs, our measure of local bias may be a noisy approximation for asymmetric information. In addition, we examine only one group of investors. All of our results on local bias and performance were restricted to mutual funds. At the individual level, the degree of local bias may be much stronger. Furthermore, for a complete documentation of the adverse selection component associated with a security, all investor (institutional and individual) holdings must be considered. Finally, other measures of asymmetric information, such as the extent of analyst following or analyst dispersion in forecasts of a stock could create a more precise measure of the degree of information asymmetries associated with a firm, yielding a more definitive answer to whether asset pricing anomalies are related to information asymmetries.\footnote{To test whether survivorship bias significantly influences the results, we chose 100 random stocks every month from our sample, and 100 from a survivorship-free sample from CRSP. We then formed equal and value weighted portfolios of the 100 stocks each month, risk-adjusting their returns using the DGTW method. The time-series average of the monthly portfolio returns is then computed. This procedure was repeated 50 times, and the average of the 50 time-series averages are compared from the two samples. The value-weighted portfolio difference between the two samples is virtually zero, suggesting that our sample of firms does not suffer from survivorship bias.} \footnote{Brennan, Chordia, and Subrahmanyam (1997), find that size is subsumed by trading volume, which they claim is a liquidity effect. This result is consistent with an adverse selection/neglected firm story. However, they find no relation between returns and bid-ask spread, another proxy for information asymmetry.}
3.7 Conclusion

We find that information asymmetries are primarily responsible for the geographic preference for local equities among mutual fund managers within the U.S. Evidence to support this conjecture is found in a diminishing intensity of local bias over time, and a negative relation between both local bias and fund size, and local bias and number of fund holdings. In addition, funds based in small, remote locations, exhibit the strongest local bias, and funds with characteristics associated with active (information seeking) management are more likely to bias their holdings toward local equities. Conversely, funds with characteristics associated with passive management (e.g., international funds and large funds) do not exhibit a preference for local equities.

Most convincing in support of our claim that local bias is related to information asymmetries, is the positive relation between performance and local investment, suggesting that fund managers are, in fact, receiving and acting upon superior information. Although local investment performance disappears in the latter half of our sample, funds associated with active management still exhibit greater performance in their local investments, while passive managers do not. These results further support our information conjecture, and are inconsistent with other theories (e.g., agency, psychology, etc.) for the geographic preference for equities.

At the mutual fund manager level, a study of local investment refines our understanding of how these institutions invest. For instance, the selection of assets with certain characteristics\textsuperscript{37} may depend on where the security is located relative to the investor. Initial evidence indicates fund managers hold highly levered, small, risky firms if these firms are local. However, other characteristics

\textsuperscript{37}See Falkenstein (1996).

152
may be equally important to managers and may depend on geography. Likewise, the use of momentum strategies and herding behavior may be very different for local and non-local investments. These may be interesting avenues for further work on the influence of geography on institutional investors' portfolio choices.

Finally, at the individual security level, the degree to which a stock is held by local investors appears to be positively related to expected returns. Thus, information may play a key role in explaining well-documented asset pricing puzzles. Additional analysis of the systematic and unsystematic components of information asymmetries, as well as a more complete and descriptive measure of information, are additional research directions worth pursuing.
Figure 3.1: Local Bias Over Time: Difference Between Market and Actual Average Distances

Figure 3.2: City Size and Local Bias
Table 3.1: Results of Tests for Local Bias Among U.S. Mutual Funds

The difference in distances $\overline{z}$, as well as the covariance measure, $z_{cov}$, averaged individually for each fund over time (from 1980:3 - 1994:4) and then averaged cross-sectionally across all funds are reported below. The same tests are run excluding all New York-based fund managers. $T_W$ is the test statistic for the non-parametric Wilcoxon signed ranks test.

<table>
<thead>
<tr>
<th></th>
<th># Funds</th>
<th>$\overline{z}$ (km)</th>
<th>T-stat</th>
<th>$T_W$</th>
<th>$z_{cov}$</th>
<th>T-stat</th>
<th>$T_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>1278</td>
<td>72.98**</td>
<td>5.42</td>
<td>6.23</td>
<td>24.84**</td>
<td>1.99</td>
<td>2.41</td>
</tr>
<tr>
<td>Overall (w/o NY)</td>
<td>1014</td>
<td>93.94**</td>
<td>6.80</td>
<td>7.46</td>
<td>2.11</td>
<td>0.14</td>
<td>0.60</td>
</tr>
<tr>
<td>Pre-1987 (1980-1986)</td>
<td>366</td>
<td>149.69**</td>
<td>5.81</td>
<td>5.84</td>
<td>33.35</td>
<td>1.50</td>
<td>1.69</td>
</tr>
<tr>
<td>1980</td>
<td>118</td>
<td>161.12**</td>
<td>3.64</td>
<td>3.20</td>
<td>28.24</td>
<td>0.91</td>
<td>1.04</td>
</tr>
<tr>
<td>1981</td>
<td>125</td>
<td>118.47**</td>
<td>2.62</td>
<td>2.27</td>
<td>13.40</td>
<td>0.41</td>
<td>0.91</td>
</tr>
<tr>
<td>1982</td>
<td>137</td>
<td>90.94**</td>
<td>2.30</td>
<td>1.86</td>
<td>2.10</td>
<td>0.06</td>
<td>0.16</td>
</tr>
<tr>
<td>1983</td>
<td>165</td>
<td>100.74**</td>
<td>2.77</td>
<td>2.14</td>
<td>-33.20</td>
<td>-1.01</td>
<td>-0.93</td>
</tr>
<tr>
<td>1984</td>
<td>188</td>
<td>70.24**</td>
<td>2.36</td>
<td>1.82</td>
<td>-52.31**</td>
<td>-1.66</td>
<td>-1.31</td>
</tr>
<tr>
<td>1985</td>
<td>306</td>
<td>223.29**</td>
<td>7.43</td>
<td>6.94</td>
<td>2.99</td>
<td>0.13</td>
<td>0.86</td>
</tr>
<tr>
<td>1986</td>
<td>366</td>
<td>149.69**</td>
<td>5.81</td>
<td>5.84</td>
<td>33.35</td>
<td>1.50</td>
<td>1.69</td>
</tr>
<tr>
<td>1987</td>
<td>404</td>
<td>137.05**</td>
<td>5.45</td>
<td>5.79</td>
<td>54.86**</td>
<td>2.39</td>
<td>2.61</td>
</tr>
<tr>
<td>1988</td>
<td>437</td>
<td>96.27**</td>
<td>3.85</td>
<td>4.67</td>
<td>54.67**</td>
<td>2.41</td>
<td>2.68</td>
</tr>
<tr>
<td>1989</td>
<td>475</td>
<td>76.60**</td>
<td>3.26</td>
<td>4.14</td>
<td>50.61**</td>
<td>2.15</td>
<td>2.82</td>
</tr>
<tr>
<td>1990</td>
<td>590</td>
<td>39.56*</td>
<td>1.75</td>
<td>2.96</td>
<td>57.35**</td>
<td>2.86</td>
<td>3.22</td>
</tr>
<tr>
<td>1991</td>
<td>708</td>
<td>5.60</td>
<td>0.26</td>
<td>1.71</td>
<td>37.46*</td>
<td>1.93</td>
<td>1.97</td>
</tr>
<tr>
<td>1992</td>
<td>827</td>
<td>36.09**</td>
<td>2.12</td>
<td>3.12</td>
<td>36.56**</td>
<td>2.51</td>
<td>2.75</td>
</tr>
<tr>
<td>1993</td>
<td>1064</td>
<td>141.49**</td>
<td>7.78</td>
<td>7.32</td>
<td>18.82</td>
<td>1.49</td>
<td>1.72</td>
</tr>
<tr>
<td>1994</td>
<td>1278</td>
<td>72.98**</td>
<td>5.42</td>
<td>6.23</td>
<td>24.84**</td>
<td>1.99</td>
<td>2.41</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>479.20</td>
<td>101.34**</td>
<td>3.79</td>
<td>4.00</td>
<td>21.98</td>
<td>1.20</td>
<td>1.51</td>
</tr>
</tbody>
</table>

*indicates significant at the 10% level. **indicates significant at the 5% level.
Table 3.2: Results of Tests for Local Bias Across Fund Investment Objective

The difference in distances $F_c$ averaged individually for each fund over time (from 1980:3 - 1994:4) and then averaged cross-sectionally across all funds are reported below for fund objective, where the null distance measure is computed using only those stocks held by funds with that investment objective. $T_{W}$ is the test statistic for the non-parametric Wilcoxon signed ranks test.

<table>
<thead>
<tr>
<th>Fund Objective</th>
<th># Funds</th>
<th>$\delta$ (km)</th>
<th>T-stat</th>
<th>$T_{W}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive Growth</td>
<td>112</td>
<td>-57.29**</td>
<td>-2.27</td>
<td>-1.71</td>
</tr>
<tr>
<td>Growth</td>
<td>559</td>
<td>38.84**</td>
<td>2.63</td>
<td>2.48</td>
</tr>
<tr>
<td>Growth &amp; Income</td>
<td>188</td>
<td>25.67</td>
<td>0.84</td>
<td>0.77</td>
</tr>
<tr>
<td>Balanced</td>
<td>87</td>
<td>81.18*</td>
<td>1.77</td>
<td>2.54</td>
</tr>
<tr>
<td>International</td>
<td>90</td>
<td>48.30</td>
<td>1.13</td>
<td>1.02</td>
</tr>
<tr>
<td>Other$^a$</td>
<td>332</td>
<td>193.91**</td>
<td>5.69</td>
<td>6.96</td>
</tr>
</tbody>
</table>

$^a$ Other funds include municipal bond, bond and preferred, metals, and unclassified sector funds. These were grouped into one category because of their small numbers and small percentage of equity holdings.

*indicates significant at the 10% level. **indicates significant at the 5% level.

Figure 3.3: Cumulative Profits on Zero Cost Portfolios Over Time (1980:3 - 1994:4)
Table 3.3: Results of Tests for Local Bias Across Fund Types

The difference in distances \( \bar{x} \), averaged individually for each fund over time (from 1980:3 - 1994:4) and then averaged cross-sectionally across all funds are reported below in Panel A for fund size, number of holdings, age, and metropolitan area categories. In addition, the time-series averages of the cross-sectional \( \bar{x} \) measures and their test statistics at the end of every quarter are reported in Panel B. \( T_W \) is the test statistic for the non-parametric Wilcoxon signed ranks test.

<table>
<thead>
<tr>
<th>Fund Size</th>
<th>Panel A</th>
<th>Panel B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cross-Sectional Average of Time-Series Measures</td>
<td>Time-Series Average of Cross-Sectional Measures</td>
</tr>
<tr>
<td></td>
<td>( \bar{x} ) (km)</td>
<td>T-stat</td>
</tr>
<tr>
<td>Size1 (0 - $0.5mm)</td>
<td>165</td>
<td>250.28**</td>
</tr>
<tr>
<td>Size2 ($0.5mm - $5mm)</td>
<td>285</td>
<td>107.58**</td>
</tr>
<tr>
<td>Size3 ($5mm - $25mm)</td>
<td>323</td>
<td>25.31</td>
</tr>
<tr>
<td>Size4 ($25mm - $100mm)</td>
<td>255</td>
<td>39.30*</td>
</tr>
<tr>
<td>Size5 ($100mm - $7000mm)</td>
<td>250</td>
<td>12.47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. Holdings</th>
<th>Panel A</th>
<th>Panel B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. Holdings ≤ 10</td>
<td>410</td>
<td>132.46**</td>
</tr>
<tr>
<td>10 &lt; No. Holdings ≤ 20</td>
<td>241</td>
<td>90.64**</td>
</tr>
<tr>
<td>20 &lt; No. Holdings ≤ 30</td>
<td>219</td>
<td>75.82**</td>
</tr>
<tr>
<td>30 &lt; No. Holdings ≤ 50</td>
<td>231</td>
<td>23.85</td>
</tr>
<tr>
<td>50 &lt; No. Holdings</td>
<td>177</td>
<td>-28.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Panel A</th>
<th>Panel B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age ≤ 2</td>
<td>639</td>
<td>62.48**</td>
</tr>
<tr>
<td>2 &lt; Age ≤ 5</td>
<td>340</td>
<td>108.35**</td>
</tr>
<tr>
<td>5 &lt; Age ≤ 10</td>
<td>165</td>
<td>64.21**</td>
</tr>
<tr>
<td>10 &lt; Age ≤ 15</td>
<td>62</td>
<td>34.98</td>
</tr>
<tr>
<td>15 &lt; Age</td>
<td>72</td>
<td>51.99**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metropolitan Area</th>
<th>Panel A</th>
<th>Panel B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Cities</td>
<td>745</td>
<td>55.19**</td>
</tr>
<tr>
<td>Small Cities</td>
<td>542</td>
<td>103.85**</td>
</tr>
<tr>
<td>Very Small Cities</td>
<td>355</td>
<td>131.37**</td>
</tr>
</tbody>
</table>

*indicates significant at the 10% level. **indicates significant at the 5% level.
Table 3.4: Covariance Between Turnover and Distance
The covariance measure $\tilde{z}_{\text{turn}}$, averaged individually for each fund over time (from 1980:3 - 1994:4) and then averaged cross-sectionally across all funds is reported below. The same tests are run excluding all New York-based fund managers. $T_W$ is the test statistic for the non-parametric Wilcoxon signed ranks test.

<table>
<thead>
<tr>
<th></th>
<th># Funds</th>
<th>$\tilde{z}_{\text{turn}}$</th>
<th>T-stat</th>
<th>$T_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>1278</td>
<td>14295.80</td>
<td>1.42</td>
<td>1.70</td>
</tr>
<tr>
<td>Overall (w/o NY)</td>
<td>1014</td>
<td>17354.44</td>
<td>1.37</td>
<td>1.81</td>
</tr>
<tr>
<td>Pre-1987 (1980-1986)</td>
<td>366</td>
<td>514.14</td>
<td>1.09</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Figure 3.4: Cumulative Risk-Adjusted Profits on Zero Cost LMD Portfolio
Table 3.5: Performance Differentials of Local and Distant Fund Portions  
(\% Monthly Return)

Results of tests of the difference in average monthly returns between local and distant portions of a fund, averaged cross-sectionally and then over time. 'Local' is defined as any holding located within 95\% of the average market distance for the fund. Returns are adjusted for size, book-to-market, and momentum investment strategies, as described in Daniel, Grinblatt, Titman, and Wermers (1997). $T_W$ is the test statistic for the non-parametric Wilcoxon signed ranks test.

<table>
<thead>
<tr>
<th>Fund Objective</th>
<th>Raw Returns $\tilde{R}_{t,z}$</th>
<th>DGTW-Adjusted Returns $\tilde{R}_{t,z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho^*$</td>
<td>$T$-stat</td>
</tr>
<tr>
<td>Total Sample</td>
<td>0.185</td>
<td>1.31</td>
</tr>
<tr>
<td>Pre-1987</td>
<td>0.450*</td>
<td>1.82</td>
</tr>
<tr>
<td>Post-1987</td>
<td>-0.080</td>
<td>-0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fund Size</th>
<th>Raw Returns $\tilde{R}_{t,z}$</th>
<th>DGTW-Adjusted Returns $\tilde{R}_{t,z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho^*$</td>
<td>$T$-stat</td>
</tr>
<tr>
<td>Size1</td>
<td>0.002</td>
<td>0.01</td>
</tr>
<tr>
<td>Size2</td>
<td>0.300</td>
<td>1.64</td>
</tr>
<tr>
<td>Size3</td>
<td>0.094</td>
<td>0.47</td>
</tr>
<tr>
<td>Size4</td>
<td>0.154</td>
<td>0.97</td>
</tr>
<tr>
<td>Size5</td>
<td>0.191</td>
<td>1.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># Holdings</th>
<th>Raw Returns $\tilde{R}_{t,z}$</th>
<th>DGTW-Adjusted Returns $\tilde{R}_{t,z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho^*$</td>
<td>$T$-stat</td>
</tr>
<tr>
<td>Num1</td>
<td>0.603*</td>
<td>1.64</td>
</tr>
<tr>
<td>Num2</td>
<td>0.291</td>
<td>1.51</td>
</tr>
<tr>
<td>Num3</td>
<td>0.308</td>
<td>1.41</td>
</tr>
<tr>
<td>Num4</td>
<td>0.146</td>
<td>0.73</td>
</tr>
<tr>
<td>Num5</td>
<td>0.105</td>
<td>0.68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Raw Returns $\tilde{R}_{t,z}$</th>
<th>DGTW-Adjusted Returns $\tilde{R}_{t,z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho^*$</td>
<td>$T$-stat</td>
</tr>
<tr>
<td>Age1</td>
<td>0.324</td>
<td>1.39</td>
</tr>
<tr>
<td>Age2</td>
<td>0.314</td>
<td>1.36</td>
</tr>
<tr>
<td>Age3</td>
<td>-0.106</td>
<td>-0.36</td>
</tr>
<tr>
<td>Age4</td>
<td>0.340</td>
<td>1.11</td>
</tr>
<tr>
<td>Age5</td>
<td>0.033</td>
<td>0.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Large Cities</th>
<th>Raw Returns $\tilde{R}_{t,z}$</th>
<th>DGTW-Adjusted Returns $\tilde{R}_{t,z}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho^*$</td>
<td>$T$-stat</td>
</tr>
<tr>
<td>Large</td>
<td>0.147</td>
<td>0.96</td>
</tr>
<tr>
<td>Small</td>
<td>0.184</td>
<td>1.22</td>
</tr>
</tbody>
</table>

* indicates significant at the 10\% level. ** indicates significant at the 5\% level.
Table 3.6: Regression of Adjusted Returns of Fund Holdings on Degree of Local Bias

Results for the cross-sectional quarterly regression of holdings abnormal returns on the degree of local bias of that holding. Dependent variable is DGTW-adjusted return of stock $j$ held by fund $i$ for all $N$ stocks and all $F$ funds. Independent variable is the scaled distance deviation or local bias between a fund $i$ and stock $j$: \( \frac{d_{j} - \hat{d}_j}{\hat{d}_j} \). Coefficients are averaged over time in the style of Fama and MacBeth (1973). Standard errors of time-series coefficient estimates are adjusted using a Newey-West procedure for 4, 8, and 12 quarter autocorrelation lags.

<table>
<thead>
<tr>
<th># Lags</th>
<th>$\alpha$</th>
<th>T-stat</th>
<th>$\beta$</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0.845**</td>
<td>98.84</td>
<td>0.0423**</td>
<td>8.89</td>
</tr>
<tr>
<td>4 quarters</td>
<td>0.845**</td>
<td>36.37</td>
<td>0.0423**</td>
<td>3.47</td>
</tr>
<tr>
<td>8 quarters</td>
<td>0.845**</td>
<td>29.20</td>
<td>0.0423**</td>
<td>2.88</td>
</tr>
<tr>
<td>12 quarters</td>
<td>0.845**</td>
<td>28.59</td>
<td>0.0423**</td>
<td>2.80</td>
</tr>
</tbody>
</table>

** Significant at the 5% level.
Table 3.7: Performance Measures of Local Investment Sorted Portfolios

(\% Monthly Returns)

Performance results for portfolios of stocks formed on degree of local investment. The average raw and DGTW-adjusted equal and value weighted monthly returns are reported for each of the portfolios. In addition, the intercept from the regression of portfolio returns on the HML, SMB, excess market, and PR1YR returns is also reported. The standard deviation of the portfolios (\(\sigma\)) is also reported as a percentage.

<table>
<thead>
<tr>
<th>Local Investment Equal-Weighted Portfolios</th>
<th>Raw Returns</th>
<th>DGTW-Adjusted Returns</th>
<th>Regression (\alpha)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{R}_{jt})</td>
<td>T-stat</td>
<td>(\sigma)</td>
</tr>
<tr>
<td>LO1 (LO (\leq) 0)</td>
<td>1.32**</td>
<td>3.43</td>
<td>5.05</td>
</tr>
<tr>
<td>LO2 (0 &lt; LO (\leq) 0.5)</td>
<td>1.14**</td>
<td>2.88</td>
<td>5.22</td>
</tr>
<tr>
<td>LO3 (0.5 &lt; LO (\leq) 1)</td>
<td>1.18**</td>
<td>2.98</td>
<td>5.19</td>
</tr>
<tr>
<td>LO4 (1 &lt; LO (\leq) 2)</td>
<td>1.28**</td>
<td>3.28</td>
<td>5.15</td>
</tr>
<tr>
<td>LO5 (2 &lt; LO)</td>
<td>1.55**</td>
<td>3.29</td>
<td>6.20</td>
</tr>
<tr>
<td>LO5 - LO1</td>
<td>0.23</td>
<td>0.99</td>
<td>3.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Local Investment Value-Weighted Portfolios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LO1 (LO (\leq) 0)</td>
</tr>
<tr>
<td>LO2 (0 &lt; LO (\leq) 0.5)</td>
</tr>
<tr>
<td>LO3 (0.5 &lt; LO (\leq) 1)</td>
</tr>
<tr>
<td>LO4 (1 &lt; LO (\leq) 2)</td>
</tr>
<tr>
<td>LO5 (2 &lt; LO)</td>
</tr>
<tr>
<td>LO5 - LO1</td>
</tr>
</tbody>
</table>

* Significant at the 10% level. ** Significant at the 5% level.

Table 3.8: Average Characteristics of the Local Investment Sorted Portfolios

<table>
<thead>
<tr>
<th>LO1 (LO (\leq) 0)</th>
<th>Size($mm)</th>
<th>B/M</th>
<th>Momentum</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO2 (0 &lt; LO (\leq) 0.5)</td>
<td>356</td>
<td>0.80</td>
<td>0.97</td>
</tr>
<tr>
<td>LO3 (0.5 &lt; LO (\leq) 1)</td>
<td>572</td>
<td>0.64</td>
<td>1.08</td>
</tr>
<tr>
<td>LO4 (1 &lt; LO (\leq) 2)</td>
<td>445</td>
<td>0.77</td>
<td>1.06</td>
</tr>
<tr>
<td>LO5 (2 &lt; LO)</td>
<td>520</td>
<td>0.69</td>
<td>1.06</td>
</tr>
<tr>
<td>LO5 - LO1</td>
<td>376</td>
<td>0.61</td>
<td>1.06</td>
</tr>
</tbody>
</table>
Table 3.9: Regression of DGTW Adjusted LO5-LO1 Value-Weighted Portfolio on Carhart Factors

Results for the regression of the value-weighted LO5-LO1 portfolio, adjusted for DGTW's characteristic-based portfolios, on the HML, SMB, excess market, and PR1YR returns. Coefficient values are multiplied by 100 and T-stats are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>HML</th>
<th>SMB</th>
<th>MKT-Rf</th>
<th>PR1YR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO5 - LO1</td>
<td>0.576**</td>
<td>-15.17</td>
<td>-3.27</td>
<td>2.68</td>
<td>-11.57</td>
</tr>
<tr>
<td>T-stat</td>
<td>(2.07)</td>
<td>(-1.39)</td>
<td>(-0.28)</td>
<td>(0.39)</td>
<td>(-1.32)</td>
</tr>
</tbody>
</table>

* Significant at the 10% level. ** Significant at the 5% level.

Table 3.10: Individual Security Fama-MacBeth Regressions

Regression results for equation (3.10). Stock characteristics are expressed as deviations from their cross-sectional mean each month. Coefficients are computed cross-sectionally each month, and then averaged over time in the style of Fama and MacBeth (1973). All coefficients are multiplied by 100. T-stats are in parentheses.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>$\tilde{R}<em>{jt} - R</em>{f1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.65*</td>
</tr>
<tr>
<td></td>
<td>(1.80)</td>
</tr>
<tr>
<td>size</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
</tr>
<tr>
<td>employees</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>B/M</td>
<td>0.49*</td>
</tr>
<tr>
<td></td>
<td>(1.74)</td>
</tr>
<tr>
<td>LO</td>
<td>0.06**</td>
</tr>
<tr>
<td></td>
<td>(2.04)</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.62**</td>
</tr>
<tr>
<td></td>
<td>(2.59)</td>
</tr>
</tbody>
</table>

* indicates significant at the 10% level. ** indicates significant at the 5% level.
References for Chapter 3


Block, Sandra, February 28. 1997. Finding returns by investing close to home, USA Today, 8B.


Coval, Joshua D., 1997, International capital flows when investors have local information, Working paper, University of Michigan Business School, Ann Arbor, MI.


Daniel, Kent, Mark Grinblatt, Sheridan Titman. and Russell Wermers, 1997,


Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, 1993, Geographic localization of knowledge spillovers as evidenced by patent citations, Quarterly Journal of Economics 63, 577-598.


University, Columbus, OH.


