Essays in Financial Economics

Essays in Financial Economics

A dissertation presented

by

Alexander Chernyakov

to

The Department of Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

August 2016

Dissertation Advisor:
Professor John Y. Campbell

Author:
Alexander Chernyakov

Essays in Financial Economics

Abstract

This dissertation consists of three essays: Chapters 1 and 2 focus on the impact of cognitive and institutional constraints on stock market efficiency while Chapter 3 examines whether shocks to the real interest rate are a priced state variable.

Chapter 1 is titled "Commodity Inattention": attention is a scarce resource for investors that must be divided among many sources of information. The commodities market is an important source of information affecting firms that operate in the economy. Investors do not fully appreciate this relationship allowing for predictability in equity returns using commodity returns. A strategy that exploits this predictability has an alpha of 1.5% per month and no meaningful factor exposure. This effect is stronger in smaller firms, firms that tend to be ignored by their owners, firms owned by investors who ignore commodity information, firms with nuanced commodity exposure and during times of high informational burden for investors.

Chapter 2 is titled "Market Crash Risk and Slow Moving Capital": index option skew (risk reversal) is a variable commonly looked at by investors to assess market conditions. In the cross-section, value stocks and junk bonds do poorly when the price of risk reversals increases. However, investors are slow to fully incorporate this information into prices leading to significant predictability in value vs. growth stocks as well as junk vs. investment grade bonds. This predictability is economically significant and poses a challenge to strictly rational models of information processing by investors.

Chapter 3 is titled "Is Real Interest Rate Risk Priced? Theory and Empirical Evidence": we propose a model in which real interest rates respond to both expected consumption growth and time preferences. Exposures to future consumption growth and time preference interest rate shocks are both priced, however, the two types of interest rate risk have different prices. The premia for time preference risk are arbitrarily large when EIS is close to 1. Empirically, we find little evidence that interest rate risk is priced in the cross-section of stocks and bonds.

Contents

List of Tables

List of Figures

Acknowledgments

I am deeply grateful to John Campbell, Lauren Cohen, Robin Greenwood and Adi Sunderam for guidance throughout my studies. Their advice has improved my work immensely and has made my time conducting research rewarding. I particularly appreciate their patience and support through the unexpected twists and turns of doctoral studies. I am also indebted to my friends and colleagues Johnny Kang, Sam Kruger, James Lee, David Smalling, and Eric Zwick for valuable discussion and feedback on my work.

I would be remiss if I did not thank my longtime friends who have provided encouragement throughout the many stages of my studies and work. While I cannot list everyone, I am thankful to Harsh Asher, Helen Belogolova, Michael Beregovsky, Timofei Gerasimov, Andrey Grigoryev, Kenny Maneles, Akhil Shashidhar and Peng Zhao for their friendship. I am exceptionally thankful to Derek Schaeffer who encouraged me to pursue graduate studies in the first place and introduced me to economics over coffees and lunches. Finally, a sincere thank you to Carol Hertz who taught me how to write, edit and revise so many years ago.

I would also like to thank James Zeitler and the staff at Baker Library for pointing me to the right data sources, participants in the Harvard Finance Lunch, and the HBS doctoral office for administrative support.

Most importantly I am grateful beyond words to my parents who have provided unwavering support throughout my life. None of this would be possible without them.

To my parents

Chapter 1

Commodity Inattention

1.1 Introduction

Much rational asset pricing work assumes that investors are able to fully incorporate all available public information into prices. Recent theory and empirical evidence has begun to cast doubt on this assumption: the ability to incorporate all available information requires investors to devote time to researching and understanding different sources of information. Rational inattention, as pioneered by Sims (2003), posits that investors have a limited amount of attention that they must allocate across information sources. Each investor will prioritize information that is most relevant to him and easiest to acquire; conversely information that is more difficult to process or less relevant to each investor may be ignored. The commodity market is one such source of information that is important for firms: commodities serve as inputs and outputs of firms that operate in the real economy. Changes in commodity prices have a real impact on the cash flows of certain firms and industries but, as I show, investors underreact to this information.

Commodities are often examined as a separate asset class to understand their risk premia and term structures as in Fama & French (1987), Schwartz (1997), Pindyck (2001), Yang (2013), however, few studies examine how information travels from the commodity market to the firms that depend on commodities. I first examine commodity returns grouped into three sectors: Energy, Agriculture (Ag) and Metal. The returns to these three commodity sectors provide a parsimonious description of the events in the commodity market. I associate equity industries with up to three of these sectors and show that price information regarding these commodity sectors travels slowly: a strategy that goes long stocks whose associated commodity sectors increased last month and short stocks whose associated commodity sectors decreased last month earns up to 1.5% per month in risk adjusted returns without a significant exposure to the commonly used equity factors. This effect is much stronger in smaller stocks: these stocks tend to have fewer analysts and be owned by fewer sophisticated investors. Therefore, smaller stocks are often ignored relative to their larger counterparts and information diffuses to them less rapidly. For instance, Hong et al. (2000) show that momentum strategies are stronger in smaller stocks and attribute this to slower information diffusion in smaller securities relative to larger securities.

To better understand the process I am describing I provide an example of just such an underreaction to information. Crosstex Energy Inc. (XTXI) is a midstream energy company that processes and transports oil and gas from producers to consumers[1]. One of the payment models in the energy industry is the percent-of-proceeds contract in which the producer and the midstream (transportation) company split the revenue from the sale of energy to consumers, exposing both parties to fluctuations in energy prices[2]. Therefore, a higher energy price means more revenue for both companies. Figure 1.1 plots the cumulative returns to the Energy commodities sector - an equal weighted average of returns to Brent Crude, Gasoil, Heating Oil, Natural Gas, RBOB[3], and WTI - and XTXI from February 2006 to March 2006. In the first two weeks of February, the Energy Information Administration (EIA) released two bearish reports showing a buildup in energy commodities which sent the prices of these commodities lower; XTXI did not react significantly to this news. On March 10th 2006,

[1] http://www.crosstexenergy.com/

[2] http://www.investingdaily.com/11887/mlps-and-natural-gas-liquids/

[3] Note that RBOB HU denotes the time series splicing together of the Unleaded Gasoline (HU) contract and the Reformulated Blendstock for Oxygenate Blending (RBOB) as HU was phased out from trading. WTI denotes the West Texas Intermediate crude oil contract.

XTXI reported its Q4 2005 and fiscal year 2005 earnings. Barry Davis, the CEO, described the announced information by saying: "We had a great fourth quarter and an outstanding year in 2005." Once again the stock does not have a significant reaction; however, on March 20th the company held an analyst meeting to discuss 2006 prospects and the stock took a significant hit. Revenue in 2006 was dependent on energy prices in 2006 which dropped by approximately 10% a month earlier. This kind of slow incorporation of information from the commodity market to the equity market will be explored in this study.

There are four potential channels through which investors could be ignoring pertinent information: they could be ignoring a particular stock because that stock is unimportant to them, they could be ignoring information regarding a stock's associated commodity sector because they are ignoring commodities, they could misunderstand the impact commodities have on a particular stock, or they may be overwhelmed with a large amount of idiosyncratic information being released by companies in a particular time period. Using mutual fund holdings data I show that portfolio managers pay attention to stocks in their portfolios with the most volatile P&L: these stocks are efficient with respect to commodity market information. Conversely, stocks that do not have much P&L variance are ignored by their owners and are inefficient. In other words, stocks that are viewed as risky by managers attract a significant amount of attention. Second, stocks whose owners hold portfolios that are not significantly exposed to the stock's associated commodity sector also underreact significantly to commodity news while stocks held by investors who do have exposure to that commodity sector do not underreact. An investor whose portfolio is exposed to a particular commodity pays attention to that commodity and incorporates that information into the stocks he owns. Alternatively, an investor whose portfolio does not have exposure to a commodity ignores that information. Third, I use news articles to understand the salience of the linkage between each firm and the commodity sector I have assigned to it. Firms that have many news articles associated with them that mention the commodity are efficient while firms whose articles do not mention the commodity often are inefficient. Fourth, I show that inattention to commodities is highest when the cross-sectional dispersion among equity returns - a proxy

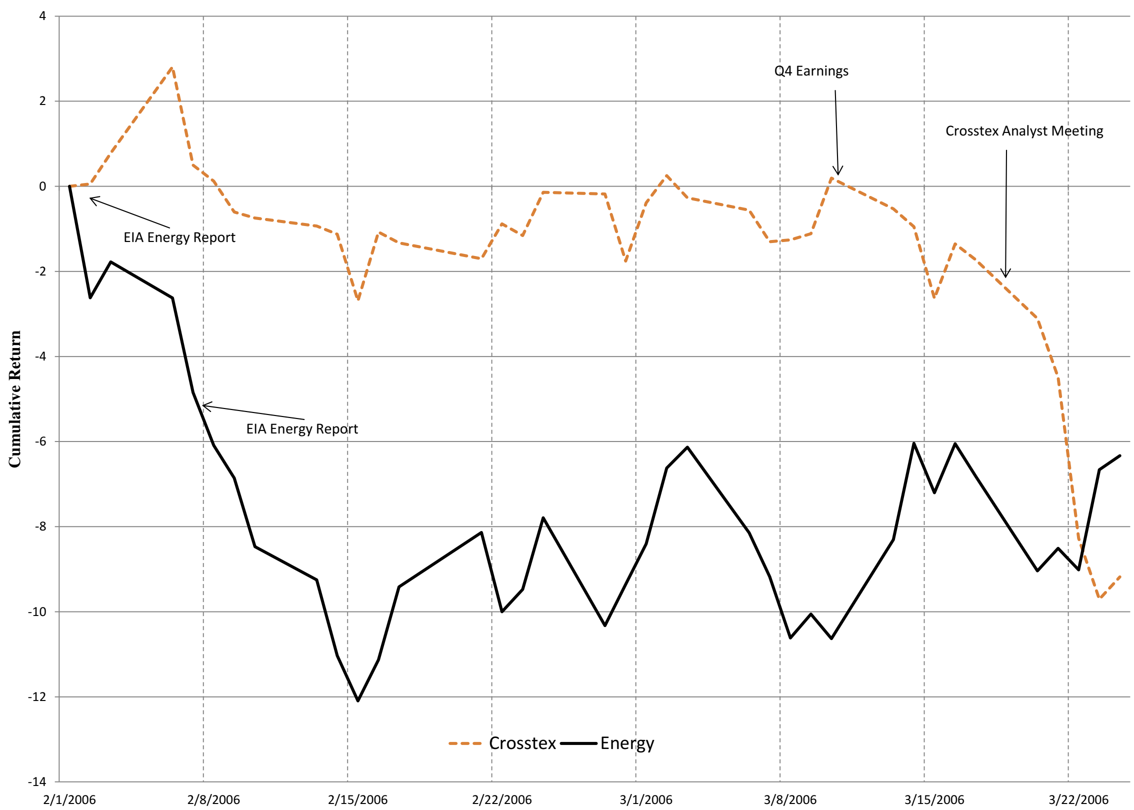


Figure 1.1: XTXI and Energy Cumulative Returns

This figure plots the cumulative returns to Energy commodities and Crosstex Energy (XTXI) between February and March 2006. In the first two weeks of February, the Energy Information Administration (EIA) released two bearish reports showing a buildup in energy commodities which sent the prices of these commodities lower; XTXI did not react significantly to this news. On March 10th 2006, XTXI reported its Q4 2005 and fiscal year 2005 earnings. Barry Davis, the CEO, described the announced information by saying: "We had a great fourth quarter and an outstanding year in 2005." Once again the stock does not have a significant reaction; however, on March 20th the company held an analyst meeting to discuss 2006 prospects and the stock took a significant hit. Revenue in 2006 was dependent on energy prices in 2006 which dropped by approximately 10% a month earlier.

for the amount of idiosyncratic news - is high. When investors are burdened with a significant amount of information they are only able to process a fraction of it which decreases efficiency of prices.

Finally I unpack the commodity sectors into individual commodities and use the entire CRSP universe to show that my results are not influenced by categorizing commodities into sectors or the sample selection procedure used throughout the study. Using the elastic net of Zou & Hastie (2005), I compute the overall commodity news to each industry stemming from individual commodities. Then I sort stocks from these industries into a long/short portfolio based on the commodity news of each industry that month, hold the portfolio for one month and then rebalance. This strategy generates an alpha of .55% per month in small securities and approximately zero in large securities. Some industries, however, have no relationship to any of the commodities in this study. Using securities in industries that are "newsworthy" instead of the entire CRSP universe generates a monthly alpha of 1.5% per month in small stocks and a .33% per month in large stocks, the latter of these being statistically insignificant.

Theoretical investigation of inattention can be traced back to Kahneman (1973) who notes that attention is a scarce resource. Sims (2003) develops a model of rational inattention suggesting that investors may have capacity constraints on their ability to process information. Hong & Stein (1999) develop a behavioral model of underreaction and overreaction to information. Recently, empirical tests of these theoretical notions have come into focus as researchers attempt to understand the pervasiveness of inattention in financial markets. In a highly influential paper, Cohen & Frazzini (2008) show that firms that are linked together through customer/supplier relationships are not always equally efficient in incorporating relevant information about each others prices. Shocks to customer firms travel to supplier firm prices slowly allowing for predictability in returns. Hong et al. (2000) show that momentum strategies are stronger in smaller stocks and interpret this as evidence of slow information diffusion because smaller stocks have lower analyst coverage. Information diffusion across asset classes has also begun to receive attention: Pollet (2005), Park & Ratti (2008) show that oil returns significantly predict some industries and the overall market return. Rizova (2010) examines informational efficiency across international stock markets and shows that stock markets in countries that are trading partners have intertemporal correlation. Empirical work has also begun to investigate the specific cognitive frictions that prevent information from being efficiently incorporated into prices. Smalling (2012) finds that stocks that comprise a large part of their owners' portfolios tend to have less post earnings announcement drift than those that comprise a small portion suggesting that investors ignore certain portions of their portfolios. Barber & Odean (2008) show that investors gravitate towards attention grabbing stocks: "preferences determine choices after attention has determined the choice set." Hirshleifer et al. (2009) show that there is a larger post earnings announcement drift after earnings announcement dates when many firms are reporting earnings compared to dates when fewer report earnings; they conclude that investors have limited attention and can be overwhelmed with information.

The remainder of this paper is organized as follows: Section 1.2 details the basic facts regarding the commodities used in the study, explains how the equity universe is selected and demonstrates a trading strategy that takes advantage of investor inattention. Section 1.3 examines several channels of friction that could prevent information about commodity returns from being efficiently incorporated into equity prices. Section 1.4 presents robustness to other effects and choice of methodologies. Finally, Section 1.5 concludes.

1.2 Commodity News and Equity Returns

Commodities are a large and important market: in 2012 over three billion contracts changed hands with trillions of dollars in outstanding notional[4]. Many of the commodities traded are used by companies in the real economy to produce everyday goods and services ranging from electricity to chocolate. I organize commodities into three commodity sectors: Agriculture (Ag), Energy and Metal. Table 1.1 lists the commodities used in the study by commodity sector; this set covers the most widely studied commodities in the literature.

[4]http://www.futuresindustry.org/downloads/FIA_Annual_Volume_Survey_2013.pdf

Table 1.1: List of Commodities: 1983-2012

List of commodities used in the study and their classification into commodity sectors. Note that RBOB HU denotes the time series splicing together of the Unleaded Gasoline (HU) contract and the Reformulated Blendstock for Oxygenate Blending (RBOB) as HU was phased out from trading. WTI denotes the West Texas Intermediate crude oil contract. Futures contract prices and specifications are obtained from Bloomberg as in Koijen et al. (2013). The sample runs from 1983-2012, though some commodities begin trading only in later years; they are added to the sample as they become available.

Ag	Energy	Metal
Cocoa	Brent Crude	Aluminum
Coffee	Gasoil	Copper
Corn	Heating Oil	Gold
Cotton	Natural Gas	Lead
Kansas Wheat	RBOB HU	Nickel
Soybeans	WTI	Silver
Sugar		Zinc
Wheat		

Futures contract prices and specifications are obtained from Bloomberg as in Koijen et al. (2013). Equity data is obtained from CRSP and Compustat.

1.2.1 Commodity Returns

The commodity sample runs from 1983-2012; the exact composition of each commodity sector changes over time as some commodities were not traded in 1983. The excess returns to each commodity are computed as a simple average of the returns to each future contract along the term structure of that commodity (up to 1 year in maturity); the excess return to each future contract is a fully margined return[5] as in Koijen et al. (2013) among others.

$$R^{fut}_{\tau,d,t} \equiv \frac{F_{\tau,d,t} - F_{\tau+1,d,t-1}}{F_{\tau+1,d,t-1}} \tag{1.1}$$

$$R_{d,t} \equiv \frac{1}{N_d} \sum_{k=1}^{N_d} R^{fut}_{k,d,t} \tag{1.2}$$

where $F_{\tau,d,t}$ is the price of a futures contract for commodity d with τ periods to maturity at time t and N_d is the number of futures contracts with maturity of less than 1 year for

[5] Unlike equity returns that require the transfer of funds equal to the price of the security, futures contracts generally require only a portion of the security price to be placed in a custodial account. To be conservative, I require that the full security price be placed into the account to avoid any issues with leverage.

commodity d. Some commodities (ex: agricultural commodities) have contracts that expire quarterly while other commodities (ex: energy commodities) have contracts that expire monthly thus N varies by commodity. To compute the return to each commodity sector (Energy, Ag, Metal), a simple average is taken across all the commodities in that sector.

$$R_{c,t} \equiv \frac{1}{M_c} \sum_{d=1}^{M_c} R_{d,t} \tag{1.3}$$

where M_c is the number of commodities in sector c.

1.2.2 Selecting Equity Universe

Commodities have an impact on many firms in the economy but not all firms. I aim to select firms that are most related to the commodity sectors described above. Individual firm returns are noisy; therefore, to select a subset of the CRSP universe that is related to commodities, I first look for industries that are related to the commodity sectors. I classify firms into industries using the two digit lagged - to prevent lookahead bias in classification - SIC code. I then define the return of industry i at time t, $IR_{i,t}$, as a value weighted return of the constituent securities of that industry. For each industry, I run a rolling (using a 5 year window with at least 3 years of returns) multivariate contemporaneous regression of industry return on the CRSP market and the commodity sectors using weekly (overlapping) data:

$$IR_{i,t} = a + \beta_{i,m,t} R_{m,t} + \beta_{i,E,t} R_{E,t} + \beta_{i,A,t} R_{A,t} + \beta_{i,Me,t} R_{Me,t} + \varepsilon_{i,t} \tag{1.4}$$

This simple regression identifies the industries that have a contemporaneous relationship to each of the commodity sectors. I associate each industry with a particular commodity sector at time t if its Newey-West p-value is at most 1% (t-statistic of 2.58). Therefore at time t a particular industry can be associated with $0-3$ different commodity sectors. If the industry is associated with 0 sectors then it is simply dropped from the sample for that period.

I further refine the sample because even the SIC categorizations are imperfect. Some businesses have multiple business segments and others may simply be misclassified. I would like to select companies that behave like the rest of their industry with respect to each commodity sector. Therefore, I also run regression (1.4) with individual stock returns on the left hand side and associate company i with commodity sector c at time t only if $\text{sign}(\beta_{i,c,t}) = \text{sign}(\beta_{j,c,t})$ where j is the industry that company i belongs to. That is, I associate a company with a particular commodity sector only if the company behaves (directionally) like the rest of its industry with respect to that commodity sector. This procedure leads to a universe of securities that have a contemporaneous relationship to these commodity sectors. Note that in all trading strategy results that I present in this study, all classification and universe selection happens using only backward looking information.

It is important to understand how well this procedure does in actually selecting companies that correlate with the aforementioned commodity sectors “out-of-sample”. Moreover, since the inattention trading strategy that I will present in the following section rebalances the portfolio monthly, as is standard in academic studies, it is important to determine if the securities that I have chosen have a contemporaneous correlation with these commodity sectors over a monthly return frequency. To answer this question, I form portfolios that are approximately market neutral but should have positive correlation with a particular commodity sector. Since individual security returns are extremely noisy, I use each security's industry β to the commodity sector as the sorting variable in this entire study (to break ties between securities having the same industry β when sorting into quintiles I use the individual security β). At the end of each month t, I select all securities that meet the filter in Section 1.2.2 and for each commodity sector, sort the associated stocks into terciles based on $\beta_{i,c,t}$. I then create a value weighted (equal weighted) portfolio within each tercile and go long the top tercile and short the bottom tercile for each commodity sector. These portfolios should have positive exposure to commodity sector c but minimal exposure to R_m. I compute the contemporaneous correlation between the return to this long/short portfolio, $R_{ec,t+1}$, and $R_{c,t+1}$: this is an “out-of-sample” correlation as the securities selected are based on information at t while the correlation is computed starting with returns at $t+1$; the portfolios are rebalanced monthly. Table 1.2 presents the results of this procedure for each commodity sector. I also report all other pairwise correlations between commodity sectors, the market and the equity portfolios.

The first row of the table shows the correlation of the CRSP market return with the commodity sectors, the equity value weighted mimicking portfolios and the equity equal weight mimicking portfolios. Ag and Metal have a fairly low correlation (.27 and .28, respectively) with R_m while Energy has an even lower correlation of .1. Among the equity mimicking portfolios, the equity value weighted metal (EQ VW Metal) portfolio has a noticeable correlation with the broader market while EQ VW Ag and EQ VW Energy have no significant correlation. In other words, the procedure to isolate only the commodity return away from the market is fairly successful. The second notable fact is that the commodities have a positive correlation among themselves: this is true for structural reasons (commodities tend to be traded by the same set of individuals and deleveraging events, for example, will have an impact on all of them) as well as fundamental economic reasons (demand for these inputs is driven by the broader economy, for example). We can see that the equity mimicking portfolios have a meaningful correlation with the actual commodities as intended (with Energy and Metal producing the best results). Finally, the correlation structure among the equity portfolios is fairly small as can be expected by specification of regression (1.4).

To get a better idea of how the selected sample of securities compares to the broader CRSP and NYSE universe (over the same time frame: 1983 - 2012), Panel A of Table 1.3 provides summary statistics on characteristics that describe the selected sample as well as CRSP and NYSE. To compute these summary statistics for a given set of securities (selected sample, CRSP, NYSE), each month I take an equal weighted (value weighted) cross sectional average of each characteristic across the sample. The time series properties of that cross sectional average are then reported. As is evident, the sample selected is very similar to the broader CRSP and NYSE universe; there are on average 723 firms per month that cover 21% of CRSP (by market capitalization; denoted as Fraction of CRSP Universe in the table). The average firm in the selected sample is larger than the average CRSP firm but smaller than the average NYSE firm; the selected firms' returns and book-to-market ratios are similar to CRSP and NYSE. I also determine what percentage of my selected securities have positive

Table 1.2: Commodity Sector and Equity Correlations

Correlations among the CRSP market, commodity sectors and "out-of-sample" commodity mimicking portfolios (constructed using equities) are presented. At the end of each month t, I select all securities that meet the filter described in Section 1.2.2 and for each commodity sector, sort the associated stocks into terciles based on $\beta_{i,c}$. I then create a value weighted (equal weighted) portfolio among each tercile and go long tercile three and short tercile one (for each commodity sector). The resulting portfolio should have positive exposure to commodity sector c but minimal exposure to R_m. I compute the contemporaneous correlation between the return to this long/short portfolio, $R_{ec,t+1}$, and $R_{c,t+1}$: this is an "out-of-sample" correlation as the securities selected are based on information at t while the correlation is computed starting with returns at $t+1$; the portfolios are rebalanced monthly. All possible pairs of correlations are presented.

	CRSP	Ag	Energy	Metal	EQ VW Ag	EQ VW Energy	EQ VW Metal	EQ EW Ag	EQ EW Energy	EQ EW Metal
CRSP	1	0.27	0.10	0.28	0.06	0.00	0.23	0.01	-0.05	0.10
Ag	0.27	1	0.27	0.43	0.14	0.14	0.32	0.10	0.17	0.39
Energy	0.10	0.27	1	0.31	0.11	0.57	0.10	0.07	0.58	0.17
Metal	0.28	0.43	0.31	1	0.07	0.18	0.46	0.01	0.20	0.45
EQ VW Ag	0.06	0.14	0.11	0.07	1	0.18	0.13	0.85	0.16	0.18
EQ VW Energy	0.00	0.14	0.57	0.18	0.18	1	0.09	0.16	0.80	0.18
EQ VW Metal	0.23	0.32	0.10	0.46	0.13	0.09	1	0.11	0.12	0.73
EQ EW Ag	0.01	0.10	0.07	0.01	0.85	0.16	0.11	1	0.17	0.16
EQ EW Energy	-0.05	0.17	0.58	0.20	0.16	0.80	0.12	0.17	1	0.20
EQ EW Metal	0.10	0.39	0.17	0.45	0.18	0.18	0.73	0.16	0.20	1

vs negative exposure to commodities: for each stock at a particular time t I compute the average β of that stock to it's associated commodities as $\overline{\beta_{i,t}} = \frac{1}{N_i}\sum_{c=1}^{N_i}\beta_{i,c,t}$ where N_i is the number of commodities associated with stock i. I then take a cross-sectional equal weighted (value weighted) average across all stocks in my universe for a particular month of $sign(\overline{\beta_{i,t}})$ and report the time-series properties of this average[6]. On average, roughly $50\% - 60\%$ of the securities in my sample have a positive commodity association with the remainder having a negative association. Therefore, the sample is fairly balanced between having a negative and positive exposure to the commodity sectors.

Panel B provides some insight regarding the types of SIC codes (equity industries) that are selected and how many commodity sectors affect each SIC code. On average there are 72.5 SIC codes per month in CRSP and my procedure deems an average of 16.1 relevant to the commodity sectors. Each SIC code is matched to an average of 1.2 commodity sectors; that is, most equity industries are only related to one commodity sector. I also list the top three equity industries (by $|\beta|$) that match to each commodity sector. For example "Oil and Gas Extraction" has the highest average absolute exposure to Energy out of all other industries just as "Agricultural Services" is most related to Ag[7]. As is evident, the selected equity industries make intuitive sense: we would expect that these equity industries have exposure to commodities.

1.2.3 Inattention Trading Strategy

The goal of this study is to show that equity investors do not fully appreciate the information available in commodity markets that is relevant for equities. Regression (1.4)

[6]The goal of this metric is to make sure that I have a sample that includes stocks with negative and positive commodity betas. An alternative methodology would have been to compute the percentage of betas each month that are positive (instead of collapsing them to the stock level and thus some stocks would enter into the average multiple times in a particular time period). Empirically this makes very little difference since most stocks have only one commodity sector associated with them.

[7]SIC category names are taken from the US Department of Labor 1987 SIC manual. Note that the name "Administration Of Environmental Quality and Housing Programs" listed under the Metal commodity sector is somewhat misleading as this SIC code is only selected between 2011 and 2012 during which it includes only one company: China Shen Zhou Mining & Resources, Inc., which is a metals mining company and hence has a high exposure to Metal.

Table 1.3: Summary Statistics

Panel A presents summary statistics describing the selected equity universe, CRSP and NYSE stocks. To compute these summary statistics for a given set of securities (selected sample, CRSP, NYSE), each month I take an equal weighted (value weighted) cross sectional average of each characteristic across the sample. The time series properties of that cross sectional average are then reported. Equity data is obtained from CRSP and Compustat spanning 1983 - 2012. "Fraction of CRSP Universe" denotes the fraction of the CRSP market capitalization that each universe comprises. "Fraction of Positive Commodity Beta Stocks" computes the average commodity β for a given (stock, month) tuple - since some stocks can have more than one commodity associated with them - and then computes the fraction of stocks that have average $\beta > 0$ for a particular month. The time series properties of this fraction are then reported as with the rest of the statistics. Panel B presents information regarding the selected SIC codes: the average number of total SIC codes, the average number of selected SIC codes and the average number of commodity sectors associated with each SIC code. It also lists the top three SIC codes associated with each commodity sector (by $|\beta|$ to the commodity sector). Note that the name "Administration Of Environmental Quality and Housing Programs" listed under the Metal commodity sector is somewhat misleading as this SIC code is only selected between 2011 and 2012 during which it includes only one company: China Shen Zhou Mining & Resources, Inc., which is a metals mining company and hence has a high exposure to Metal.

(a) Selected Universe Characteristics

	Selected Universe				CRSP	NYSE
Statistic	Mean	SD	Min	Max	Mean	Mean
Book-To-Market EW	0.95	0.42	0.40	3.45	1.31	1.03
Book-To-Market VW	0.46	0.11	0.19	0.88	0.46	0.48
Size (in thousands)	2,897,212	2,012,648	456,139	7,382,285	2,060,697	5,226,981
Excess Returns EW	0.78	6.07	-28.40	19.52	0.82	0.78
Excess Returns VW	0.63	4.37	-22.19	13.80	0.57	0.66
Number of Firms	723	309	233	1697	4940	1434
Fraction of CRSP Universe	0.21	0.11	0.05	0.51	1.00	0.80
Fraction of Positive Commodity Beta Stocks EW	0.60	0.17	0.25	1.00		
Fraction of Positive Commodity Beta Stocks VW	0.55	0.16	0.22	1.00		

(b) Selected SIC Code Summary

Total SIC Codes	Selected SIC Codes	Mean Assocations/SIC Code
72.5	16.1	1.2

Energy	Ag	Metal
Oil And Gas Extraction	Agricultural Services	Administration Of Environmental Quality And Housing Programs
Coal Mining	Coal Mining	Metal Mining
Water Transportation	Agriculture Production Livestock and Animal Specialties	Miscellaneous Repair Services

characterizes firms based on contemporaneous relationships with the commodity sectors. I define commodity news for stock i, $R_{i,c,t}$, in equation (1.5) as the dot product of its associated commodity sector returns and its industry β to those commodity sectors (I once again rely on industry β rather than individual stock β because individual stock returns are noisy). If investors are not able to fully appreciate these relationships then purchasing (selling) securities whose associated commodity news was positive (negative) should yield a profitable trading strategy; this is the hypothesis that will be tested in this section.

A stock can have several commodity sectors associated with it. For example fertilizer production is a very energy intensive activity so fertilizer producers might be exposed to energy returns. The procedure described in Section 1.2.2 associates each stock with the commodity sectors that have a significant effect on its returns. As noted earlier, I define commodity news for stock i at time t as the dot product of exposure and commodity return:

$$R_{i,c,t} \equiv \boldsymbol{\beta'_{j,c,t}} \boldsymbol{R_{c,t}} \tag{1.5}$$

where $\boldsymbol{\beta_{j,c,t}}$ is a vector of commodity sector exposures (with exposures to commodity sectors not associated with i set to 0) of industry j that contains stock i, and $\boldsymbol{R_{c,t}}$ is a vector of monthly commodity sector returns. In words, this is simply the total commodity news that will be experienced by stock i at time t.

At the end of each month I sort securities into quintiles based on $R_{i,c,t}$, form value weighted (equal weighted) portfolios and rebalance monthly. If investors are not fully attentive, then securities that experienced positive commodity news should continue to appreciate in value the following month while those that experienced negative commodity news should decline in value. Table 1.4 and 1.5 present the results of this experiment. As hypothesized, a strategy that goes long securities that have positive commodity news and short securities that have negative commodity news earns approximately 1% per month - in risk adjusted returns - in value weight portfolios and roughly 1.5% per month in equal weight portfolios. The alphas are monotonically increasing from the short portfolio to the long portfolio. Approximately half of the trading strategy alpha comes from the short portfolio and half from the long. The strategy has modest Sharpe ratios, no significant skewness, and some excess kurtosis without having any meaningful factor exposure. These facts suggest that the reason for this alpha has little to do with common explanations for equity anomalies such as shorting constraints or the phenomenon being limited to a small subset of securities.

Notably, this strategy generates an extra .5% per month in equal weighted portfolios as compared to value weighted portfolios suggesting that small securities may have stronger underreaction to commodity news. This is precisely what an inattention hypothesis would have predicted ex-ante: smaller securities tend to have fewer analysts covering them and have fewer institutional owners as noted by Hong et al. (2000). Thus there are fewer channels through which information could be incorporated into prices in a timely manner, relative to larger stocks. I test this hypothesis explicitly in Table 1.6. At the end of each month I split stocks into small and large securities along the NYSE median market capitalization and then sort securities into value weighted quintiles in each size category based on $R_{i,c,t}$. The table presents the results of a long-short portfolio that goes long (short) stocks with positive (negative) $R_{i,c,t}$: it is denoted as $5-1$. As suggested by earlier results, small securities have a significantly higher trading strategy alpha - and thus underreaction - than their big counterparts. A commodity underreaction strategy generates approximately 1.8% per month four factor alpha in small securities but a statistically insignificant .5% in large securities (difference of 1.331% with Newey-West t-statistic of 4.341). Clearly small securities have a significantly larger underreaction to commodity news than large securities. This highlights the importance of analysts and sophisticated investors to having efficient equity prices.

Another important prediction of an inattention hypothesis is that this trading strategy not revert the following month: if this month a stock incorporates some information that was available the previous month, it should not reverse next month. To check this, I form the the $5-1$ portfolio presented in Table 1.6 in the same month as the commodity news is available, one month after, two months after, etc. with the 0 lag indicating the contemporaneous relationship between news and returns. Figure 1.2 plots the results of this experiment. The results are consistent with underreaction to information: both small and large stocks

Table 1.4: Quintile Equity Sorts: Value Weight

Basic inattention trading strategy: securities are sorted at the end of each month into quintile value weight portfolios using $R_{i,c,t} \equiv \boldsymbol{\beta}'_{\boldsymbol{j,c,t}} \boldsymbol{R_{c,t}}$ (where $\boldsymbol{\beta_{j,c,t}}$ is a vector of commodity sector exposures, with exposures to commodity sectors not associated with i set to 0, of industry j that contains stock i and $\boldsymbol{R_{c,t}}$ is a vector of monthly commodity sector returns), held for one month and rebalanced. Equity returns are obtained from CRSP; commodity futures price and characteristics are obtained from Bloomberg. The sample runs from 1983 - 2012 with commodities added to each commodity sector as they become available.

q	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
1	0.087	-0.442	-0.500	-0.465	0.043	-0.005	3.153	0.985	0.021	0.139	-0.044
		[-1.329]	[-1.633]	[-1.598]							
2	0.346	-0.051	-0.095	-0.095	0.224	-0.651	3.577	0.765	-0.094	0.119	0.000
		[-0.216]	[-0.354]	[-0.405]							
3	0.708	0.249	0.154	0.146	0.461	-0.508	3.081	0.866	0.166	0.257	0.010
		[1.138]	[0.811]	[0.716]							
4	0.810	**0.407**	0.293	0.254	0.561	-0.148	2.149	0.793	0.109	0.320	0.051
		[1.980]	[1.470]	[1.244]							
5	1.180	**0.694**	0.511	0.484	0.642	-0.311	3.235	0.967	0.153	0.501	0.034
		[2.448]	[1.956]	[1.824]							
5-1	1.093	**1.136**	**1.011**	**0.949**	0.515	-0.178	1.225	-0.018	0.132	0.362	0.078
		[2.642]	**[2.257]**	**[2.200]**							

Table 1.5: Quintile Equity Sorts: Equal Weight

Basic inattention trading strategy: securities are sorted at the end of each month into quintile equal weight portfolios using $R_{i,c,t} \equiv \boldsymbol{\beta}'_{\boldsymbol{j,c,t}} \boldsymbol{R}_{\boldsymbol{c,t}}$ (where $\boldsymbol{\beta}_{\boldsymbol{j,c,t}}$ is a vector of commodity sector exposures, with exposures to commodity sectors not associated with i set to 0, of industry j that contains stock i and $\boldsymbol{R}_{\boldsymbol{c,t}}$ is a vector of monthly commodity sector returns), held for one month and rebalanced. Equity returns are obtained from CRSP; commodity futures price and characteristics are obtained from Bloomberg. The sample runs from 1983 - 2012 with commodities added to each commodity sector as they become available.

q	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
1	-0.059	-0.657	**-0.890**	**-0.725**	-0.027	-0.100	2.799	1.064	0.584	0.553	-0.209
		[-1.733]	**[-2.541]**	**[-2.327]**							
2	0.311	-0.212	-0.415	-0.292	0.165	-0.144	2.431	0.905	0.706	0.492	-0.157
		[-0.689]	[-1.487]	[-1.104]							
3	0.777	0.247	0.094	0.157	0.423	-0.230	2.652	0.878	0.912	0.383	-0.079
		[0.878]	[0.430]	[0.725]							
4	1.100	0.557	0.374	0.510	0.575	-0.076	2.125	0.915	0.787	0.429	-0.172
		[1.722]	[1.447]	[1.910]							
5	1.460	**0.880**	**0.635**	**0.773**	0.667	0.095	2.964	1.015	0.779	0.597	-0.175
		[2.427]	**[2.329]**	**[2.370]**							
5-1	1.519	**1.537**	**1.525**	**1.498**	0.748	0.134	2.966	-0.049	0.196	0.044	0.033
		[3.866]	**[3.788]**	**[3.497]**							

Table 1.6: Size Double Sort (NYSE Median Break)

At the end of each month I split stocks into small and large securities along the NYSE median market capitalization and then sort securities into value weighted quintiles in each size category based on $R_{i,c,t}$. The table presents the results of a long-short portfolio that goes long (short) stocks with positive (negative) $R_{i,c,t}$ in each size category: it is denoted as $5-1$.Equity data is obtained from CRSP and commodity data from Bloomberg from 1983 - 2012.

Size	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Small	1	-0.271	**-0.910**	**-1.178**	**-1.053**	-0.113	-0.212	2.647	1.170	0.586	0.662	-0.159
			[-2.18]	**[-3.18]**	**[-3.10]**							
	2	0.183	-0.391	**-0.582**	-0.506	0.091	-0.604	2.317	0.995	0.789	0.477	-0.097
			[-1.23]	**[-2.01]**	[-1.82]							
	3	0.658	0.064	-0.078	-0.061	0.339	-0.573	2.544	1.003	0.925	0.372	-0.023
			[0.25]	[-0.35]	[-0.28]							
	4	0.943	0.349	0.164	0.259	0.485	-0.265	1.986	1.007	0.880	0.454	-0.121
			[1.23]	[0.75]	[1.23]							
	5	1.577	**0.961**	**0.691**	**0.776**	0.695	-0.083	2.651	1.098	0.850	0.686	-0.107
			[2.70]	**[2.31]**	**[2.49]**							
	5-1	1.848	**1.872**	**1.869**	**1.828**	0.818	-0.110	2.571	-0.073	0.264	0.024	0.052
			[3.95]	**[3.96]**	**[3.80]**							
Big	1	0.276	-0.263	-0.298	-0.255	0.136	0.042	3.576	1.005	-0.078	0.075	-0.054
			[-0.86]	[-1.03]	[-0.90]							
	2	0.331	-0.093	-0.126	-0.126	0.197	-0.643	3.665	0.821	-0.171	0.088	0.000
			[-0.38]	[-0.45]	[-0.44]							
	3	0.729	0.280	0.198	0.142	0.456	0.170	3.100	0.882	0.004	0.243	0.070
			[1.21]	[0.87]	[0.63]							
	4	0.749	0.351	0.221	0.192	0.502	-0.737	4.803	0.807	-0.014	0.361	0.037
			[1.54]	[1.05]	[0.83]							
	5	0.936	0.462	0.310	0.242	0.510	-0.470	3.467	0.955	0.084	0.436	0.085
			[1.65]	[1.15]	[0.89]							
	5-1	0.660	0.725	0.607	0.498	0.312	-0.336	1.445	-0.050	0.162	0.361	0.139
			[1.73]	[1.42]	[1.19]							
Small - Big	5-1	1.188	**1.147**	**1.262**	**1.331**	0.830	0.185	0.524	-0.023	0.102	-0.337	-0.087
			[3.77]	**[3.65]**	**[4.34]**							

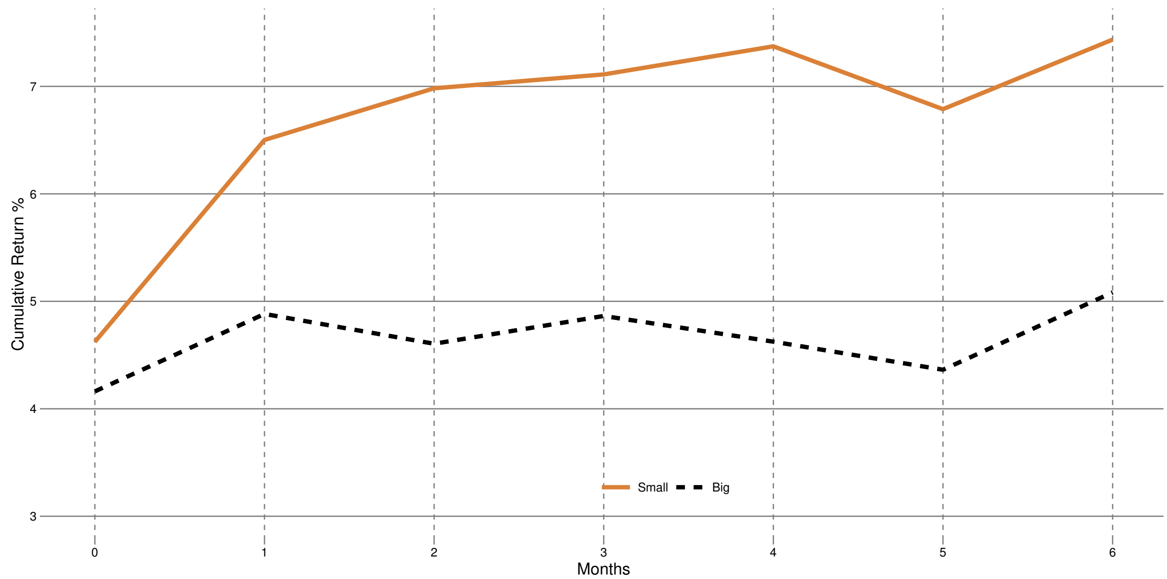


Figure 1.2: Inattention Horizon

A plot of returns to the $5-1$ inattention portfolio by varying the time between commodity news and portfolio formation month.

have a contemporaneous reaction to commodity news, however, small stocks have a significant amount of underreaction as evidenced by their continued rise the following month. Importantly, this effect does not reverse in the following months.

1.3 Inattention Channels

There are four channels through which investors can incorporate commodity news into equity prices with a lag: they can ignore a particular stock so that stock incorporates information slowly, they can be attentive to news released by the company but ignore the commodity market, they may not understand that a particular stock is affected by commodity prices or they may be generally inattentive because they are overwhelmed with many sources of idiosyncratic information in the spirit of Hirshleifer et al. (2009). In this section I will use mutual fund holdings to show that stocks owned by investors who are ignoring the particular stock or ignoring commodity market information have a larger underreaction than stocks owned by attentive investors. I will also show that stocks that are frequently mentioned alongside their associated commodities in news articles are efficient in incorporating

Table 1.7: Mutual Fund Universe Summary Statistics

Summary statistics regarding the mutual funds used to construct inattention measures. Each month I take an equal weighted average among all funds of a particular statistic (i.e. number of funds), then I report the time series properties of that average.

Statistic	Mean	SD	Min	Max
Number of Funds	2299.29	1268.32	383	4084
Number of Stocks Held by All Funds	4141.55	836.25	3022	5872
Fraction of CRSP Number of Stocks	0.75	0.14	0.50	0.92
Number of Stocks Held by Each Fund	77.59	19.00	23.00	102.28
Fund's Portfolio Value (in $B)	0.46	0.27	0.00	1.28

commodity information into their prices; stocks that are rarely mentioned together with the commodity are inefficient at doing this. In other words, companies that are clearly associated with a commodity by investors are efficient while those that have more nuanced connections to the commodity market (and therefore not mentioned together in the press) are inefficient. Furthermore, cognitive burden for investors varies through time. Some periods have a lot of idiosyncratic news, and thus investors must pay attention to many different information sources, while other periods have less and investors only need to pay attention to the overall market. I show that underreaction to commodities is significantly larger in periods with high informational burden.

I obtain data on mutual fund holdings - the sophisticated investors - from Thomson Reuters Mutual Fund Holdings (S12) database and include only domestic actively managed mutual funds following Kacperczyk et al. (2008). I remove any fund from the sample that holds more than $1,000$ securities in their portfolio or contains the word "index" in the fund name. A significant percentage of funds report quarterly holdings data though they are required to report their holdings every six months. If a particular fund has not reported holdings within a one year period, I assume that fund has disappeared and remove it from the sample at that time. Securities held by fewer than 5 mutual funds are excluded. Table 1.7 provides summary statistics on the funds whose holdings are used to construct inattention measures.

1.3.1 Individual Stock Inattention

One specific channel by which a stock can be slow to fully incorporate all available information is investors simply ignoring this particular security. An investor that has a limited capacity for information processing has to prioritize the items that he pays attention to. Specifically, an investor that owns a portfolio of securities will pay more attention to securities that generate a volatile P&L stream within his portfolio relative to other securities he owns - they appear "riskier"[8]. This could happen because the position the investor holds in that security is very large and thus even small swings in value translate to large P&L swings. It could also happen because this particular security is experiencing anomalous volatility due to fundamental news about the company. Both of these causes lead to the same effect: they create a volatile P&L stream causing the investor to look more deeply at the company to see what is driving the increased volatility and if position adjustment in that security is necessary. This extra attention given to the security by investors increases its efficiency to publicly available information.

I examine this hypothesis with respect to commodity news using mutual fund holdings data. I show that stocks that deliver high P&L variance for their investors (relative to other stocks that those investors hold) are efficient in reacting to commodity information. Alternatively, stocks that do not have a high P&L variance within an investor's portfolio don't attract much attention and are slow to incorporate all available commodity information in their prices. For a particular stock i held by fund f at time t, I define the amount (in dollars) held of that stock by f as:

$$H_{i,f,t} = SHARES_{i,f,t} P_{i,t} \tag{1.6}$$

where $SHARES_{i,f,t}$ is the number of shares held by f of i and P is the price of i. The P&L on a particular day is simply the change in the value of the holding:

$$\Delta H_{i,f,t+1} = SHARES_{i,f,t} P_{i,t+1} - SHARES_{i,f,t} P_{i,t}$$

[8]This is simply a heuristic and surely does not capture the correlation a security has with other securities within the portfolio which is clearly important for risk measurement.

$$= \quad SHARES_{i,f,t} \Delta P_{i,t+1} \tag{1.7}$$

During a particular month, I compute the variance of ΔH for each security in a fund's portfolio by taking the variance of $\Delta H_{i,f,t}$ within the month:

$$\sigma^2(\Delta H_{i,f,t}) = \frac{1}{T-1} \sum_{t=1}^{T} \left(\Delta H_{i,f,t} - \overline{\Delta H}_{i,f,t} \right)^2 \tag{1.8}$$

This quantity is simply the variance of the realized P&L that fund f experienced from security i during a particular month. To determine if this is important or not for fund f (since the importance of this quantity is relative for each fund: funds that hold very volatile securities may view a particular security as uneventful while those that hold less volatile securities may view this security as highly anomalous), I scale this quantity by the sum of $\sigma^2(\Delta H)$ of the other securities in fund f's portfolio:

$$RAWATTN^s_{i,f,t} = \frac{\sigma^2(\Delta H_{i,f,t})}{\sum_{j=1}^{K_f} \sigma^2(\Delta H_{j,f,t})} \tag{1.9}$$

where K_f is the number of securities held by fund f. This is simply the variance of a particular security's P&L scaled by the sum of the variances of the P&L of the other securities. It gives us a measure of how anomalous the P&L stream of security i has been in a particular month for fund f relative to the other securities they hold. If a security is experiencing highly anomalous P&L then fund f may take a closer look to see what is driving the high variance as it has a material impact on their portfolio.

Finally, attention paid to a particular security is cumulative across sophisticated investors: the more sophisticated investors pay attention the higher the chance that stock i will be efficient. To capture this notion I collapse $RAWATTN^s_{i,f,t}$ to the stock level by simply summing across all the funds that hold i in their portfolio during month t:

$$RAWATTN^s_{i,t} = \sum_{f \in F_i} RAWATTN^s_{i,f,t} \tag{1.10}$$

Implicitly, this measures how much attention is devoted to i by its sophisticated owners treating every one of the owners as equally capable (that is, no fund's attention to i is more important than any other only the quantity of attention devoted by f to i matters).

Furthermore, funds that are not in my universe are assigned a capability score of 0: they may be paying attention but they are unsophisticated and thus their expertise is irrelevant in increasing efficiency of i. This is done largely because holdings information is unavailable for hedge funds and other classes of investors and retail investors would likely be unsophisticated participants. If a stock has a high $RAWATTN^s$ then it should be more efficient than a stock that has low $RAWATTN^s$.

This particular attention metric likely has significant loadings on characteristics that are already known to influence stock efficiency. For example, we know that breadth of ownership (defined as the number of funds that hold a particular stock), $BREADTH_{i,t}$, has an impact on efficiency as noted in Chen et al. (2002), among others. Other such variables may also be important such as institutional ownership, $IO_{i,t}$ defined as the total mutual fund ownership of a stock relative to its market capitalization[9], security market capitalization, $ME_{i,t}$ - which I have already shown affects this particular anomaly, book-to-market ratio, $BM_{i,t}$ (where $bm_{i,t} = log(BM_{i,t})$), security market beta, $\beta_{i,m,t}$, computed from a four factor model on daily data during month t, last month's security return, $R_{i,t-1}$, security momentum $R_{i,t-13\to t-2}$ defined as the 12 month security return up to the previous month, and idiosyncratic volatility, $IV_{i,t}$, defined as the standard deviation of residuals from a four factor model attribution regression in month t using daily data. It is important to residualize for these quantities because their effects are already known and it is not my goal to capture them. Second, they may be obfuscating the true metric that I am attempting to measure. Finally, I want to show that this is truly a new and unique channel of inattention that has not yet been shown in previous research. To residualize $RAWATTN^s_{i,t}$ to this set of control variables, I run monthly Fama-MacBeth regressions of the form:

$$\begin{aligned} RAWATTN^s_{i,t} &= \theta_{0,t} + \theta_{bm,t}log(BM_{i,t}) + \theta_{me,t}log(ME_{i,t}) + \theta_{io,t}IO_{i,t} + \\ &+ \theta_{br,t}BREADTH_{i,t} + \theta_{mb,t}\beta_{m,i,t} + \theta_{r,t}R_{i,t} + \theta_{mom,t}R_{i,t-12\to t-1} + \\ &+ \theta_{iv,t}IV_{i,t} + \varepsilon^s_{i,t} \end{aligned} \tag{1.11}$$

[9] I use mutual fund ownership since this is the universe of investors I am concerned with in this article.

Table 1.8: Stock Inattention Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (1.11). The $RAWATTN^s$ metric has highly significant loadings on many known factors that affect stock efficiency; by residualizing to these metrics and using $ATTN^s$, equation (1.12), as the attention sorting variable I am able to purge their effects and focus on the unique portion of the variable that captures the effects I am demonstrating.

	$bm_{i,t}$	$log(ME_{i,t})$	$IO_{i,t}$	$BREADTH_{i,t}$	$\beta_{i,m,t}$	$R^e_{i,t}$	$R^e_{i,t-12\to t-1}$	$IV_{i,t}$	(Intercept)	R^2	N
(1)	**-0.181**	**-0.288**	**-1.587**	**0.021**					**3.291**	50.31%	490.191
	[-12.72]	**[-7.49]**	**[-4.95]**	**[28.47]**					**[7.06]**		
(2)	**-0.178**	**-0.294**	**-1.633**	**0.021**	**0.064**				**3.307**	50.76%	490.159
	[-12.76]	**[-7.56]**	**[-5.17]**	**[28.56]**	**[4.27]**				**[7.13]**		
(3)	**-0.125**	**-0.314**	**-1.626**	**0.022**	**0.063**	**0.005**	**0.002**		**3.536**	51.86%	484.191
	[-11.95]	**[-7.49]**	**[-5.27]**	**[28.39]**	**[4.25]**	**[4.72]**	**[4.86]**		**[7.11]**		
(4)	**-0.089**	**-0.186**	**-1.405**	**0.021**	**0.041**	**0.002**	**0.002**	**0.018**	**1.364**	54.24%	484.191
	[-7.87]	**[-5.03]**	**[-4.61]**	**[26.06]**	**[3.26]**	**[2.35]**	**[5.99]**	**[12.40]**	**[3.16]**		

Each month I extract the residuals, $\varepsilon^s_{i,t}$ and define a residualized attention metric as

$$ATTN^s_{i,t} \equiv \varepsilon^s_{i,t} \tag{1.12}$$

This particular metric captures the effects that I would like to demonstrate while controlling for already known factors affecting anomalies. Table 1.8 presents the results of these Fama-MacBeth regressions.

It is interesting to briefly look at the results of these regressions to understand the loadings that $RAWATTN^s$ contains: it is negatively related to BM indicating that stocks commanding higher attention have a lower book to market. This result is interesting in it's own right given the long standing debate regarding the value effect being a behavioral phenomenon or a rational one, for example Porta et al. (1997). The results of my regression are certainly supportive of a behavioral connection between attention and the book-to-market metric. $RAWATTN^s$ also loads negative on ME suggesting that larger stocks have less attention paid to them. This loading is counter intuitive and serves to obfuscate the true effect I am attempting to show highlighting the importance of controlling for these factors. I have shown for my particular return signal (commodity market information) in Table 1.6: small stocks underreact significantly more to commodity information than larger ones. $RAWATTN^s$ also loads negatively on mutual fund ownership surprisingly: stocks that have more sophisticated investors holding them should be more efficient (of course these are loadings conditional on other control variables). It loads positively, however, on the number of owners partly by construction (since a sum is taken across owners of a particular stock) but this loading also conforms to the intuition that more sophisticated investors is better for efficiency. Finally $RAWATTN^s$ loads positively on idiosyncratic volatility: part of this can happen by construction since $RAWATTN^s$ includes the variance of price changes but this result once again is supported by literature (for example Barber & Odean (2008) find that idiosyncratic volatility draws the attention of retail investors). The R^2 of these regressions is sizable indicating that a large amount of variation of $RAWATTN^s$ is captured by controls underscoring the importance of residualizing this metric. I present results based on $RAWATTN^s$ as well as $ATTN^s$ in Tables 1.9 and 1.10.

To test this channel of underreaction, at the end of each month I sort stocks into low and high attention stocks based on $ATTN^s_{i,t}$ ($RAWATTN^s_{i,t}$). Then within each attention category I sort stocks into quintiles based on $R_{i,c,t}$ and form value weight quintile portfolios. I go long stocks that have had positive commodity news and short those having negative commodity news. I hold the portfolio over the following month and rebalance monthly. The inattention hypothesis (and this channel specifically) predicts that the trading strategy implemented in low attention securities would have significantly higher alpha than the same strategy implemented in high attention securities. Table 1.9 and 1.10 present exactly this result.

The $5-1$ portfolio in the low attention category as measured by $RAWATTN^s$ produces a monthly alpha of approximately 1.4% (Newey-West t-stat of 3.19) while the same $5-1$ portfolio within the high attention category produces a statistically insignificant .7% per month of alpha (N-W tstat of 1.47). Examining the factor loadings within each attention category among the quintile sorts it is obvious that the low attention category portfolios one through five have a larger exposure to size and value factors (i.e. these stocks are smaller and have higher book-to-market). The $5-1$ portfolio in the low attention category has a positive loading on small stocks and no significant value loading.

Table 1.9: Individual Stock Attention - $RAWATTN^s$

At the end of each month stocks are sorted into low and high attention stocks using $RAWATTN^s_{i,t}$ and then into value weight quintile portfolios using $R_{i,c,t}$. Stocks that are ignored by investors underreact to commodity news significantly more than those that are closely monitored.

$RAWATTN^s$	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Low	1	0.096	-0.458	**-0.711**	-0.571	0.045	0.182	-1.422	1.044	0.336	0.617	-0.178
			[-1.28]	**[-2.21]**	[-1.81]							
	2	0.419	-0.087	-0.187	-0.113	0.219	-0.702	0.561	0.842	0.675	0.236	-0.093
			[-0.23]	[-0.53]	[-0.33]							
	3	0.542	0.034	-0.158	-0.125	0.309	-0.612	-1.035	0.915	0.627	0.499	-0.042
			[0.13]	[-0.79]	[-0.63]							
	4	1.123	**0.615**	0.428	**0.480**	0.611	0.093	0.210	0.921	0.538	0.476	-0.065
			[2.33]	[1.82]	**[2.05]**							
	5	1.626	**1.037**	**0.773**	**0.827**	0.768	-0.280	0.002	1.085	0.672	0.682	-0.069
			[3.25]	**[2.57]**	**[2.55]**							
	5-1	1.529	**1.495**	**1.484**	**1.398**	0.764	0.100	-1.602	0.041	0.336	0.065	0.109
			[3.43]	**[3.47]**	**[3.19]**							
High	1	0.187	-0.351	-0.397	-0.354	0.090	-0.043	0.365	1.003	-0.052	0.104	-0.054
			[-1.02]	[-1.23]	[-1.06]							
	2	0.434	0.043	-0.035	-0.093	0.279	-0.113	-2.079	0.789	-0.089	0.234	0.074
			[0.17]	[-0.15]	[-0.38]							
	3	0.820	0.368	0.290	0.236	0.495	0.392	0.332	0.851	0.202	0.233	0.069
			[1.37]	[1.07]	[0.85]							
	4	0.708	0.263	0.149	0.139	0.460	-0.697	0.709	0.881	-0.023	0.309	0.013
			[1.20]	[0.74]	[0.64]							
	5	1.063	**0.581**	0.406	0.334	0.561	-0.477	0.532	0.966	0.180	0.500	0.091
			[1.96]	[1.45]	[1.15]							
	5-1	0.876	**0.932**	0.802	0.688	0.392	-0.268	-1.580	-0.037	0.232	0.397	0.145
			[1.98]	[1.68]	[1.47]							
Low - High	5-1	0.653	0.562	**0.682**	**0.710**	0.401	0.331	-0.222	0.078	0.104	-0.332	-0.036
			[1.77]	**[2.03]**	**[2.07]**							

Table 1.10: Individual Stock Attention - $ATTN^s$

At the end of each month stocks are sorted into low and high attention stocks using $ATTN^s_{i,t}$ and then into value weight quintile portfolios using $R_{i,c,t}$. Stocks that are ignored by investors underreact to commodity news significantly more than those that are closely monitored.

$ATTN^s$	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Low	1	0.141	-0.433	-0.488	-0.466	0.067	-0.061	0.660	1.078	-0.034	0.139	-0.029
			[-1.34]	[-1.57]	[-1.54]							
	2	0.552	0.105	0.011	0.027	0.333	-0.318	-1.741	0.874	-0.074	0.244	-0.020
			[0.39]	[0.04]	[0.10]							
	3	0.680	0.210	0.096	0.136	0.421	-0.483	-0.887	0.911	-0.022	0.287	-0.051
			[0.90]	[0.47]	[0.65]							
	4	0.861	**0.465**	0.352	0.363	0.590	0.074	-1.459	0.765	0.089	0.297	-0.014
			[2.12]	[1.72]	[1.63]							
	5	1.287	**0.757**	**0.594**	**0.614**	0.665	-0.215	-0.289	1.026	0.133	0.426	-0.026
			[2.68]	**[2.16]**	**[2.23]**							
	5-1	1.145	**1.189**	**1.082**	**1.080**	0.521	-0.268	-0.194	-0.052	0.168	0.287	0.002
			[2.78]	**[2.42]**	**[2.45]**							
High	1	0.333	-0.195	-0.340	-0.327	0.153	-0.072	-0.693	1.002	0.217	0.381	-0.016
			[-0.50]	[-1.06]	[-0.98]							
	2	0.347	-0.104	-0.128	-0.107	0.167	-1.482	5.811	0.796	0.211	0.057	-0.027
			[-0.27]	[-0.37]	[-0.31]							
	3	0.599	0.156	0.053	-0.065	0.317	0.194	0.121	0.842	0.368	0.328	0.149
			[0.52]	[0.19]	[-0.22]							
	4	0.499	-0.003	-0.092	-0.198	0.275	-0.950	1.569	0.961	0.230	0.284	0.135
			[-0.01]	[-0.38]	[-0.79]							
	5	0.783	0.271	0.048	0.023	0.371	-0.349	-1.084	0.999	0.364	0.608	0.031
			[0.81]	[0.16]	[0.08]							
	5-1	0.450	0.467	0.387	0.350	0.185	-0.196	-1.717	-0.003	0.147	0.227	0.047
			[1.07]	[0.86]	[0.79]							
Low - High	5-1	0.696	**0.723**	0.695	**0.730**	0.370	0.377	-1.451	-0.050	0.021	0.060	-0.045
			[2.09]	[1.95]	**[2.03]**							

Results presented in Table 1.10 prevent such implicit sorting from taking place by sorting on a residualized version of $RAWATTN^s$. The $5-1$ portfolio in the low attention category as measured by $ATTN^s$ generates approximately 1% of alpha (N-W tstat of 2.45; a lower number than when measured with $RAWATTN^s$) vs .35% (N-W tstat of .79) in high attention securities. The spread between these two portfolios of .73% per month remains roughly the same as when sorts were done using $RAWATTN^s$; rather, the average return of the $5-1$ portfolio within each category has been decreased. We can also see that quintile portfolios formed in low attention securities don't have a significant exposure to small stocks (in fact, they have lower average exposure to SMB than quintile portfolios in the high attention category). This is reassuring as breaking this linkage was the intention of the residualization procedure - equation (1.11). Same with exposure to HML (value): quintile portfolios formed in the low and high attention categories have approximately the same loading on this factor.

I have just shown that the attention allocated to stocks by sophisticated funds is important: a higher number of funds paying high amounts of attention creates more efficient pricing. However, as I will show in the next section, the characteristics of the actual sophisticated investors are important as well. Sophisticated investors that pay attention to the commodity market are better at making prices of stocks that depend on the commodity market efficient.

1.3.2 Commodity News Inattention

The second type of inattention that could occur relates to what information sources investors observe. Certain funds may simply ignore commodities as a source of information because it is not of first order importance. Consider a fund whose overall portfolio does not have any commodity exposure: this can be, for example, because the fund owns a combination of securities with offsetting exposure. Since the actual portfolio does not move in response to commodities, the fund has less incentive to pay attention to commodities as an information source relative to an investor whose portfolio is strongly correlated with commodities. Therefore, stocks that are owned by funds that have no incentive to pay attention to commodities will be less efficient at incorporating news from the commodity market. In this section I am concerned with discriminating among sophisticated investors that own shares of a particular stock based on their incentive to monitor commodity information.

To compute the exposure of fund f to commodity sector c at time t, $\beta_{f,c,t}$, I simply take a value weighted average of $\beta_{i,c,t}$ based on the fund's holdings:

$$\beta_{f,c,t} \equiv \sum_{i=1}^{K_f} w_{i,f,t}\beta_{i,c,t} \tag{1.13}$$

where K_f is the number of stocks held by fund f at time t and $w_{i,f,t}$ is the weight of stock i in f's portfolio. Note that $\beta_{i,c,t}$ is defined to be the β of the industry that stock i is in to commodity sector c and stocks that are in industries with insignificant β to c are assigned a $\beta_{i,c,t} \equiv 0$ as described in Section 1.2.2. Attention paid by fund f to a commodity sector c, is therefore proportional to $|\beta_{f,c,t}|$ (since the sign does not matter: a fund that owns airline stocks should be equally concerned with Energy movements as a fund that owns oil producers).

As opposed to the previous section which was concerned with the attention being paid to a particular stock by classes of investors, in this section I would like to discriminate among the sophisticated investors that participate in ownership of a particular stock. Therefore, for a particular stock i, the measure of attention devoted by i's sophisticated investors to commodity sector c will simply be the average of their absolute exposures to that commodity sector:

$$SRAWATTN^n_{i,c,t} = \frac{1}{F_i}\sum_{f=1}^{F_i} |\beta_{f,c,t}| \tag{1.14}$$

where F_i is the number of funds that own stock i. Note that a stock can be associated with several commodity sectors as previously discussed and will therefore have multiple values of $SRAWATTN^n_{i,c,t}$: its owners pay attention to each commodity sector differently. The relevance of each commodity sector to stock i is, of course, proportional to $|\beta_{i,c,t}|$. If a stock has a large exposure to Metal and a small exposure to Ag, then the amount of attention being paid by the owners of stock i to Metal is much more important than the amount of attention being paid to Ag. Therefore, to aggregate this to the stock level, I take a $|\beta_{i,c,t}|$

weighted average of $SRAWATTN_{i,c,t}$ for each i and t and define

$$RAWATTN^n_{i,t} \equiv \sum_{c=1}^{C_i} w_{i,c,t} SRAWATTN_{i,c,t} \tag{1.15}$$

where C_i is the number of commodities that i has exposure to at time t and $w_{i,c,t} \equiv \frac{|\beta_{i,c,t}|}{\sum_{c=1}^{C_i} |\beta_{i,c,t}|}$.

$RAWATTN^n_{i,t}$ abstracts away from investor classes and discriminates among the sophisticated owners of stock i. The goal is to show that in addition to attention being paid to a particular stock by sophisticated investors, the types of sophisticated investors also have an important effect on the efficiency of prices. A stock that has investors highly focused on its associated commodity sector will likely be efficient at incorporating information from that commodity sector because those investors have significant incentives to pay attention to that information source. On the other hand, a stock owned by sophisticated investors whose portfolios have very little to do with the associated commodity sector will likely be inefficient with respect to information from that sector.

Similar to $RAWATTN^s$, $RAWATTN^n$ will also have important loadings on various previously known factors. In order to purge any effect those factors may have, I employ monthly Fama-MacBeth regressions with $RAWATTN^n$ as the dependent variable and the same set of controls as in (1.11):

$$\begin{aligned} RAWATTN^n_{i,t} &= \theta_{0,t} + \theta_{bm,t} log(BM)_{i,t} + \theta_{me,t} log(ME_{i,t}) + \theta_{io,t} IO_{i,t} + \\ &+ \theta_{br,t} BREADTH_{i,t} + \theta_{mb,t} \beta_{m,i,t} + \theta_{r,t} R_{i,t} + \theta_{mom,t} R_{i,t-12\to t-1} + \\ &+ \theta_{iv,t} IV_{i,t} + \varepsilon^n_{i,t} \end{aligned} \tag{1.16}$$

I define a residualized commodity attention metric for each stock as the residual of this regression:

$$ATTN^n_{i,t} \equiv \varepsilon^n_{i,t} \tag{1.17}$$

and report results for $RAWATTN^n$ as well as $ATTN^n$ as measures of attention to commodity news. The results of these orthogonalizing regressions are presented in Table 1.11. Similarly to $RAWATTN^s$ it has a negative loading on BM and a positive loading on IV. The R^2

Table 1.11: News Inattention Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (1.16). The $RAWATTN^n$ metric has highly significant loadings on many known factors that affect stock efficiency; by residualizing to these metrics and using $ATTN^n$, equation (1.17), as the attention sorting variable I am able to purge their effects and focus on the unique portion of the variable that captures the effects I am demonstrating.

	$log(BM_{i,t})$	$log(ME_{i,t})$	$IO_{i,t}$	$BREADTH_{i,t}$	$\beta_{i,m,t}$	$R^e_{i,t}$	$R^e_{i,t-12\rightarrow t-1}$	$IV_{i,t}$	(Intercept)	R^2	N
(1)	**-0.007**	0.000	-0.009	**0.000**					**0.029**	4.90%	441.638
	[-3.76]	[0.54]	[-0.54]	**[-2.31]**					**[2.30]**		
(2)	**-0.007**	0.000	-0.009	**0.000**	**-0.002**				**0.031**	6.98%	441.613
	[-3.99]	[0.53]	[-0.60]	**[-3.00]**	**[-4.16]**				**[2.86]**		
(3)	**-0.007**	0.000	-0.004	**0.000**	**-0.002**	0.000	0.000		**0.030**	10.34%	436.238
	[-4.70]	[0.45]	[-0.26]	**[-3.96]**	**[-4.58]**	[-0.71]	[-1.62]		**[2.84]**		
(4)	**-0.007**	**0.003**	0.003	**0.000**	**-0.002**	0.000	0.000	**0.000**	-0.010	12.07%	436.238
	[-4.33]	**[4.94]**	[0.19]	**[-4.53]**	**[-5.24]**	[-1.44]	[-1.48]	**[9.56]**	[-1.21]		

from these regressions are significantly lower which suggests that this metric is capturing a significant amount of information outside of the control variables. I present return sorts categorized by $RAWATTN^n$ and $ATTN^n$ to verify that both produce the results we would expect: stocks owned by funds that don't have a significant commodity exposure are slow to react to commodity news.

Table 1.12 presents the results of sorting stocks into low and high attention categories based on $RAWATTN^n$. Then within each attention category I sort stocks into quintiles based on $R_{i,c,t}$ and form value weight quintile portfolios. I go long stocks that have had positive commodity news and short those having negative commodity news. I hold the portfolio over the following month and rebalance monthly.

Stocks owned by investors who do not pay attention to the commodity sector (indicated by the "Low" $RAWATTN^n$ category in Table 1.12) have a significant inefficiency with respect to the commodity sector. The $5-1$ long/short portfolio formed within this group of stocks generates an alpha of approximately 1.3% per month (N-W tstat of 3.64) vs. the $5-1$ portfolio formed within the "High" attention category that generates a statistically insignificant alpha of .2% per month (N-W tstat of .32). I form the same quintile portfolios in the residualized version of the commodity attention metric, $ATTN^n$ and present results in Table 1.13. The $5-1$ portfolio in the "Low" attention category generates an alpha of 1.4% per month (N-W

Table 1.12: News Attention - $RAWATTN^n$

At the end of each month stocks are sorted into low and high attention stocks using $RAWATTN^n$ and then into value weight quintile portfolios using $R_{i,c,t}$. Stocks owned by investors who ignore the commodity market are inefficient with respect to commodity news.

$RAWATTN^n$	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Low	1	-0.193	**-0.760**	**-0.892**	**-0.927**	-0.102	-0.773	0.026	1.108	0.059	0.366	0.045
			[-2.81]	**[-3.49]**	**[-3.67]**							
	2	0.551	0.082	0.022	0.099	0.344	-0.696	-1.731	0.876	-0.058	0.127	-0.097
			[0.31]	[0.08]	[0.38]							
	3	0.201	-0.266	-0.272	-0.134	0.120	-0.955	3.917	0.833	-0.126	-0.043	-0.175
			[-1.13]	[-1.22]	[-0.61]							
	4	0.603	0.142	-0.017	0.110	0.367	-0.265	-1.652	0.859	0.166	0.369	-0.160
			[0.55]	[-0.07]	[0.45]							
	5	1.000	**0.471**	0.366	0.383	0.577	-0.511	0.432	1.004	0.083	0.271	-0.021
			[2.06]	[1.66]	[1.68]							
	5-1	1.194	**1.231**	**1.258**	**1.310**	0.667	0.400	-1.266	-0.103	0.024	-0.095	-0.065
			[3.42]	**[3.56]**	**[3.64]**							
High	1	0.768	0.269	0.161	0.195	0.337	0.322	1.216	0.911	0.291	0.273	-0.042
			[0.64]	[0.40]	[0.48]							
	2	0.733	0.171	0.068	-0.074	0.330	0.203	0.139	1.067	0.375	0.336	0.180
			[0.56]	[0.19]	[-0.20]							
	3	1.017	**0.578**	0.480	0.333	0.552	-0.034	-1.266	0.877	0.145	0.324	0.187
			[2.32]	[1.82]	[1.17]							
	4	0.855	0.433	0.263	0.192	0.483	-0.366	0.139	0.852	0.187	0.485	0.090
			[1.52]	[1.00]	[0.72]							
	5	1.119	0.621	0.464	0.372	0.533	-0.201	-0.243	0.967	0.339	0.459	0.116
			[1.78]	[1.43]	[1.10]							
	5-1	0.351	0.352	0.303	0.178	0.148	-0.443	0.015	0.056	0.049	0.186	0.159
			[0.69]	[0.56]	[0.32]							
Low - High	5-1	0.842	0.879	0.956	**1.132**	0.340	-0.001	0.393	-0.160	-0.025	-0.280	-0.224
			[1.66]	[1.73]	**[2.03]**							

tstat of 2.65) vs. the $5-1$ portfolio in the "High" attention category that has a statistically insignificant alpha of .3% per month (N-W tstat .59).

Both the residualized and the raw attention metrics deliver a sorting procedure that is able to categorize securities into those that are efficient and inefficient with respect to commodity news by discriminating among the sophisticated investors that own the stock. The alphas are economically and statistically significant and offer a challenge to the purely rational version of asset pricing that does not take into account the cognitive limitations of investors. There is another channel through which securities can fail to be fully efficient with respect to the commodity market: investors may not realize that they are affected by commodities because their exposure is nuanced. This idea is taken up in the next section.

1.3.3 Stock-Commodity Association Salience

Certain companies are inherently easy to recognize as those that are affected by commodity prices. For example, a November 2, 2012 New York Times article titled "Exxon and Shell Earnings, Hurt by Natural Gas, Are Helped by Refining" discussing the earnings of Exxon Mobil states: "Exxon Mobil and Royal Dutch Shell reported lackluster earnings on Thursday because of declining oil and natural gas production and weak domestic gas prices. ... Energy analysts were not surprised by the results since natural gas prices in the United States were roughly 30 percent lower than the year before." Clearly, investors are aware of the impact that the energy complex has on this particular company. Other companies may be connected to commodity prices in a more nuanced way that isn't clearly obvious to investors. Therefore, securities whose earnings are overtly related to commodity prices should be efficient in incorporating news from the commodity market into their prices; on the other hand, securities whose earnings have a more complicated connection to commodities may take time to fully incorporate commodity news. The most direct way to understand the salience of the association between a particular stock and commodity is by seeing how frequently news articles mention the two together in the same article (as a percentage of total articles about the company). If the commodity is a primary concern for investors, then it is likely mentioned

Table 1.13: Residualized News Attention - $ATTN^n$

At the end of each month stocks are sorted into low and high attention stocks using $ATTN^n$ (a residualized version of $RAWATTN^n$ using eq (1.16)) and then into value weight quintile portfolios using $R_{i,c,t}$. Stocks owned by investors who ignore the commodity market are inefficient with respect to commodity news.

$ATTN^n$	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Low	1	-0.204	-0.789	**-0.948**	**-0.846**	-0.092	-0.177	0.106	1.105	0.089	0.383	-0.130
			[-1.87]	**[-2.73]**	**[-2.49]**							
	2	0.270	-0.236	-0.351	-0.348	0.147	-1.159	1.433	0.963	0.135	0.306	-0.003
			[-0.57]	[-0.83]	[-0.78]							
	3	0.410	-0.117	-0.133	-0.030	0.229	-1.010	2.333	0.915	0.151	-0.002	-0.130
			[-0.43]	[-0.51]	[-0.11]							
	4	0.863	0.360	0.323	0.369	0.491	0.037	2.669	0.840	0.514	0.079	-0.059
			[1.29]	[1.24]	[1.45]							
	5	1.213	**0.682**	0.520	0.563	0.647	-0.464	0.027	0.999	0.256	0.414	-0.055
			[2.31]	[1.78]	[1.91]							
	5-1	1.417	**1.470**	**1.468**	**1.409**	0.660	-0.714	3.291	-0.106	0.167	0.032	0.075
			[2.78]	**[2.90]**	**[2.65]**							
High	1	0.564	0.055	0.006	0.031	0.270	0.218	1.250	0.948	0.005	0.119	-0.031
			[0.16]	[0.02]	[0.09]							
	2	0.640	0.145	0.100	-0.026	0.331	0.152	-0.408	0.947	0.140	0.175	0.160
			[0.49]	[0.30]	[-0.08]							
	3	0.656	0.275	0.173	0.126	0.410	0.228	-1.569	0.755	0.066	0.291	0.060
			[1.05]	[0.70]	[0.50]							
	4	0.708	0.296	0.125	0.021	0.432	-0.468	0.575	0.885	-0.057	0.502	0.131
			[1.18]	[0.61]	[0.10]							
	5	1.034	**0.555**	0.396	0.318	0.527	-0.316	-0.194	0.955	0.177	0.459	0.099
			[1.77]	[1.35]	[1.07]							
	5-1	0.471	0.501	0.390	0.287	0.201	-0.467	-1.488	0.007	0.172	0.340	0.129
			[1.07]	[0.80]	[0.59]							
Low - High	5-1	0.946	**0.970**	**1.078**	**1.121**	0.491	0.128	-0.673	-0.113	-0.005	-0.309	-0.055
			[2.36]	**[2.33]**	**[2.31]**							

very frequently in articles discussing the company.

To examine this degree of salience, I use news articles from the Financial Times, New York Times and Wall Street Journal (searched using Factiva). For each company that appears in my sample (associated with a particular commodity sector using the procedure of Section 1.2.2), I split the time period (from the first date it appears to the last date) into five year intervals and search for the company name to determine how many articles are written about that company during a particular five year period, $CONEWS_{i,t}$. The online Appendix provides details on how I process the company names from CRSP to retrieve the most relevant matches. I convert the number of articles to daily units (dividing $CONEWS_{i,t}$ by the number of days in the time interval) since not all time intervals are going to be exactly five years. For example, some companies may only exist for a year and to compare their news coverage to companies that exist for five years, one must scale by time since companies existing for a longer duration will have more articles written about them. I also search for articles that contain the company name and at least one name of a commodity from its associated commodity sector, $CMDTYCONEWS_{i,c,t}$, during the same time interval (once again converted to daily units). The names of the commodities are from Table 1.1 and the online Appendix provides details on the exact construction of search strings.

For ease of exposition, I would like to provide an example of this procedure using Exxon Mobil as the example company and Energy as the commodity sector. Exxon Mobil first appears in the sample on 05/31/1986 and remains until 12/31/2012. I split this period into five year intervals: (05/31/1986, 05/31/1991), (05/31/1991, 05/31/1996), (05/31/1996, 12/31/1999), (12/31/1999, 12/31/2004), (12/31/2004, 12/31/2009), (12/31/2009, 12/31/2012). Note that the 1996 - 1999 interval is only 3 years: the reason for this is that prior to that date Exxon Mobil was known as Exxon Corp. and it merged with Mobil at the end of 1999, thus becoming Exxon Mobil Corp. This is important because this name change causes a change in the search string used for news processing as well as highlighting why it is important to convert all results to daily units. This company is associated with the Energy sector and thus its particular commodity keyword set is ("Brent", "Crude Oil", "Gasoil", "Heating Oil", "Oil", "Natural Gas", "Gasoline", "Gas", "WTI", "West Texas Intermediate"). The processed company name is "exxon corp" prior to the merger and "exxon mobil" after the merger. Therefore, for the (12/31/2009, 12/31/2012) period the search string without keywords typed into Factiva is:

> exxon mobil and date from 12/31/2009 to 12/31/2012 and (rst=FTFT or rst=J or rst=NYTF)

and this yields 1068 matches. The string including the keywords is:

> exxon mobil and (Brent or Crude Oil or Gasoil or Heating Oil or Oil or Natural Gas or Gasoline or Gas or WTI or West Texas Intermediate) and date from 12/31/2009 to 12/31/2012 and (rst=FTFT or rst=J or rst=NYTF)

which yields 890 matches. Both of these numbers are then converted to daily units by dividing the match number by 1095 - the number of days in the time period. The online Appendix provides details of the general string construction procedure.

It is important to understand the validity of search results that appear: each additional news article that is written about a company increases the information availability about this company but at a decreasing rate. In other words, in the case of companies that have thousands of articles discussing them, additional articles are unlikely to provide the same amount of marginal information as they would for companies that only have tens of articles (for widely covered companies, news sources tend to simply report the same information in different form). This is analogous to analyst coverage for companies: the value of each additional analyst covering a company decreases as the total number of analysts covering the company grows. To properly account for decreasing marginal value of analysts, Hong et al. (2000) apply the *log* transform to the number of analysts covering a company. I apply their logic to the number of articles released about a company. Specifically, I define

$$conews_{i,t} \equiv log(1 + CONEWS_{i,t}) \tag{1.18}$$

where $CONEWS_{i,t}$ is the number of articles released about a particular company during time period t. Similarly, I define

$$cmdtyconews_{i,c,t} \equiv log(1 + CMDTYCONEWS_{i,c,t}) \tag{1.19}$$

where $CMDTYCONEWS_{i,c,t}$ is the number of articles that mention company i and commodity sector c together in the same article during time period t. Furthermore, certain periods may have more articles discussing commodities than other periods (for geopolitical reasons, for example). However, I am interested in a relative ranking among companies regarding their commodity salience. Therefore, I Z-Score $conews_{i,t}$ and $cmdtyconews_{i,c,t}$ within each time period. This provides me a relative ranking regarding how much news coverage each company receives during each time period with and without its associated commodities. Finally, I define the commodity salience for a particular stock i, commodity sector c and time period t as the proportion of news stories that mention the commodity and the company as a total fraction of news stories that mention the company:

$$cs_{i,c,t} \equiv \frac{cmdtyconews_{i,c,t}}{conews_{i,t}} \tag{1.20}$$

Companies may have multiple commodity sectors associated with them, as explained in Section 1.2.2, so I follow the same procedure as in Section 1.3.2 to collapse this information to the company level. Specifically, for each company i I define the level of salience to be a $|\beta_{i,c,t}|$ weighted average of $cs_{i,c,t}$:

$$RAWCS_{i,t} \equiv \sum_{c=1}^{C} w_{i,c,t} cs_{i,c,t} \tag{1.21}$$

where C_i is the number of commodities that i has exposure to at time t and $w_{i,c,t} \equiv \frac{|\beta_{i,c,t}|}{\sum_{c=1}^{C_i} |\beta_{i,c,t}|}$. In words: if a stock's returns have a large exposure to Metals but only a small exposure to Ag, then the cs of Metals should be much more important in our understanding of whether investors properly associate stock i with the commodity sector.

As was done previously, it is prudent to check if $RAWCS_{i,t}$ has loadings on any of the controls (size, book-to-market, etc.) used to residualize $RAWATTN^s$ and $RAWATTN^n$. To do this I follow the same methodology by running monthly Fama-MacBeth regressions of

Table 1.14: Commodity Salience Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (1.22). The $RAWCS$ metric does not have a significant loading on any of the variables known to influence equity returns or accentuate return anomalies.

	$log(BM_{i,t})$	$log(ME_{i,t})$	$IO_{i,t}$	$BREADTH_{i,t}$	$\beta_{i,m,t}$	$R^e_{i,t}$	$R^e_{i,t-12\rightarrow t-1}$	$IV_{i,t}$	(Intercept)	R^2	N
(1)	0.214	-0.088	5.105	-0.019					2.257	0.94%	441.638
	[0.86]	[-0.22]	[0.89]	[-0.70]					[0.49]		
(2)	0.259	-0.137	4.616	-0.018	0.207				2.720	1.23%	441.613
	[1.05]	[-0.35]	[0.80]	[-0.64]	[0.84]				[0.57]		
(3)	0.152	-0.109	5.198	-0.021	0.227	-0.029	-0.004		2.453	1.92%	436.238
	[0.55]	[-0.29]	[0.84]	[-0.74]	[0.89]	[-0.87]	[-0.60]		[0.53]		
(4)	0.067	-0.322	5.168	-0.018	0.263	-0.023	-0.006	-0.022	5.692	2.16%	436.238
	[0.24]	[-0.87]	[0.84]	[-0.68]	[1.04]	[-0.72]	[-0.82]	[-1.26]	[1.23]		

the form:

$$\begin{aligned} RAWCS_{i,t} &= \theta_{0,t} + \theta_{bm,t}log(BM)_{i,t} + \theta_{me,t}log(ME_{i,t}) + \theta_{io,t}IO_{i,t} + \\ &+ \theta_{br,t}BREADTH_{i,t} + \theta_{mb,t}\beta_{m,i,t} + \theta_{r,t}R_{i,t} + \theta_{mom,t}R_{i,t-12\rightarrow t-1} + \\ &+ \theta_{iv,t}IV_{i,t} + \varepsilon^{cs}_{i,t} \qquad (1.22) \end{aligned}$$

and defining the residualized version of commodity salience as

$$CS_{i,t} \equiv \varepsilon^{cs}_{i,t} \qquad (1.23)$$

Table 1.14 presents the loadings of $RAWCS_{i,t}$ on the common variables known to influence equity returns. It is not evident that any of the control variables have a significant correlation with $RAWCS$, however, I remain prudent by presenting results using $RAWCS$ and CS.

Each month (belonging to one of the non-overlapping five year periods) I sort companies into low and high salience categories based on $RAWCS_{i,t}$ ($CS_{i,t}$); within each salience category I sort stocks into value weighted quintiles based on $R_{i,c,t}$ and go long (short) stocks that have positive (negative) commodity return news[10]. If investors have trouble understanding

[10]Strictly speaking this introduces forward looking information into the sorting procedure since sorting stocks based on $CS_{i,t}$ includes news counts from the five year period that month happens to fall into. The main reason this is done is to avoid the computational cost of searching for news articles each month: this would require approximately 60 times the number of Factiva searches currently done. There is also little reason to believe this simplification would bias the results as a 5 year period is long time frame and having many news stories associated with a company is not correlated with a positive or negative return relationship. Note

that certain companies are affected by commodities then the trading strategy in the low salience category should be highly profitable while it should produce no alpha in the high salience category; this is exactly what I find. Tables 1.15 and 1.16 provide the results of this experiment.

Examining Table 1.15, the results are striking: companies in the low salience category are inefficient in incorporating commodity information allowing one to generate approximately 1.1% of four factor alpha per month (N-W t-statistic 2.47) while companies that have a high commodity salience generate much less (statistically insignificant) alpha: .4% per month (t-statistic of .74). The difference between these two categories of .8% per month is significant economically as well as statistically (N-W t-statistic of 1.96). Table 1.16 presents results that are largely similar: portfolios in the low salience category generate approximately 1% of alpha per month vs .3% in the high salience category.

This highlights a particular source of market inefficiency: investors sometimes do not fully understand all the factors that can influence a company either because the company is too complicated and there may be too many factors to take into account, as explored by Cohen & Lou (2012), or because the connection to the information source may be nuanced.

1.3.4 Time Varying Cognitive Burden

Section 1.3.1 and 1.3.2 examined how inattention varies cross-sectionally showing that stocks owned by attentive investors are more efficient than stocks owned by inattentive investors. However, information burden for investors varies over time: some periods - for example earnings season - tend to be a particularly busy and cognitively constrained time. When companies are reporting significant amounts of idiosyncratic information then investors must keep up with many different sources of news. At these times, inattention to commodities should be exacerbated (for example PEAD is larger at times of high cognitive burden as shown by Hirshleifer et al. (2009)).

---

that while news counts have forward looking information, there is absolutely no forward looking information in $R_{i,c,t}$.

Table 1.15: Commodity Salience Using $RAWCS$

Stocks are sorted into low and high salience categories based on $RAWCS_{i,t}$ (1.21). Within each salience category I sort stocks into value weighted quintiles based on $R_{i,c,t}$ and go long (short) stocks that have positive (negative) commodity return news. Equity data is from CRSP, news article data is from the Financial Times, Wall Street Journal and New York Times accessed using Factiva and commodity data is from Bloomberg. Securities that have a more obvious connection to commodities are quicker at incorporating commodity sector information into their returns.

$RAWCS$	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Low	1	-0.132	**-0.755**	**-0.776**	**-0.677**	-0.057	-0.185	0.799	1.115	0.033	0.014	-0.126
			[-2.18]	**[-2.23]**	**[-2.05]**							
	2	0.488	-0.042	-0.080	-0.131	0.249	-0.397	1.973	1.027	-0.122	0.124	0.064
			[-0.13]	[-0.26]	[-0.40]							
	3	0.466	-0.014	-0.102	-0.087	0.256	-0.709	0.738	0.890	0.175	0.227	-0.018
			[-0.05]	[-0.38]	[-0.32]							
	4	0.992	0.529	0.433	0.314	0.529	0.136	1.693	0.855	0.495	0.307	0.151
			[1.62]	[1.41]	[1.07]							
	5	1.223	**0.687**	0.527	0.456	0.591	-0.167	-0.784	1.030	0.338	0.458	0.089
			[2.19]	[1.80]	[1.64]							
	5-1	1.354	**1.442**	**1.303**	**1.133**	0.556	-0.196	-1.165	-0.084	0.305	0.444	0.215
			[3.03]	**[2.74]**	**[2.47]**							
High	1	0.605	0.140	0.041	-0.002	0.310	0.229	-0.112	0.901	0.100	0.282	0.055
			[0.38]	[0.13]	[-0.01]							
	2	0.877	0.461	0.375	0.376	0.534	-0.050	-1.334	0.786	0.117	0.228	-0.001
			[1.84]	[1.50]	[1.49]							
	3	0.748	0.282	0.165	0.172	0.449	0.165	-0.571	0.836	0.455	0.311	-0.009
			[1.17]	[0.75]	[0.80]							
	4	0.702	0.199	0.076	0.072	0.414	-0.847	0.886	0.969	0.097	0.330	0.005
			[0.88]	[0.38]	[0.31]							
	5	0.953	0.504	0.361	0.344	0.556	-0.458	1.444	0.890	0.058	0.390	0.022
			[1.87]	[1.40]	[1.21]							
	5-1	0.347	0.365	0.320	0.346	0.169	-0.345	-0.616	-0.010	-0.042	0.108	-0.032
			[0.83]	[0.75]	[0.74]							
Low - High	5-1	1.007	**1.077**	**0.983**	**0.788**	0.500	0.182	-1.506	-0.074	0.347	0.335	0.248
			[2.71]	**[2.44]**	**[1.96]**							

Table 1.16: Commodity Salience Using CS

Stocks are sorted into low and high salience categories based on $CS_{i,t}$ (1.23). Within each salience category I sort stocks into value weighted quintiles based on $R_{i,c,t}$ and go long (short) stocks that have positive (negative) commodity return news. Equity data is from CRSP, news article data is from the Financial Times, Wall Street Journal and New York Times accessed using Factiva and commodity data is from Bloomberg. Securities that have a more obvious connection to commodities are quicker at incorporating commodity sector information into their returns.

CS	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
	1	0.165	-0.422	-0.491	-0.359	0.074	-0.021	-0.324	1.047	0.122	0.129	-0.167
			[-1.17]	[-1.31]	[-1.06]							
	2	0.349	-0.161	-0.191	-0.180	0.175	-1.208	3.372	0.939	0.039	0.076	-0.014
			[-0.46]	[-0.54]	[-0.48]							
	3	0.677	0.163	0.051	0.117	0.393	-0.608	-0.493	0.958	0.116	0.270	-0.083
Low			[0.62]	[0.21]	[0.48]							
	4	0.891	0.435	0.331	0.336	0.526	-0.399	2.223	0.819	0.415	0.277	-0.007
			[1.72]	[1.33]	[1.36]							
	5	1.285	**0.745**	0.615	0.626	0.628	-0.155	0.127	1.008	0.268	0.342	-0.013
			[2.33]	[1.95]	[1.89]							
	5-1	1.120	**1.167**	**1.106**	0.985	0.469	-0.021	-0.873	-0.040	0.145	0.213	0.154
			[2.40]	**[2.16]**	[1.92]							
	1	0.546	0.034	-0.035	-0.083	0.276	0.088	0.769	0.981	0.043	0.206	0.060
			[0.11]	[-0.12]	[-0.29]							
	2	0.657	0.259	0.161	0.120	0.396	-0.261	-1.792	0.786	0.046	0.278	0.052
			[0.92]	[0.59]	[0.42]							
	3	0.639	0.183	0.085	0.032	0.368	0.164	-1.443	0.863	0.234	0.285	0.067
High			[0.72]	[0.37]	[0.12]							
	4	0.844	**0.440**	0.302	0.289	0.556	-0.218	-1.909	0.805	0.055	0.375	0.016
			[2.02]	[1.53]	[1.43]							
	5	0.908	0.444	0.231	0.191	0.498	-0.203	-0.585	0.940	0.195	0.585	0.051
			[1.63]	[0.95]	[0.75]							
	5-1	0.363	0.410	0.267	0.274	0.174	-0.434	-0.456	-0.040	0.152	0.379	-0.009
			[1.06]	[0.70]	[0.71]							
Low - High	5-1	0.757	**0.757**	**0.840**	**0.711**	0.489	0.203	-1.769	0.001	-0.007	-0.166	0.163
			[2.41]	**[2.43]**	**[2.15]**							

To test this hypothesis, each month I compute a measure of cross-sectional return dispersion defined as the cross-sectional standard deviation of returns that month of all securities in CRSP:

$$XD_t \equiv \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(R_{i,t}-\overline{R_t}\right)^2} \tag{1.24}$$

Months where XD is high, are periods when stocks are behaving particularly differently from one another. In other words, there is a significant amount of idiosyncratic information being released. I categorize all trading months into high and low information burden periods. Then I create two trading strategies: one that trades only in high information burden (HIB) periods and another that trades only low information burden (LIB) periods. Specifically, at the end of month t, I sort stocks into value weight and equal weight quintile portfolios (as in Section 1.2.3) and go long (short) stocks with positive (negative) commodity sector news. If XD_t falls into the high (low) category then the resulting return in month $t+1$ is attributed to the HIB (LIB) strategy. I then run a factor attribution regression for the HIB and LIB strategy using only periods in which there is trading in each. By construction, half of the months in the sample will be HIB periods and half will be LIB periods. Another way to think about this exact situation is simply running a factor attribution regression for the basic quintile sort strategy of section 1.2.3 that includes indicator variables on all independent variables (including the constant) taking the value 1 (0) if the *previous* period was an HIB (LIB) period. Notably, I allow factor exposures of each strategy to be different to make sure that the results are not driven by regime changes in factor exposure.

Tables 1.17 and 1.18 report the results of this hypothesis test. The difference between HIB and LIB periods is large and significant in value weight and equal weight portfolios. The commodity inattention strategy has alphas of over 2% per month in HIB periods and alphas that are indistinguishable from 0 in LIB periods. During periods when investors are burdened with too much idiosyncratic information, they are more likely to ignore the commodity market.

Table 1.17: Commodity Inattention During Times of High and Low Information Burden VW

This table presents results of two trading strategies: one that takes advantage of commodity inattention during high information burden (HIB) periods and the other that trades only during low information burden (LIB) periods. The sample is split into these categories based on XD_t so that each month falls into HIB or LIB. Section 1.3.4 presents details of the construction. Equity data is from CRSP and commodity data is from Bloomberg.

Information Burden	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
	1	-0.937	**-1.129**	**-1.168**	**-1.150**	-0.424	0.046	-0.653	0.962	-0.039	0.144	-0.026
			[-2.25]	**[-2.20]**	**[-2.29]**							
	2	-0.276	-0.406	-0.400	-0.374	-0.165	-0.987	0.487	0.659	-0.105	0.098	-0.036
			[-0.88]	[-1.03]	[-0.93]							
	3	0.173	0.007	-0.178	-0.169	0.108	-0.323	-2.239	0.824	0.191	0.276	-0.012
High			[0.02]	[-0.60]	[-0.59]							
	4	0.255	0.123	-0.032	-0.015	0.178	0.072	-2.494	0.659	0.134	0.259	-0.023
			[0.46]	[-0.11]	[-0.05]							
	5	1.222	**1.064**	**0.853**	**0.911**	0.646	0.387	-1.526	0.812	0.097	0.437	-0.080
			[2.67]	**[2.02]**	**[2.64]**							
	5-1	2.159	**2.193**	**2.021**	**2.060**	0.930	-0.419	-1.703	-0.151	0.136	0.293	-0.054
			[3.40]	**[3.07]**	**[3.25]**							
	1	1.191	0.285	0.435	0.524	0.701	0.262	1.298	0.973	0.389	0.062	-0.192
			[0.94]	[1.30]	[1.63]							
	2	1.015	0.233	0.211	0.220	0.746	0.241	-1.179	0.873	0.039	0.107	-0.020
			[1.00]	[0.80]	[0.80]							
	3	1.286	0.475	**0.498**	0.472	0.887	-0.705	3.770	0.880	0.183	0.174	0.056
Low			[1.95]	**[1.96]**	[1.84]							
	4	1.409	**0.601**	**0.483**	0.394	0.978	-0.388	1.093	0.918	0.102	0.473	0.191
			[2.36]	**[1.99]**	[1.52]							
	5	1.134	0.136	0.113	-0.053	0.635	-1.201	2.347	1.070	0.332	0.519	0.354
			[0.37]	[0.32]	[-0.15]							
	5-1	-0.057	-0.149	-0.322	-0.577	-0.031	0.023	-2.091	0.098	-0.057	0.457	0.546
			[-0.26]	[-0.56]	[-1.04]							
High - Low	5-1	2.216	**2.342**	**2.342**	**2.637**	0.961	-0.442	0.388	-0.248	0.194	-0.164	-0.600
			[2.76]	**[2.62]**	**[3.05]**							

Table 1.18: Commodity Inattention During Times of High and Low Information Burden EW

This table presents results of two trading strategies: one that takes advantage of commodity inattention during high information burden (HIB) periods and the other that trades only during low information burden (LIB) periods. The sample is split into these categories based on XD_t so that each month falls into HIB or LIB. Section 1.3.4 presents details of the construction. Equity data is from CRSP and commodity data is from Bloomberg.

Information Burden	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
High	1	-0.584	-0.819	**-1.277**	**-1.126**	-0.232	-0.007	-1.012	1.086	0.515	0.583	-0.206
			[-1.27]	**[-2.59]**	**[-2.15]**							
	2	0.267	0.082	-0.403	-0.265	0.132	-0.474	-1.787	0.793	0.681	0.469	-0.189
			[0.15]	[-0.90]	[-0.66]							
	3	0.885	0.675	0.125	0.187	0.418	-0.079	-1.642	0.901	0.919	0.395	-0.085
			[1.69]	[0.37]	[0.68]							
	4	1.079	0.885	0.386	0.553	0.541	-0.200	-1.851	0.775	0.798	0.356	-0.229
			[1.88]	[1.06]	[1.58]							
	5	1.981	**1.777**	**1.223**	**1.450**	0.849	0.415	-0.750	0.831	0.771	0.511	-0.309
			[3.35]	**[3.29]**	**[3.10]**							
	5-1	2.565	**2.596**	**2.500**	**2.576**	1.163	0.160	0.223	-0.255	0.257	-0.072	-0.104
			[4.15]	**[4.11]**	**[3.94]**							
Low	1	0.506	-0.413	-0.113	-0.036	0.281	-0.103	0.254	0.945	0.956	0.372	-0.165
			[-1.18]	[-0.30]	[-0.11]							
	2	0.359	-0.613	-0.415	-0.314	0.206	0.430	1.249	1.050	0.816	0.470	-0.215
			[-1.88]	[-1.47]	[-1.01]							
	3	0.661	-0.168	0.119	0.116	0.447	-0.732	1.903	0.824	0.895	0.348	0.006
			[-0.55]	[0.42]	[0.43]							
	4	1.123	0.123	0.233	0.303	0.614	0.101	0.358	1.108	0.739	0.619	-0.151
			[0.34]	[0.83]	[0.98]							
	5	0.898	-0.208	-0.129	-0.220	0.446	-0.530	0.484	1.187	0.725	0.723	0.195
			[-0.52]	[-0.34]	[-0.62]							
	5-1	0.393	0.205	-0.016	-0.184	0.221	-0.211	-1.719	0.242	-0.230	0.351	0.360
			[0.47]	[-0.03]	[-0.36]							
High - Low	5-1	2.172	**2.390**	**2.516**	**2.760**	0.942	0.371	1.942	-0.497	0.487	-0.422	-0.464
			[3.01]	**[3.08]**	**[3.30]**							

1.3.5 Channel Uniqueness

I have presented several channels through which investors may be inattentive to information in the commodity market: they may be ignoring particular portions of their portfolio, they may be ignoring the commodity market as a source of information, they may misunderstand the dependence some equities have on the commodity market or they may simply be overwhelmed with a large amount of idiosyncratic information and thus not have the capacity to fully process everything that is relevant. I have also shown that the metrics I use to proxy for these channels are robust to a set of controls commonly associated with equity anomalies. However, it is important to check that they truly represent unique channels affecting attention. The methodology I have used until this point (sorting into portfolios) is useful because of its non-parametric nature allowing me to capture any non-linearities inherent in the relationship between the intertemporal correlation of $R_{i,c,t}$ and $R^e_{i,t+1}$. However, it is difficult to test multiple effects in that framework; the typical way to combine cross-sectional effects (and the methodology I use in the robustness section of the paper, Section 1.4) is Fama-MacBeth. However, one of my metrics - time varying cognitive burden - is strictly a time series variable and thus it is more natural to test results in a pooled panel regression to estimate these effects. To focus attention of the cross-sectional metrics on the cross-sectional influence on stocks, I use monthly (time) dummy variables as would be done in a Fama-MacBeth estimation except now I am free to interact XD_t with $R_{i,c,t}$. I briefly detail it here to set specifics: the model driving returns:

$$\begin{aligned} R^e_{i,t+1} &= a_{t+1} + \left(\beta_0 + \beta_1 \mathbf{1}_{HIB,t} + \beta_2 RAWCS_{i,t} + \beta_3 RAWATTN^n_{i,t} + \beta_4 RAWATTN^s_{i,t}\right) R_{i,c,t} + \quad (1.25) \\ &\quad + \eta_1 \mathbf{1}_{HIB,t} + \eta_2 RAWCS_{i,t} + \eta_3 RAWATTN^n_{i,t} + \eta_4 RAWATTN^s_{i,t} + \boldsymbol{\theta'} \boldsymbol{X_{i,t}} + \varepsilon_{i,t+1} \\ &= a_{t+1} + \boldsymbol{\beta'} \left[R_{i,c,t}\ \boldsymbol{\Gamma_{i,t}} R_{i,c,t}\right]^T + \boldsymbol{\eta'} \boldsymbol{\Gamma_{i,t}} + \boldsymbol{\theta'} \boldsymbol{X_{i,t}} + \varepsilon_{i,t+1} \quad (1.26) \end{aligned}$$

where $\mathbf{1}_{HIB,t}$ is an indicator variable that is 1 (0) if XD_t falls into the high (low) burden period - half of the periods are high and half are low as in Section 1.3.4; $RAWCS_{i,t}$ is the commodity salience associated with stock i at time t from Section 1.3.3, $RAWATTN^n_{i,t}$ is the attention being paid to the commodity sector by the owners of i at time t as in Section 1.3.2,

$RAWATTN^s_{i,t}$ is the amount of attention being paid to stock i by sophisticated investors at time t as in Section 1.3.1, and $\boldsymbol{X_{i,t}}$ is a vector of control variables:

$$\boldsymbol{X_{i,t}} = [log(BM_{i,t}),\ log(ME_{i,t}),\ IO_{i,t},\ BREADTH_{i,t},\ \beta_{m,i,t},\ R_{i,t-1},\ R_{i,t-13\to t-2},\ IV_{i,t}] \tag{1.27}$$

as in Eq 1.11. Regression (1.25) allows me to control for a set of variables that affect equity returns, control for the effect of the attention modifying variables on equity returns and, most importantly, determine how they affect the intertemporal relationship between $R_{i,c,t}$ and $R^e_{i,t+1}$ by observing the estimates of β_1 through β_4. By including the time dummies (a_{t+1}), this regression focuses on estimating the cross-sectional effects of the return modifying variables (except time varying cognitive burden which clearly has no cross-sectional effects). To estimate (1.25), I take the cross-sectional mean of the equation (noting that the cross-sectional mean of a variable that is constant for a particular time period is simply that variable):

$$\overline{R^e}_{t+1} = a_{t+1} + \boldsymbol{\beta'} \left[\overline{R}_{c,t}\ \overline{\boldsymbol{\Gamma_{i,t}} R_{i,c,t}}\right]^T + \boldsymbol{\eta'}\overline{\boldsymbol{\Gamma_{i,t}}} + \boldsymbol{\theta'}\overline{\boldsymbol{X_{i,t}}} + \overline{\varepsilon}_{t+1} \tag{1.28}$$

Subtracting (1.28) from (1.25) yields:

$$\widetilde{R^e}_{i,t+1} = \boldsymbol{\beta'} \left[\widetilde{R}_{i,c,t}\ \widetilde{\boldsymbol{\Gamma_{i,t}} R_{i,c,t}}\right]^T + \boldsymbol{\eta'}\widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{i,t}} + \boldsymbol{\theta'}\widetilde{\boldsymbol{X}}_{\boldsymbol{i,t}} + \widetilde{\varepsilon}_{i,t+1} \tag{1.29}$$

where the tilde over the variable indicates the cross-sectional demeaned variable for that time period, for example $\boldsymbol{\widetilde{X}_{i,t} = X_{i,t} - \overline{X}_t}$. Equation (1.29) is estimated by OLS as usual. This is the usual centering method for estimating a regression with time dummies. In addition to centering, I scale all my variables by their cross-sectional standard deviation so that the obtained estimates can be evaluated (it is much easier to understand the regression coefficient when each variable is in units of standard deviation). Standard errors are corrected for correlation by double clustering on time and stock. Table 1.19 presents the results of this regression.

The first set of regressions labeled (1) verify that the regression coefficient of $R_{i,c,t}$ is positive and highly significant as we would expect given the exhaustive results presented

Table 1.19: Panel Regressions of Inattention Channels

To verify the uniqueness of each inattention channel, I run a pooled panel regression, eq (1.29), that includes interactions of all the attention modifying variables with $R_{i,c,t}$ as well as a set of control variables. The channels presented have independent effects on the ability of $R_{i,c,t}$ to predict $R^e_{i,t+1}$.

	$R_{i,c,t}$	$R_{i,c,t} \cdot RAWATTN^n_{i,t}$	$R_{i,c,t} \cdot RAWCS_{i,t}$	$R_{i,c,t} \cdot 1_{HIB}$	$R_{i,c,t} \cdot RAWATTN^s_{i,t}$	$RAWATTN^s_{i,t}$	$RAWATTN^n_{i,t}$	$RAWCS_{i,t}$	$R^e_{i,t-12\to t-1}$	$bm_{i,t}$	$log(ME_{i,t})$	$R^e_{i,t}$	$\beta_{i,m,t}$	$IV_{i,t}$	R^2	N
(1)	**0.04**								**0.02**	**0.02**	0.01				0.22%	133,820
	[3.357]								**[2.877]**	**[2.770]**	[1.271]					
	0.04								**0.02**	**0.02**	0.01	-0.01			0.23%	133,820
	[3.510]								**[2.839]**	**[2.620]**	[1.329]	[-1.093]				
	0.04								**0.02**	**0.02**	0.01	-0.01	-0.01		0.23%	133,819
	[3.508]								**[2.861]**	**[2.600]**	[1.339]	[-1.072]	[-1.024]			
	0.04								**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.34%	133,819
	[3.507]								**[2.674]**	**[2.051]**	[-1.298]	[-0.809]	[-0.534]	**[-4.742]**		
(2)	**0.05**	**-0.01**					0.00		**0.02**	**0.02**	0.01				0.26%	133,820
	[3.813]	**[-2.577]**					[0.060]		**[2.873]**	**[2.718]**	[1.257]					
	0.05	**-0.01**					0.00		**0.02**	**0.02**	0.01	-0.01			0.26%	133,820
	[3.970]	**[-2.573]**					[0.050]		**[2.836]**	**[2.571]**	[1.314]	[-1.075]				
	0.05	**-0.01**					0.00		**0.02**	**0.02**	0.01	-0.01	-0.01		0.27%	133,819
	[3.969]	**[-2.579]**					[0.028]		**[2.857]**	**[2.548]**	[1.324]	[-1.054]	[-1.028]			
	0.05	**-0.01**					0.00		**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.37%	133,819
	[3.957]	**[-2.603]**					[0.466]		**[2.681]**	**[2.035]**	[-1.311]	[-0.784]	[-0.519]	**[-4.798]**		
(3)	**0.04**		**0.00**					0.00	**0.02**	**0.02**	0.01				0.23%	133,820
	[3.360]		**[-2.343]**					[-0.223]	**[2.873]**	**[2.771]**	[1.272]					
	0.04		**0.00**					0.00	**0.02**	**0.02**	0.01	-0.01			0.23%	133,820
	[3.513]		**[-2.329]**					[-0.226]	**[2.836]**	**[2.621]**	[1.331]	[-1.090]				
	0.04		**0.00**					0.00	**0.02**	**0.02**	0.01	-0.01	-0.01		0.23%	133,819
	[3.511]		**[-2.334]**					[-0.225]	**[2.857]**	**[2.601]**	[1.341]	[-1.069]	[-1.025]			
	0.04		**0.00**					0.00	**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.34%	133,819
	[3.510]		**[-2.375]**					[-0.212]	**[2.671]**	**[2.051]**	[-1.296]	[-0.806]	[-0.534]	**[-4.744]**		
(4)	-0.05			**0.06**					**0.02**	**0.02**	0.01				0.31%	133,820
	[-1.297]			**[2.552]**					**[3.000]**	**[2.834]**	[1.276]					
	-0.05			**0.06**					**0.02**	**0.02**	0.01	-0.01			0.31%	133,820
	[-1.265]			**[2.548]**					**[2.962]**	**[2.682]**	[1.334]	[-1.077]				
	-0.05			**0.06**					**0.02**	**0.02**	0.01	-0.01	-0.01		0.31%	133,819
	[-1.266]			**[2.549]**					**[2.982]**	**[2.663]**	[1.343]	[-1.056]	[-1.028]			
	-0.05			**0.06**					**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.42%	133,819
	[-1.246]			**[2.527]**					**[2.795]**	**[2.117]**	[-1.266]	[-0.794]	[-0.536]	**[-4.722]**		
(5)	**0.04**				**-0.01**	0.00			**0.02**	**0.02**	0.01				0.26%	133,820
	[3.603]				**[-5.418]**	[-1.166]			**[2.882]**	**[2.758]**	[1.450]					
	0.04				**-0.01**	-0.01			**0.02**	**0.02**	0.01	-0.01			0.26%	133,820
	[3.759]				**[-5.429]**	[-1.217]			**[2.845]**	**[2.607]**	[1.519]	[-1.103]				
	0.04				**-0.01**	0.00			**0.02**	**0.02**	0.01	-0.01	-0.01		0.26%	133,819
	[3.757]				**[-5.448]**	[-1.197]			**[2.866]**	**[2.587]**	[1.522]	[-1.082]	[-1.025]			
	0.04				**-0.01**	0.01			**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.37%	133,819
	[3.753]				**[-5.577]**	[1.424]			**[2.661]**	**[2.054]**	[-1.607]	[-0.781]	[-0.540]	**[-4.903]**		
(6)	-0.04	**-0.01**	**0.00**	**0.06**	**-0.01**	-0.01	0.00	0.00	**0.03**	**0.02**	0.01				0.36%	133,820
	[-1.015]	**[-2.165]**	**[-2.406]**	**[2.486]**	**[-5.092]**	[-1.452]	[0.306]	[-0.375]	**[3.000]**	**[2.795]**	[1.489]					
	-0.04	**-0.01**	**0.00**	**0.06**	**-0.01**	-0.01	0.00	0.00	**0.02**	**0.02**	0.01	-0.01			0.37%	133,820
	[-0.984]	**[-2.162]**	**[-2.394]**	**[2.482]**	**[-5.104]**	[-1.509]	[0.298]	[-0.376]	**[2.962]**	**[2.645]**	[1.555]	[-1.069]				
	-0.04	**-0.01**	**0.00**	**0.06**	**-0.01**	-0.01	0.00	0.00	**0.03**	**0.02**	0.01	-0.01	-0.01		0.37%	133,819
	[-0.984]	**[-2.167]**	**[-2.397]**	**[2.482]**	**[-5.126]**	[-1.484]	[0.272]	[-0.375]	**[2.981]**	**[2.624]**	[1.556]	[-1.048]	[-1.021]			
	-0.04	**-0.01**	**0.00**	**0.06**	**-0.01**	0.00	0.00	0.00	**0.02**	**0.01**	-0.01	-0.01	0.00	**-0.04**	0.48%	133,819
	[-0.979]	**[-2.163]**	**[-2.387]**	**[2.471]**	**[-5.339]**	[1.428]	[0.546]	[-0.387]	**[2.784]**	**[2.116]**	[-1.525]	[-0.746]	[-0.521]	**[-4.901]**		

thus far. The magnitude is not significantly affected by any of the presented controls (more controls will be added in Section 1.4). A one standard deviation increase in commodity news causes a .04 standard deviation increase in $R^e_{i,t+1}$ relative to the average stock return that period after controlling for momentum, book-to-market, size, market β, and idiosyncratic volatility. Examining set (2) of regressions that include $RAWATTN^n$ attention modification variable one can see that the estimate is negative and highly significant, as expected, showing that the relationship between $R_{i,c,t}$ and $R^e_{i,t+1}$ becomes smaller as $RAWATTN^n$ increases. In words, the more attention that is paid to commodity news by owners of i, the more efficiently commodity news is priced into i this month without spilling over into next month. The magnitude is economically significant as well: a one standard deviation increase in $RAWATTN^n$ causes a 20% decrease in the effect of $R_{i,c,t}$ into next period. Set (3) includes $RAWCS$ as the attention variable once again with a negative slope: a larger proportion of news stories that mention stock i and its associated commodities the clearer the connection becomes for i's investors and the more efficiently i is priced. Set (4) includes the dummy variable for high information burden months: the coefficient estimate is positive in this case suggesting that the amount of commodity news that gets priced (inefficiently) the following month increases during periods with high informational burden. The coefficient estimate is quite large suggesting that all of the time series effects of inattention occur during high informational burden periods. Set (5) includes $RAWATTN^s$ interacted with $R_{i,c,t}$: the estimate is negative and significant. More attention paid to stock i by it's sophisticated owners yields higher efficiency in pricing and therefore a lower lag in incorporating information from the commodity sector.

In all of these regressions we see that the attention modifying variables alone (without interacting them with $R_{i,c,t}$) have negligible estimates as we'd expect: high attention paid to a particular stock has no affect on returns by itself (only as a modifying variable for $R_{i,c,t}$). Finally, set (6) includes all of the attention modifying variables jointly: the results are unchanged, they all have an affect on the ability of $R_{i,c,t}$ to predict $R^e_{i,t+1}$. Furthermore, their magnitudes to not change when included jointly suggesting that they truly are independent sources of inattention. This is to be expected as the goal was to target disjoint portions of cognition, however, the panel regression verifies that this effort was successful.

1.4 Robustness

The previous sections have demonstrated that equities are slow to fully incorporate all available information from the commodity market into their price. I would like to show that this phenomenon is distinct from several other anomalies that are already known, is not due to lookahead bias in commodity information or staleness of small security prices, and isn't an artifact of the particular methodology I have used (categorizing commodities into sectors and using a p-value cutoff for selecting stocks).

1.4.1 Fama-MacBeth

To show that the underreaction to commodities is a unique phenomenon, I am going to use monthly Fama-MacBeth regressions with the stock universe selected using the methodology in Section 1.2.2 (as has been used in the rest of the study). There are several phenomena which serve as good candidates for robustness. First, there are the usual stock level candidates that have been known to produce anomalous returns in the past: stock reversals, size, value, and stock momentum. Stock reversals have been documented in the literature (for example Avramov et al. (2006)) and are particularly strong in smaller securities. Size, value and momentum anomalies are well studied and exposure to their respective factors has been controlled for in all the previous trading strategy attribution results. However, as Daniel & Titman (1997) show, characteristics can still play a role in pricing even after controlling for factor exposure therefore I include them as variables in the Fama-MacBeth regressions.

The next set of controls revolves around industries. Since individual stock commodity sector β is noisy, I have proxied for it using the β of the industry that the stock belongs to. This can cause my results to be driven by industry phenomena that are already known: various forms of industry momentum reported by Moskowitz & Grinblatt (1999) and intra-industry large to small stock information diffusion of Hou (2007). The industry momentum

Table 1.20: Fama-MacBeth Regressions

Fama-MacBeth regressions are run each month to control for various other known anomalies. Newey-West t-statistics are reported in brackets along with R^2 and average number of firms across the monthly regressions.

	$R_{i,c,t}$	$R^e_{i,t}$	$log(ME_{i,t})$	$R^e_{i,t-12\to t-1}$	$bm_{i,t}$	$IR^e_{i,t-11\to t}$	$IR^e_{i,t-12\to t-1}$	$IR^e_{i,t}$	$IR^e_{i,t-1}$	$LIR^e_{i,t}$	(Intercept)	R^2	N
(1)	**0.460**										**0.794**	2.01%	704.891
	[2.373]										**[1.971]**		
(2)	**0.396**	**-0.047**	-0.066	0.002	**0.362**	**0.023**					1.089	6.79%	704.891
	[2.239]	**[-8.047]**	[-1.091]	[0.902]	**[3.352]**	**[3.620]**					[1.096]		
(3)	**0.410**	**-0.047**	-0.064	0.002	**0.363**		0.011				1.042	6.83%	704.891
	[2.225]	**[-8.032]**	[-1.054]	[1.066]	**[3.338]**		[1.600]				[1.040]		
(4)	**0.405**	**-0.049**	-0.061	0.003	**0.368**			**0.107**			1.125	6.75%	704.891
	[2.109]	**[-8.405]**	[-1.055]	[1.113]	**[3.502]**			**[3.769]**			[1.170]		
(5)	**0.575**	**-0.047**	-0.066	0.002	**0.374**				0.029		1.099	6.77%	704.891
	[2.757]	**[-8.108]**	[-1.115]	[1.088]	**[3.539]**				[1.253]		[1.088]		
(6)	**0.401**	**-0.049**	-0.062	0.003	**0.367**					**0.105**	1.121	6.75%	704.891
	[2.082]	**[-8.389]**	[-1.059]	[1.115]	**[3.492]**					**[3.779]**	[1.163]		

controls are the last 12 month return of the industry that a particular stock belongs to ($IR^e_{i,t-11\to t}$), a 1 month lag of this return ($IR^e_{i,t-12\to t-1}$), last month's industry return ($IR^e_{i,t}$), and two month lag industry return ($IR^e_{i,t-1}$). I control for the effect of Hou (2007) by splitting each industry into large and small stocks (at that industry's median market capitalization) and creating a value weighted portfolio of large stock returns in that industry ($LIR^e_{i,t}$).

The left hand side of the regression is individual excess stock returns in my universe, $R^e_{i,t+1}$, and the predictive quantity I am concerned with is $R_{i,c,t}$ as before. I report the results, average R^2, and number of firms for each Fama-MacBeth regression in Table 1.20. The controls do not decrease the significance of the result that I have shown in previous sections: stocks underreact to commodity news.

1.4.2 Stale Pricing and Lookahead Bias

Another potential problem to guard against is that smaller stocks might not be trading at the end of the day and since this effect is stronger in smaller securities, there may be lookahead bias in commodity information. Imagine a situation where a stock does not trade in the last hour of the day and the return computed in CRSP is actually based on the midpoint of a very wide bid-ask spread. In that case, I would be using end of day information in commodities but my equity returns would be based on prices that were only true an hour before close

Table 1.21: Commodity Exchange Closing Times

A list of all the commodities used in this study, the exchange they are traded on and its current closing time. Commodities end their trading day prior to equities.

Commodity	Settlement Time (EST)	Exchange
Cocoa	11:50 AM	ICE
Coffee	1:25 PM	ICE
Corn	2:15 PM	CME
Cotton	2:15 PM	ICE
Kansas Wheat	2:15 PM	CME
Soybeans	2:15 PM	CME
Sugar	1:00 PM	ICE
Wheat	2:15 PM	CME
Brent Crude	2:30 PM	ICE
Gasoil	11:30 AM	ICE
Heating Oil	2:30 PM	CME
Natural Gas	2:30 PM	ICE
RBOB HU	2:30 PM	CME
WTI	2:30 PM	CME
Aluminum	8:15 AM	LME
Copper	1:00 PM	CME
Gold	1:30 PM	CME
Lead	8:15 AM	LME
Nickel	8:15 AM	LME
Silver	1:25 PM	CME
Zinc	8:15 AM	LME

leading to lookahead bias. In this section I show that this problem is not corrupting my analysis in two ways: first commodities actually stop trading earlier than equities and some actually stop trading many hours earlier. Table 1.21 lists the exchange closing times for the commodities used in this study. The exchanges that trade energy commodities are the latest to close among the commodities used and finish trading by 2:30 PM; agricultural commodities and metals finish trading even earlier. Given these circumstances, it is unlikely that there is any lookahead bias in my analysis.

However, to be sure, there is a more conservative way to conduct the analysis. At the end of the month, I simply sort stocks based on commodity information that was known on the second to last day of the month. This skipping of one day guarantees a conservative result that is immune to lookahead bias. Specifically, at the end of each month I sort stocks into small and large categories based on the NYSE median as before. Within each size category I sort stocks into quintiles using commodity news that only uses information on the second to last day of the month (instead of the last day) and form value weighted portfolios. As before I go long stocks that have positive commodity news and short stocks that have had negative commodity news. Table 1.22 presents the results of this robustness experiment. This inattention strategy using the second to last day of the month commodity information yields 1.7% per month four factor alpha in small stocks and an insignificant .5% in large stocks (difference of 1.2%, Newey-West t-statistic of 3.68). Clearly the results presented earlier are not driven by any sort of lookahead bias in commodity information.

1.4.3 Individual Commodities and Alternative Methodology

While the stock universe selection mechanism described in Section 1.2.2 is straightforward and transparent, it is nonetheless important to make sure that a different methodology does not produce contradicting or vastly different results. There were three choices that I made in the classification scheme: 1) to classify commodities into sectors, 2) to use a cutoff to retain only the industries that have a statistically significant association with a commodity sector, and 3) retain securities that behave like the rest of their industry with respect to the commodity sector. In this section I will relax these assumptions by using individual commodities (rather than commodity sectors) and I will use the elastic net framework of Zou & Hastie (2005) to fit a sparse commodity model to each equity industry using leave-one-out cross validation rather than using a t-statistic cutoff as I had done in Section 1.2.2. I will refrain from removing securities that behave differently than the rest of their industry with respect to individual commodities, thus relaxing assumption 3. In other words, I will attempt to use a different statistical mechanism that relaxes my earlier assumptions to determine if the results still indicate that equity investors are not fully attentive to the commodity market.

The elastic net framework is a combination of ridge regression of Hoerl & Kennard (1970)and LASSO of Tibshirani (1996). OLS suffers from the problem that highly correlated independent variables produce a poorly conditioned covariance matrix that leads to extreme coefficients on those variables (often of different signs). Ridge regression attempts to overcome this problem by penalizing the square of the coefficients to shrink the coefficient

Table 1.22: Skip A Day Size Double Sort

To verify that there is no lookahead bias driving previous results, this table skips a day between commodity information and equity portfolio formation. Specifically, at the end of each month securities are sorted into small and large stocks using the NYSE median, then into quintile value weight portfolios based on commodity news that was known on the second to last day of the month.

Size	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Small	1	-0.143	-0.767	**-1.054**	**-0.961**	-0.060	-0.317	2.454	1.168	0.565	0.727	-0.118
			[-1.87]	**[-2.87]**	**[-2.84]**							
	2	0.319	-0.257	-0.434	-0.355	0.163	-0.443	2.232	0.999	0.735	0.440	-0.101
			[-0.81]	[-1.60]	[-1.37]							
	3	0.480	-0.122	-0.277	-0.266	0.243	-0.539	2.296	1.028	0.915	0.409	-0.014
			[-0.44]	[-1.23]	[-1.18]							
	4	0.904	0.312	0.126	0.205	0.459	-0.337	2.142	0.998	0.947	0.462	-0.100
			[1.12]	[0.53]	[0.93]							
	5	1.520	**0.893**	**0.639**	**0.764**	0.658	-0.185	2.537	1.093	0.873	0.625	-0.159
			[2.50]	**[2.10]**	**[2.39]**							
	5-1	1.663	**1.661**	**1.694**	**1.726**	0.731	0.044	1.732	-0.074	0.308	-0.102	-0.040
			[3.43]	**[3.46]**	**[3.40]**							
Big	1	0.413	-0.109	-0.170	-0.150	0.212	-0.270	2.609	0.998	-0.102	0.156	-0.026
			[-0.35]	[-0.61]	[-0.54]							
	2	0.109	-0.310	-0.349	-0.370	0.066	-0.726	3.708	0.808	-0.079	0.114	0.027
			[-1.25]	[-1.29]	[-1.30]							
	3	0.673	0.252	0.172	0.106	0.445	0.273	3.507	0.840	-0.039	0.241	0.083
			[1.21]	[0.85]	[0.52]							
	4	0.737	0.319	0.182	0.163	0.479	-0.729	4.600	0.840	0.017	0.374	0.024
			[1.40]	[0.86]	[0.72]							
	5	1.017	**0.532**	0.415	0.375	0.550	-0.388	3.136	0.947	0.089	0.332	0.050
			[1.98]	[1.63]	[1.52]							
	5-1	0.604	0.641	0.585	0.525	0.296	-0.196	1.088	-0.051	0.190	0.176	0.076
			[1.60]	[1.47]	[1.34]							
Small - Big	5-1	1.059	**1.020**	**1.109**	**1.200**	0.722	0.032	1.751	-0.023	0.118	-0.277	-0.116
			[3.32]	**[3.77]**	**[3.68]**							

magnitude on correlated variables toward zero. This has the effect of "averaging" several correlated variables and using a modest coefficient on that average. However, ridge regression retains all the variables in the model even if some of them are irrelevant. LASSO, on the other hand, explicitly sets some coefficients to zero. This has the effect of selecting a parsimonious model to represent the dependent variable. Tibshirani (1996) provides some intuition on how the LASSO objective function differs from ridge regression: in a simplified problem with orthonormal independent variables the LASSO objective function shifts the magnitude (absolute value) of the OLS coefficient by an amount related to λ (the magnitude of penalization). Ridge regression, on the other hand, scales the OLS coefficients by an amount related to λ. In this sense, LASSO is similar to subset selection but operates in a continuous fashion if the resulting coefficient is non-zero. Figures 1 and 2 in Tibshirani (1996) illustrate this difference. Elastic net, therefore, selects a parsimonious model but "averages" correlated variables by shrinking their coefficients toward zero rather than simply selecting a particular variable of the correlated subset (as LASSO might).

Formally elastic net solves the following optimization problem:

$$\min_{a,\beta} \frac{1}{T} \sum_{t=1}^{T} \left(IR_{j,t} - a_{j,t} - \boldsymbol{\beta'_{j,t}} \boldsymbol{R_{d,t}} \right)^2 + \lambda \left[(1-\alpha) \|\boldsymbol{\beta_{j,t}}\|_2^2 / 2 + \alpha \|\boldsymbol{\beta_{j,t}}\| \right] \tag{1.30}$$

where $\boldsymbol{R_{d,t}}$ is a vector of individual commodity returns and $IR_{j,t}$ is an industry return. There are two choices that need to be made: how much penalization is done overall (λ) and how to combine the features of LASSO and ridge regression (α). λ is selected using leave-one-out cross-validation while α is usually selected apriori (in my analysis I set $\alpha = .5$ as described in Zou & Hastie (2005)). To implement cross-validation I use a backward looking rolling 5 year window of non-overlapping weekly returns. For a particular value of λ I leave one week out (validation sample), fit the model over the remaining weeks (test sample), and compute the RMSE in the validation sample. I repeat this process in the same 5 year window leaving out a different week and again calculating the model error. I then average the RMSE over all the validation samples and pick the λ that yields the lowest average error. Note that $\boldsymbol{R_{d,t}}$ is standardized prior to the fit so that commodities are treated equally in the selection procedure (if returns were not standardized then commodities that are unusually volatile would be penalized less since their β would naturally be smaller).

Using this fitted model I compute $R^{alt}_{i,c,t} \equiv \boldsymbol{\beta'_{j,t} R_{d,t}}$ at the end of each month for all stocks in CRSP and sort securities into small and large categories using the NYSE median. For each size category I then form value weighted quintile portfolios based on $R^{alt}_{i,c,t}$ going long (short) stocks with positive (negative) commodity news. Many of the stocks in CRSP simply do not have strong commodity associations so that $\boldsymbol{\beta_{j,t} = 0}$ and, therefore, $R^{alt}_{i,c,t} = 0$. I first report results that include these stocks with 0 commodity news to simply show that my conclusions hold when using the entire CRSP universe. This sample will have significantly smaller alpha simply because a lot of the securities have extremely small relationships to commodities and thus the spread in commodity news of the long short portfolio will be smaller. I also report results after eliminating securities that have "low" commodity news: to do this, each month I compute a commodity news Z-Score using the entire CRSP cross-section.

$$RZ^{alt}_{i,c,t} = \frac{R^{alt}_{i,c,t} - \overline{R^{alt}_{c,t}}}{\sigma_t(R^{alt}_{i,c,t})} \tag{1.31}$$

Each month I only form portfolios using stocks that have "high" commodity news by keeping securities where $|RZ^{alt}_{i,c,t}| > 1$: that is, this procedure keeps securities that have commodity news larger than 1 standard deviation relative to the rest of the CRSP universe. Table 1.23 and 1.24 report the results of these two experiments.

Using the entire CRSP universe, the inattention trading strategy using this methodology yields a four factor alpha of .55% per month in small securities and an insignificant $-.02\%$ per month in large securities (difference .572%, Newey-West t-statistic 2.476). Using a sample of securities that have high commodity news, the inattention strategy yields a four factor alpha of 1.5% per month in small securities and an insignificant .33% per month in large securities (difference 1.187%, Newey-West t-statistic 3.11). There are several inferences that are worth noting from these results: first, the inattention phenomenon is not dependent on a specific type of methodology selected in this study. It is robust to using an alternative model selection procedure and using all commodities individually instead of classifying them into

Table 1.23: Individual Commodity Size Double Sort

An alternate model selection methodology is used to verify that all the results are robust and not affected by particular assumptions made in Section 1.2.2. This table uses all stocks in CRSP (even those that have negligible commodity news).

Size	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Small	1	0.409	-0.289	**-0.388**	-0.277	0.222	-0.625	1.156	1.029	0.911	0.173	-0.138
			[-1.39]	**[-2.90]**	[-1.77]							
	2	0.495	-0.203	**-0.306**	**-0.229**	0.277	-0.593	-0.272	1.048	0.869	0.196	-0.096
			[-1.11]	**[-2.49]**	**[-1.96]**							
	3	0.813	0.114	0.019	0.154	0.460	-0.563	-0.386	1.035	0.821	0.156	-0.168
			[0.68]	[0.17]	[1.57]							
	4	0.685	-0.044	-0.096	0.018	0.373	-0.652	-0.361	1.068	0.852	0.069	-0.142
			[-0.27]	[-0.88]	[0.15]							
	5	0.927	0.233	0.146	0.274	0.513	-0.325	-0.795	1.024	0.841	0.141	-0.160
			[1.32]	[1.03]	[1.92]							
	5-1	0.519	**0.522**	**0.533**	**0.551**	0.464	-0.261	4.016	-0.005	-0.070	-0.032	-0.022
			[2.25]	**[2.19]**	**[2.15]**							
Big	1	0.447	-0.138	-0.160	-0.122	0.311	-0.607	-0.514	0.971	0.028	0.034	-0.047
			[-1.17]	[-1.36]	[-1.02]							
	2	0.618	0.039	0.017	0.019	0.456	-0.664	-1.142	1.009	-0.200	0.048	-0.002
			[0.40]	[0.20]	[0.21]							
	3	0.744	0.172	**0.176**	0.155	0.554	-0.686	-0.819	0.982	-0.130	0.000	0.026
			[1.95]	**[2.01]**	[1.64]							
	4	0.785	0.190	0.233	0.217	0.557	-0.742	-0.223	0.979	0.000	-0.090	0.020
			[1.44]	[1.87]	[1.81]							
	5	0.498	-0.101	-0.096	-0.143	0.336	-0.643	-0.848	1.019	-0.033	0.007	0.059
			[-0.87]	[-0.76]	[-1.13]							
	5-1	0.051	0.036	0.064	-0.022	0.041	-0.485	2.300	0.048	-0.061	-0.028	0.106
			[0.17]	[0.28]	[-0.10]							
Small - Big	5-1	0.468	**0.486**	**0.469**	**0.572**	0.473	1.293	2.362	-0.053	-0.009	-0.004	-0.128
			[2.34]	**[2.17]**	**[2.48]**							

Table 1.24: Individual Commodity Size Double Sort - Newsworthy Stocks

An alternate model selection methodology is used to verify that all the results are robust and not affected by particular assumptions made in Section 1.2.2. This table uses all stocks that are newsworthy.

Size	Commodity News	$\overline{r^e}$	CAPM α	Fama-French α	Four Factor α	Annualized Sharpe	Skewness	Excess Kurtosis	β_m	β_{smb}	β_{hml}	β_{umd}
Small	1	-0.062	**-0.725**	**-0.999**	**-0.897**	-0.030	-0.717	0.834	1.060	0.882	0.569	-0.127
			[-2.02]	**[-3.38]**	**[-2.94]**							
	2	0.443	-0.273	-0.340	-0.257	0.205	-0.003	1.247	1.074	0.779	0.114	-0.103
			[-0.95]	[-1.43]	[-1.06]							
	3	0.859	0.223	0.168	0.409	0.431	0.145	-0.966	0.888	0.766	0.025	-0.300
			[0.84]	[0.62]	[1.64]							
	4	0.837	0.189	0.101	0.208	0.427	-0.433	-1.126	0.951	0.861	0.150	-0.133
			[0.76]	[0.43]	[0.86]							
	5	1.343	**0.690**	**0.524**	**0.627**	0.643	-0.266	-0.804	1.005	0.813	0.328	-0.129
			[2.58]	**[2.16]**	**[2.82]**							
	5-1	1.406	**1.415**	**1.523**	**1.524**	0.705	-0.013	-0.271	-0.055	-0.069	-0.241	-0.002
			[3.32]	**[3.58]**	**[3.65]**							
Big	1	0.431	-0.147	-0.275	-0.130	0.236	0.035	-0.366	0.959	0.133	0.227	-0.179
			[-0.64]	[-1.25]	[-0.55]							
	2	0.562	-0.089	-0.110	-0.087	0.291	0.132	0.795	1.072	0.100	0.039	-0.028
			[-0.37]	[-0.41]	[-0.35]							
	3	0.466	-0.146	-0.173	-0.205	0.264	-0.567	0.653	1.054	-0.051	0.073	0.039
			[-0.69]	[-0.82]	[-0.86]							
	4	0.699	0.140	0.104	0.054	0.417	-0.339	-1.989	0.980	-0.086	0.099	0.062
			[0.70]	[0.48]	[0.29]							
	5	0.852	0.280	0.218	0.207	0.466	-0.685	-0.701	0.976	0.063	0.143	0.013
			[1.13]	[0.93]	[0.96]							
	5-1	0.421	0.428	0.493	0.338	0.211	-0.241	-0.388	0.017	-0.070	-0.083	0.193
			[1.14]	[1.32]	[1.00]							
Small - Big	5-1	0.985	**0.987**	**1.030**	**1.187**	0.655	0.767	-0.347	-0.072	0.001	-0.158	-0.194
			[3.27]	**[3.23]**	**[3.11]**							

a commodity sector. Second, these results confirm the size phenomenon identified earlier: small stocks underreact more to commodity news than large stocks.

In summary, equity investors do not fully appreciate the relevance of the commodity market for equity returns. The results are robust to other previously known phenomena, are not driven by any sort of stale pricing or lookahead bias in commodity data, and are robust to a completely different model selection methodology.

1.5 Conclusion

Understanding how information is incorporated into prices is an important research area that informs market design and asset allocation. Much of the rational asset pricing literature has argued that prices incorporate all publicly available information into prices instantaneously. More recently, empirical and theoretical work has started examining the psychological impediments that could prevent investors from acting exactly like infinite capacity computers. Investors may have capacity constraints on how much information they can process per unit of time.

This study has examined how equity investors incorporate information from the commodity market into stock prices. I have shown that investors underreact to commodity information leading to predictable stock returns. In particular, investors ignore information that is least important to their overall portfolio. Therefore stocks that are "unimportant" to their owners are slow to fully incorporate all available information. Similarly investors ignore commodity news if their overall portfolio is not significantly affected by this news; stocks affected by commodity news that are owned by investors who ignore commodity news underreact to commodity information. Investors also fail to appreciate a firm's connection to a commodity when that connection is nuanced and not frequently mentioned in the press. Finally, inattention to commodities is significantly stronger in periods when investors have many idiosyncratic information sources to process. Future areas of research should continue to explore the costs investors face in acquiring information and incorporating it into prices.

Chapter 2

Market Crash Risk and Slow Moving Capital[1]

2.1 Introduction

Professional investors commonly look to the options market to assess market conditions through the premium the market places on portfolio protection. Specifically, they look at the difference in implied volatility between the put and call options on the S&P 500 index. As investors become more cautious about future market outcomes, they purchase more puts - often financing the investment by selling calls - which increases the difference in the implied volatility between puts and calls; this trade is known as a risk reversal. Our work attempts to understand if the cross-section of securities efficiently incorporates information regarding changes in the price of risk reversals. We use the cross-section of equity securities as our primary test assets and verify our conclusions in the corporate bond market as well.

Our findings are two-fold: the value-minus-growth (HML)[2] trade performs poorly when the price of risk reversals increases. HML incorporates information into its price slowly and

[1]This chapter is co-authored jointly with Johnny Kang

[2]We refer to a trade that goes long value and short growth as well as the formal value factor constructed by Fama & French (1992) as HML.

thus continues to do poorly the following month as well. These conclusions also hold in the corporate bond market: the junk-minus-investment grade trade performs poorly when the price of risk reversals increases but also continues to perform poorly the following month. Using mutual fund flow data, we find that following increases in the price of risk reversals - which can happen because investors become more risk averse or because the distributional characteristics of the underlying market change - investors rotate out of value funds and into growth funds. However, this rotation happens slowly contributing to the predictability we find in HML.

Our work contributes to understanding how information diffuses from the options market to the equity market. Several papers have explored how individual option skew relates to future returns Conrad et al. (2013) find that risk neutral skewness of individual options forecasts higher future returns while Xing et al. (2010) find that a larger differential between out of the money put options and call options forecasts a lower future return attributing the result to slow information diffusion (and forces out risk neutral skewness in a multivariate setting). Ang et al. (2012) show that individual stock options contain information regarding future returns of individual stocks. Brunnermeier et al. (2008) examines the difference between put and call volatility in currency pairs showing that currencies involved in carry trades exhibit significant crash risk as measured by the respective options. Pan (2002) shows that the difference between out of the money put volatility and out of the money call volatility is largely driven by a downward jump risk premium. Bollen & Whaley (2004) finds that S&P skew is largely driven by buying pressure from investors for puts but remains agnostic to the reasons. Unlike the aforementioned papers, we examine how investor risk neutral estimates of market crash risk derived from S&P 500 options affects the cross-section of stock and credit returns.

Our findings that the value minus growth (HML) factor is highly correlated with innovations in skew changes ties with another literature on understanding the value premium. In a series of papers Fama & French (1992, 1993, 1996, 1998) show that high book-to-market securities earn high returns relative to low book-to-market stocks. The rational asset pricing literature has advocated for an ICAPM style model where growth and value stocks have covariance with state variables. Campbell & Vuolteenaho (2004) examine the covariance of securities with discount rates and cash flows. Using an ICAPM they show that the price of risk associated with discount rate covariance is lower; value stock betas are mostly composed of cash flow betas while growth stock betas are discount rate betas. Therefore value stocks should command a higher return per unit of market beta. Petkova & Zhang (2005) show that market β of value stocks is higher in bad states of the world. Our work poses a challenge for a strictly rational interpretation of HML since the factor that is supposed to explain returns is itself predictable. This is also important for empirical papers that document anomalies in the stock market: if certain factors are predictable then examining exposure to static factors is no longer a high enough threshold. We need to think about how dynamic portfolios of factors affect the strength of other anomalies.

2.2 Risk Reversals

Intuitively, risk reversals correspond to a self financing position that an individual uses to express concern about negative market outcomes. It has also been used extensively, for example Brunnermeier et al. (2008), as a measure of crash risk. Specifically, if the implied volatility of puts is higher than calls, this implies that traders believe the risk neutral distribution of returns is negatively skewed. The theoretical motivation for using the option skew to forecast returns comes from Pan (2002) who constructs a model to attempt to reconcile the dynamics of spot prices of the S&P 500 with option prices on the same. Pan finds that a model that only includes a volatility risk premium is inadequate to fit the option series and a model that includes jumps with state dependent intensity is necessary. Specifically, she uses a model that incorporates a time varying jump intensity that rises with volatility: when market volatility increases, so does the probability of a large downward jump. Unlike a model that only includes a risk premium for volatility, this model isn't rejected by the data and is able to price the full term structure of options well. The model estimates a risk premium of 5.5% for diffusive risk and 3.5% for jump risk on the S&P 500. Jump risk is specifically reflected in out of the money puts: for at the money options, approximately 55% of the overall risk premia is due to jump risk. For a 5% out of the money put, 80% of the risk premia is due to jumps while for a 5% out of the money call, only 30% is due to jumps. This differential ability by out of the money puts and calls to capture the magnitude of the jump risk premium is reflected in our construction of $SKEW$. Specifically define

$$SKEW_t \equiv \frac{1}{N}\sum_{j=1}^{N} IV(P)_{j,t} - \frac{1}{N}\sum_{j=1}^{N} IV(C)_{j,t} \tag{2.1}$$

where $IV(P)$ ($IV(C)$) is the volatility of 1 year S&P 500 out of the money (OTM) put (call) options at equally spaced delta grid points. Thus this is simply the average volatility on out of the money (OTM) S&P 500 put options minus the average volatility on OTM S&P 500 call options. We use the interpolated volatility surface from OptionMetrics to compute $SKEW$ monthly from 1996-2012 based on data availability. The surface provides a set of volatilities for put and call options in increments of five delta units and thus is symmetric around the at the money (ATM) point. Since our work will focus on monthly data and option markets close 15 minutes after equity markets, we use the second to last day of the month to compute our implied volatility related metrics to prevent any look ahead bias.

To get a better sense of how $SKEW$ relates, empirically, to the moments of the risk neutral distribution we follow the methodology of Bakshi et al. (2003) and extract the model free moments from the cross section of S&P 500 option prices. We proxy for market volatility using the VIX index and extract the risk neutral skewness, $RNSKEW$. As noted earlier, a high positive $SKEW$ represents risky states of the world from the perspective of the investor. Analogously a highly positive VIX represents high (risk neutral) forecasted levels of volatility; a highly *negative* $RNSKEW$ represents a *negatively* skewed distribution. The point to keep in mind is that a *negative* innovation in $RNSKEW$ is an *increase* in risk from the perspective of an investor. The opposite is true for VIX and $SKEW$: a *positive* innovation to these variables corresponds to a riskier distribution. We extract innovations at the monthly frequency using univariate ARMA models chosen by BIC for each quantity ($SKEW$, VIX and $RNSKEW$). All of the quantities are can be described by a low order

ARMA model.

Our goal is to understand how $SKEW$ innovations, ε_{skew}, relate to innovations in risk neutral moments. We do this parametrically using linear regression and non-parametrically using local polynomial regressions. Namely, we run two forms of regressions:

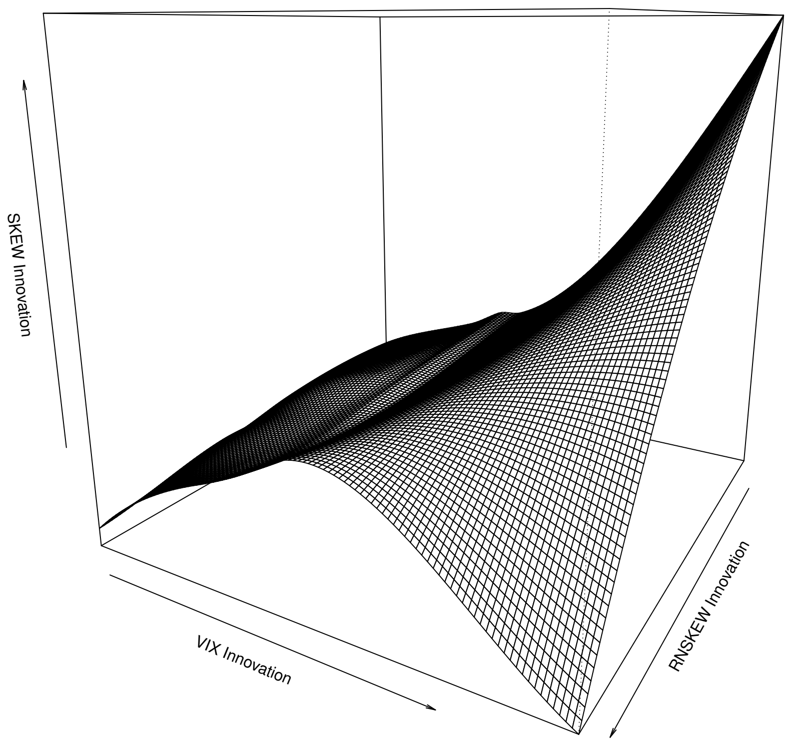
$$\varepsilon_{skew,t} = f(\varepsilon_{rnskew,t}, \varepsilon_{vix,t}) \tag{2.2}$$

$$\varepsilon_{skew,t} = a + \beta_v \varepsilon_{vix,t} + \beta_{rns} \varepsilon_{rnskew,t} + \beta_i (\varepsilon_{vix,t} \varepsilon_{rnskew,t}) + \eta_t \tag{2.3}$$

where $f(\cdot)$ is a second order local polynomial. Figure 2.1 presents a surface representing the results of fitting (2.2). Variables are standardized prior to running these regressions so that magnitudes can be more easily interpreted. The arrows point in the positive direction for each variable.

The figure plots the fitted values of $\varepsilon_{skew,t}$ from regression (2.2) against $\varepsilon_{vix,t}$ and $\varepsilon_{rnskew,t}$. It shows ε_{skew} is high when there is an increase in volatility *and* the distribution of returns becomes more negatively skewed. Thus $SKEW$ is a measure of the joint behavior of volatility and skewness. This empirical result is consistent with the theoretical model of Pan (2002) who specifies risk neutral jump intensity as a function of volatility. Specifically, in her model the jump intensity increases with market volatility and this is our finding as well. This can be examined in a linear regression context also: the table below the figure presents results of regression (2.3). The results are the same as those explained by the plot: $SKEW$ increases when the distribution becomes more negatively skewed ($\varepsilon_{rnskew} < 0$) and volatility increases ($\varepsilon_{vix} > 0$). Additionally, there is a significant interaction effect that was highlighted by the plots: $SKEW$ increases particularly strongly when there is an increase in volatility and the distribution of returns becomes more negatively skewed.

It is helpful to understand how $SKEW$ varies through time in relation to the VIX and the business cycle since investors are generally familiar with the time-series pattern of these quantities. Figure 2.2 presents a plot of $SKEW$ for our sample along with the VIX and $SKEW$ orthogonalized to the VIX (using linear regression) labeled $OrthSKEW$; recessions are highlighted using gray bars. The variables have been standardized. One obvious pattern



Dependent Variable	$\varepsilon_{rnskew,t}$	$\varepsilon_{rnskew,t} * \varepsilon_{vix,t}$	$\varepsilon_{vix,t}$	(Intercept)	R^2	N
$\varepsilon_{skew,t}$	**-0.167**	**-0.266**	**0.516**	0.062	41.02%	201
	[-3.421]	**[-2.074]**	**[9.369]**	[1.013]		

Figure 2.1: Nonparametric Relationship Between $SKEW$, VIX and $RNSKEW$ Innovations

We show a nonparametric surface that relates innovations in $SKEW$ to innovations in the VIX and $RNSKEW$. The surface is the result of a local polynomial regression using second order polynomials and 75% span. Below the plot we present the parametric version of the relationship from equation (2.3).

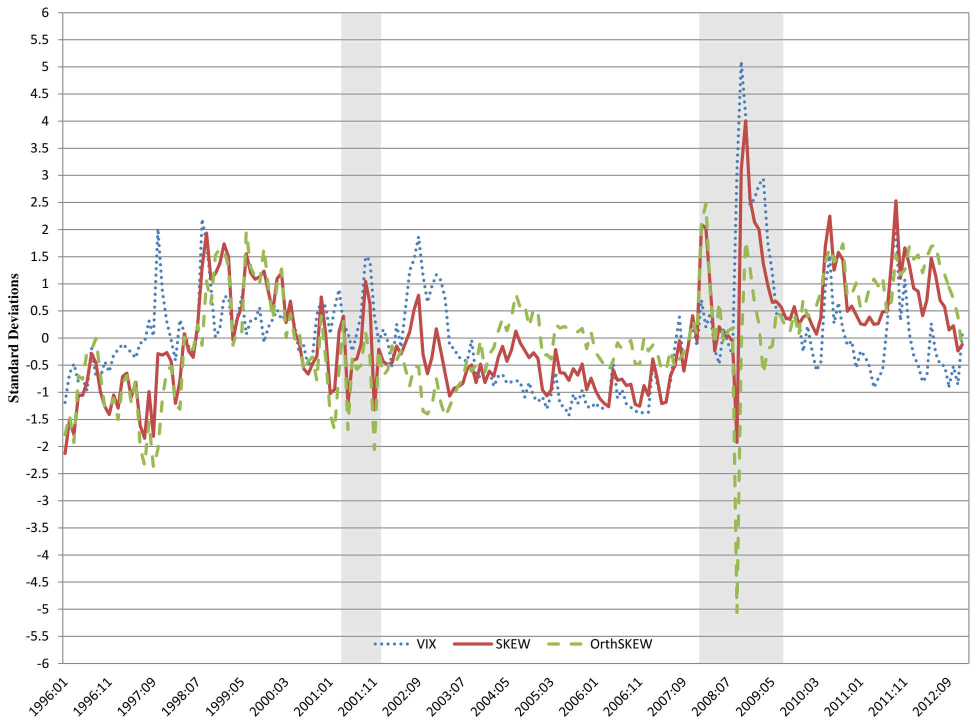


Figure 2.2: Time Series Plot of $SKEW$

Plot of $SKEW$, VIX and $OrthSKEW$ ($SKEW$ orthogonalized to the VIX through linear regression) from 1996-2012.

is the correlation that $SKEW$ exhibits with the VIX: roughly 60%. However, there are subtle differences in the pattern: $SKEW$ was higher relative to its normal levels than the VIX (relative to its normal levels) prior to the dot-com bubble bursting. It also spiked up prior to the 2008 financial crisis and has remained elevated after the crisis. We will see further in the article that these features are important in predicting HML returns.

2.3 Security Underreaction

2.3.1 Equities

Pan (2002) provides the motivation for examining $SKEW$ as a measure of aversion by investors to jumps in the S&P 500: she finds that the overwhelming majority of the difference in OTM put and call volatilities corresponds to the jump premium (as opposed to actual jump realizations). We document that $SKEW$ does, in fact, have significant forecasting power for the overall market return consistent with it being a measure of risk premium. If investors who trade S&P 500 options (and the underlying index since options investors often delta hedge their positions) are more sophisticated than the average investor in the cross-section of equities, then information from $SKEW$ will diffuse through the cross-section slowly.

The finance literature has already identified several variables that forecast future market returns. We are careful to control for these other variables to understand the multivariate implications of $SKEW_t$ on market returns. The variables that we consider are the dividend yield on the S&P 500, the smoothed earnings yield defined as the 10 year trailing moving average of aggregate earnings on the S&P 500 divided by the index price level, the term premium defined as the difference between the 10 year and 3 month treasury bond yield and the default premium defined as the difference in yield on Moody's AAA index and Moody's BAA index. To understand the relationship between $SKEW_t$ and expected market returns, we run the following regression:

$$R^e_{m,t} = a + \rho_m R^e_{m,t-1} + \beta_s SKEW_{t-1} + \beta_{dp} dp_{t-1} + \beta_{dy} dy_{t-1} + \beta_{sey} sey_{t-1} + \beta_{tp} tp_{t-1} + \varepsilon_{m,t} \quad (2.4)$$

using monthly data from 1996 - 2012. Our sample is constrained by the availability of options data. Additionally, much of the expected return literature - see Cochrane (2011) for a recent summary - has reported stronger effects at longer horizons. Therefore, we also run a regression of 6 month market excess returns on predictor variables:

$$R^e_{m,t\to t+5} = a + \rho_m R^e_{m,t-1} + \beta_s SKEW_{t-1} + \beta_{dp} dp_{t-1} + \beta_{dy} dy_{t-1} + \beta_{sey} sey_{t-1} + \beta_{tp} tp_{t-1} + \varepsilon_{m,t} \quad (2.5)$$

Table 2.1 reports the results. As noted earlier, since options markets close later than equity markets, we skip an extra day between information on $SKEW_t$ and any equity return to prevent look-ahead bias. Thus in monthly data, $SKEW_t$ represents observations on the

Table 2.1: $SKEW$ Forecasts of Future Market Excess Returns

We report the forecasting power of $SKEW$ for future market returns across the 1 month and 6 month horizon. We are careful to control for other variables that have been found to predict market returns in the literature.

Dependent Variable	$SKEW_{t-1}$	$R^e_{m,t-1}$	dp_{t-1}	dy_{t-1}	sey_{t-1}	tp_{t-1}	(Intercept)	R^2	N
$R^e_{m,t}$	0.260						-0.734	0.35%	201
	[1.221]						[-0.724]		
$R^e_{m,t}$	**0.377**	**0.162**					-1.357	2.31%	201
	[2.221]	**[1.978]**					[-1.429]		
$R^e_{m,t}$	**0.591**	0.171	**-3.659**	1.741	0.741	-0.315	**-4.078**	5.45%	201
	[2.176]	[1.934]	**[-2.051]**	[0.986]	[1.178]	[-1.031]	**[-2.291]**		
$R^e_{m,t\to t+5}$	**1.674**						-4.803	4.32%	196
	[2.025]						[-0.907]		
$R^e_{m,t\to t+5}$	**1.873**	0.272					-5.847	4.78%	196
	[2.346]	[1.233]					[-1.140]		
$R^e_{m,t\to t+5}$	2.142	0.300	-12.345	10.952	2.754	-1.208	-22.535	16.75%	196
	[1.516]	[1.305]	[-1.089]	[0.672]	[0.469]	[-0.466]	[-1.795]		

second to last day of the month. $SKEW$ has significant forecasting power for market returns even in the presence of other state variables. A one volatility point higher $SKEW$ corresponds to roughly 30 - 60 basis points of expected market returns the following month. Similarly the second half of the table shows that a one volatility point increase in $SKEW$ is related to 1.8% higher expected return over the following 6 months.

We next turn to the cross-section of equities: to determine if innovations in $SKEW$ are differentially important in the cross-section (this is not a pre-determined conclusion) we use the Fama-French 25 portfolios as our basis assets (since individual equity security returns are noisy) and regress excess returns of each portfolio, $R^e_{i,t}$, on ε_{skew} controlling for market returns and volatility innovations proxied by innovations in the VIX index:

$$R^e_{i,t} = a + \beta_m R^e_{m,t} + \beta_v \varepsilon_{vix,t} + \beta_{\varepsilon_{skew}} \varepsilon_{skew,t} + \nu_t \tag{2.6}$$

We are careful to control for volatility innovations because Pan's model (and our earlier empirical results) specifies jump intensity to be a function of volatility so we want to be sure we aren't picking up changes in the volatility risk premium. Table 2.2 presents the multiple β of each portfolio with respect to ε_{skew}. The results here are clear: growth stocks have a

Table 2.2: Cross Sectional Sort by β to ε_{skew}

Multiple β of Fama-French 25 portfolio returns to innovations in $SKEW$; equation (2.6). Newey-West t-statistics are in brackets.

	Book-to-Market					
Size	Growth	2	3	4	Value	Value - Growth
Small	0.59	0.22	0.05	-0.13	-0.41	**-1.00**
	[1.07]	[0.47]	[0.15]	[-0.37]	[-0.99]	**[-2.89]**
2	0.67	0.17	-0.24	-0.26	-0.23	**-0.90**
	[1.41]	[0.51]	[-0.94]	[-0.84]	[-0.74]	**[-2.18]**
3	0.53	0.05	**-0.33**	-0.29	**-0.56**	**-1.09**
	[1.66]	[0.30]	**[-2.03]**	[-1.59]	**[-2.45]**	**[-2.67]**
4	0.47	-0.20	-0.32	-0.21	-0.24	-0.71
	[1.85]	[-1.10]	[-1.66]	[-1.03]	[-0.98]	[-1.87]
Large	0.02	-0.17	-0.34	**-0.52**	-0.27	-0.29
	[0.15]	[-1.11]	[-1.84]	**[-2.41]**	[-1.20]	[-1.14]
Large - Small	-0.57	-0.39	-0.39	-0.39	0.14	
	[-0.91]	[-0.70]	[-0.90]	[-0.92]	[0.27]	

positive exposure to increases in the jump risk premium while value stocks have a negative exposure. That is, when the risk neutral distribution becomes riskier, value stocks do poorly while growth stocks do well. The results are especially strong in small and medium stocks.

Since HML is the factor that has exposure to ε_{skew} (while the small-minus-big factor does not), we use the HML portfolio directly rather than each individual Fama-French portfolio in subsequent results. As mentioned in the introduction, we are interested in understanding how efficient securities in the cross-section are at incorporating information regarding changes in jump risk premium into their prices; we find a significant lag in the price adjustment process. In addition to doing poorly at time t when there is a positive innovation to $SKEW_t$, $R_{hml,t+1}$ is also highly negative; value stocks continue to underperform the following month. Figure

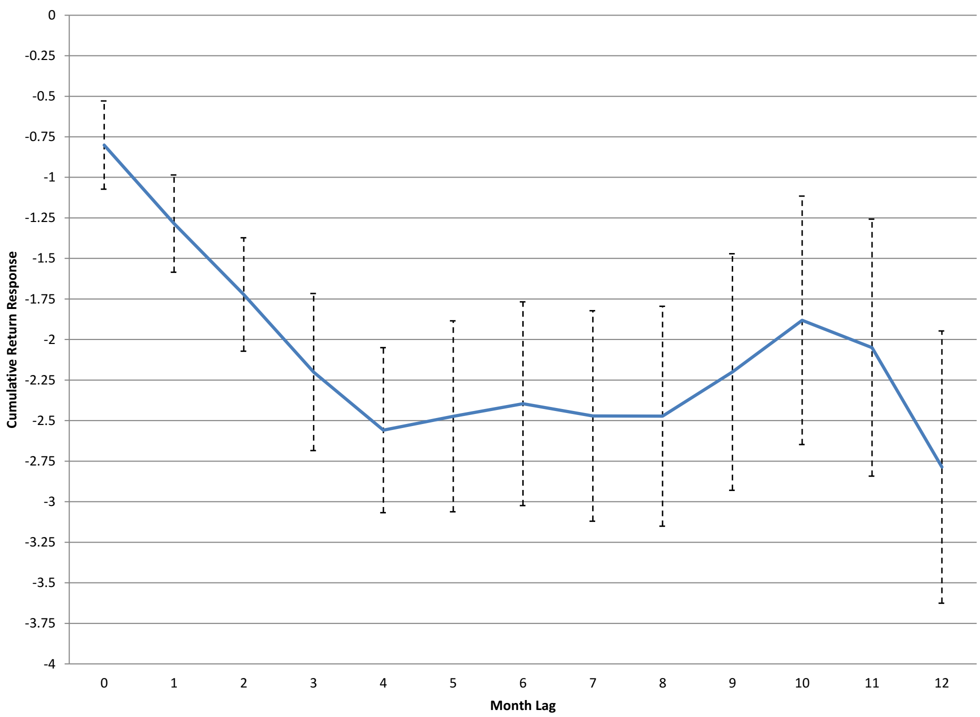


Figure 2.3: HML Underreaction to $SKEW$ Innovations

This plot demonstrates the underreaction of $R_{hml,t}$ to innovations in $SKEW_t$. We run regressions of $R_{hml,t} = a + \beta_m R_{m,t} + \beta_{\varepsilon,j}\varepsilon_{skew,t-j} + \epsilon_t$ for $j = 0...12$ and plot $\sum_{k=0}^{j} \beta_{\varepsilon,k}$. The error bars are plotted in dashes and assume that estimates of $\beta_{\varepsilon,j}$ are independent. As is clear from the plot, there is a significant predictability to $R_{hml,t}$ based on lagged innovations in $SKEW_t$.

2.3 plots the cumulative response to $\varepsilon_{skew,t}$ from the regression of

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_{\varepsilon,j}\varepsilon_{skew,t-j} + \epsilon_t \tag{2.7}$$

for $j = 0...12$. The Newey-West error bounds are presented in the plot in dashes; j is measured on the x-axis. For each j, the figure plots $\sum_{k=0}^{j} \beta_{\varepsilon,k}$ and assumes that estimates of $\beta_{\varepsilon,j}$ are independent. This regression can be interpreted as asking: if there is a one volatility point increase in $SKEW$ at time t and a one volatility point increase in $SKEW$ at $t-1$, ..., $t-j$ then what is the cumulative effect on HML at time t?

We see a striking pattern: while $R_{hml,t}$ is indeed highly negatively correlated with $\varepsilon_{skew,t}$, there is a significant delay in the price adjustment process. The plot shows that the $j = 0$ and $j = 1$ coefficients for $R_{hml,t}$ are highly negative and significant (with associated t-statistics of -2.94 and -3.84, respectively). Following an increase in $SKEW$, HML continues to underperform the following month. In fact, we see that the price adjustment process actually takes up to 4 months to fully realize. This is a startling finding: to confirm these results we extend the sample internationally to Europe and Japan[3]. Note that our analysis is done from the perspective of a US investor (thus, for example, market returns refer to the CRSP value weighted market return). We also attempt to control for the effect of time-varying HML market β. As noted earlier, Petkova & Zhang (2005) found that the market β of HML varies through time; it is possible that a high level of $SKEW_{t-1}$ forecasts a higher market β at time t. To be precise, imagine that

$$R_{hml,t} = \beta_{m,t} R^e_{m,t} + \eta_t \tag{2.8}$$

$$\beta_{m,t} = \beta_0 + \beta_1 SKEW_{t-1} \tag{2.9}$$

then the appropriate attribution regression to run for HML is

$$R_{hml,t} = \beta_0 R^e_{m,t} + \beta_1 (SKEW_{t-1} \cdot R^e_{m,t}) + \eta_t \tag{2.10}$$

We would like to eliminate the possibility that a time varying market β is driving our results so we include the interaction of $SKEW_{t-1}$ and $R^e_{m,t}$ into the regression. Therefore, for each region we run a forecasting regression of the form:

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s \left(\sum_{j=1}^{4} \varepsilon_{skew,t-j} \right) + \beta_i (R^e_{m,t} \cdot SKEW_{t-1}) + \nu_t \tag{2.11}$$

where we use the sum of the last four innovations in $SKEW$ based on the results of Figure 2.3. Table 2.3 presents the summary statistics of HML returns in each region (Panel A) and reports the results from this regression (Panel B).

Panel A shows that HML in all three regions has a high return and a high CAPM alpha

[3]While this gives us comfort against data snooping concerns, we note that there is a 40% correlation between US HML returns and JPY HML returns; a 60% correlation between US and Europe HML returns.

during our sample period. The units are left in their natural monthly frequency (the Sharpe ratio is annualized). Thus HML, in the US, has a CAPM alpha of .347% per month. The sample statistics in all three regions are similar with Europe having the highest Sharpe ratio. Panel B reports that a one volatility point unexpected increase in $SKEW$ over the last four months leads to a .6% lower HML return the following month in the US, .25% lower HML return in Japan and .37% lower HML return in Europe. The adjusted R^2 statistics are also quite large: a simple univariate forecasting regression in the US can explain 10% of the variation in HML returns in monthly data. The t-statistics are also very large: in the US the t-statistic associated with $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ is -6; this is directly linked to the Sharpe ratio of a strategy that one can construct (ie Sharpe (1994)).

To construct a realistic trading strategy from these regressions, we would like to avoid estimating an ARMA model for innovation extraction and also avoid estimating the distributed lag model for forecasting R_{hml}. While this is certainly the optimal method, our sample is relatively small and we want to avoid all possibility of look-ahead bias in parameters. To get around this constraint, we will simply use the level of $SKEW$: since several lags of innovations seem important (as noted earlier) and $SKEW$ is not terribly persistent, we hope that older innovations that are irrelevant will have decayed sufficiently and thus not erode our forecasting performance. To do this we simply run regression (2.11) but replace the term $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ with $SKEW_{t-1}$:

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s SKEW_{t-1} + \beta_i (R^e_{m,t} \cdot SKEW_{t-1}) + \nu_t \tag{2.12}$$

Table 2.4 reports these results. Comparing the two tables we can immediately see that our forecasting performance is worse using the level of $SKEW$ as to be expected: the t-statistic is cut in half in the US and so is the R^2. Thus, backtest results that we report using $SKEW$ as a forecasting variable is a lower bound on the true performance that an investor can achieve.

These results imply a strategy, which we refer to as active HML, that would selectively rotate into and out of HML being long value (growth) and short growth (value) stocks at different points in time. We are careful to prevent any look-ahead bias in the parameters and

Table 2.3: Forecast of $R_{hml,t}$ Using $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ 1996-2012

Panel A shows summary statistics for HML in US, Japan (JPY) and Europe (EUR): all of these regions have a significant value effect. The results are from the perspective of a US investor (CAPM α is with respect to the US market). Panel B shows that lagged innovations in $SKEW$ have significant forecasting power for HML in all of these regions.

(a) HML Summary Statistics By Region

Region	$\overline{R_t}$	$\min R_t$	$\max R_t$	$\sigma(R_t)$	Sharpe	CAPM α	N
US	0.267	-12.600	13.840	3.495	0.265	0.347	201
JPY	0.457	-13.820	10.080	3.092	0.512	0.545	201
EUR	0.487	-9.570	10.960	2.653	0.636	0.481	201

(b) R_{hml} Forecasts

Region	$\sum_{j=1}^{j=4} \varepsilon_{skew,t-j}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.600**			**0.403**	10.87%	197
	[-6.010]			**[1.689]**		
US	**-0.576**	-0.154		**0.468**	15.07%	197
	[-5.159]	[-1.406]		**[1.773]**		
US	**-0.564**	-0.419	0.053	**0.428**	17.00%	197
	[-6.190]	[-1.692]	[1.403]	**[1.785]**		
JPY	**-0.255**			**0.505**	2.09%	197
	[-2.312]			**[1.939]**		
JPY	**-0.227**	**-0.183**		**0.582**	9.89%	197
	[-2.648]	**[-3.560]**		**[2.380]**		
JPY	**-0.223**	**-0.268**	0.017	**0.569**	9.74%	197
	[-2.412]	**[-3.108]**	[1.048]	**[2.399]**		
EUR	**-0.376**			**0.554**	7.15%	197
	[-4.621]			**[2.052]**		
EUR	**-0.379**	0.023		**0.545**	6.84%	197
	[-4.849]	[0.302]		**[1.837]**		
EUR	**-0.371**	-0.147	0.034	**0.520**	8.02%	197
	[-4.169]	[-0.822]	[1.261]	**[1.876]**		

Table 2.4: Forecast of $R_{hml,t}$ Using $SKEW_{t-1}$ 1996-2012

Reports the results of equation (2.12): forecasts of HML returns using the $SKEW$ level which is useful in turning our results into a trading strategy that avoids any look-ahead in parameters.

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.524**			**2.695**	6.26%	201
	[-3.778]			**[4.217]**		
US	**-0.484**	-0.154		**2.581**	10.41%	201
	[-3.506]	[-1.559]		**[4.298]**		
US	**-0.507**	-0.480	0.065	**2.650**	13.54%	201
	[-4.396]	[-1.885]	[1.568]	**[4.671]**		
JPY	**-0.285**			**1.778**	2.05%	201
	[-2.203]			**[3.067]**		
JPY	**-0.238**	**-0.180**		**1.646**	9.58%	201
	[-2.014]	**[-3.573]**		**[3.152]**		
JPY	**-0.246**	**-0.286**	0.021	**1.668**	9.61%	201
	[-2.002]	**[-3.131]**	[1.253]	**[3.084]**		
EUR	**-0.380**			**2.245**	5.65%	201
	[-3.651]			**[4.756]**		
EUR	**-0.386**	0.025		**2.263**	5.39%	201
	[-3.602]	[0.346]		**[4.221]**		
EUR	**-0.401**	-0.181	0.041	**2.307**	7.41%	201
	[-3.203]	[-0.956]	[1.397]	**[3.818]**		

ensure this strategy is realistic. We begin by computing a forecast for returns on HML using

$$R_{hml,t-1} = a + \beta_s SKEW_{t-2} + \epsilon_{t-1} \tag{2.13}$$

allowing 36 months burn in period for estimation of β_s. Based on this model we compute the forecast for the following month's HML return, $\hat{R}_{hml,t}$. Assuming an endowment of $\mathcal{W}_{t-1}$, we build a portfolio by putting $w_{hml,t-1}$ of $\mathcal{W}_{t-1}$ into HML. Assuming that margin accounts pay no interest rate and require 50% of the absolute value of the position (so that if one wants to go long HML by purchasing "H" and selling "L" then one has to put up margin equivalent to half of the position), the remainder of the endowment, $(1 - |w_{hml,t-1}|)\mathcal{W}_{t-1}$, is invested into the risk free rate[4]. The weight is defined as $w_{hml,t-1} = \tanh\left(\frac{\hat{R}_{hml,t}}{\sqrt{\frac{1}{t}\sum_{j=1}^{t}(\hat{R}_{hml,j} - \overline{\hat{R}}_{hml,t})^2}}\right)$. This is the hyperbolic tangent of the forecasted HML return scaled by the standard deviation of previous HML forecasts. The hyperbolic tangent is applied so that $w_{hml,t-1} \in [-1, 1]$. The gross return to this portfolio $\frac{\mathcal{W}_t}{\mathcal{W}_{t-1}} = 1 + w_{hml,t-1}R_{hml,t} + (1 - |w_{hml,t-1}|)R_{free,t}$ where $R_{free,t}$ is the risk free rate realized at time t. Each successive month, the window over which the model is estimated expands but always only includes historical data. Rebalancing is done monthly for both active and passive HML. The monthly rebalancing for passive HML assures that the investor puts half of his wealth in being long "H" and half into being short "L". The cumulative return to passive and active HML is presented in Figure 2.4. The shaded regions represent times when the strategy is short HML while the white areas represent times when the strategy is long HML.

As is evident from the plot, the position direction is quite persistent; since HML is such a large aggregate, transaction costs here are also minimal. However, the performance of active HML is significantly better than passive HML. An investment of \$1.00 in HML in 1999 becomes roughly \$1.50 by the end of 2012. This same dollar invested in active HML becomes roughly \$2.50 by the end of 2012. The annualized information ratio, $IR \equiv \frac{E(R^{active}_{hml,t} - R^{passive}_{hml,t})}{\sigma(R^{active}_{hml,t})}$, for this strategy is .36. The times when this strategy is short HML correspond to significantly anomalous market conditions. For example we see that this strategy is short HML during

[4] We assume that $\mathcal{W}_{t-1} > 0$; if at any point this condition is violated the strategy stops.

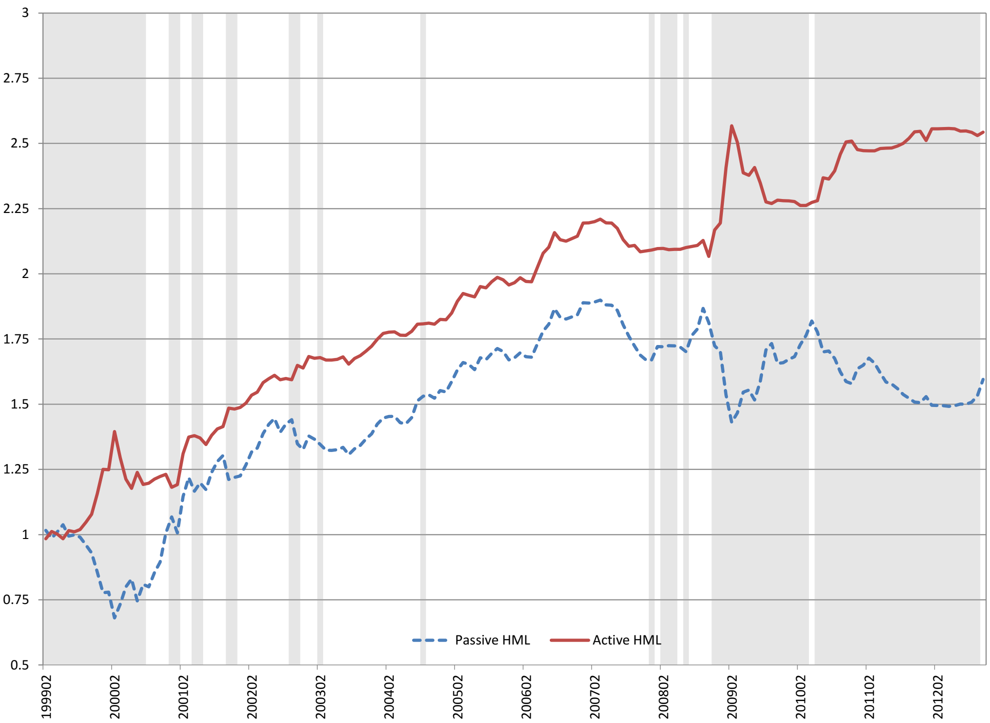


Figure 2.4: Cumulative Returns to Passive and Active HML Strategy

This plot demonstrates the out-of-sample forecasting performance of $SKEW_t$ in timing returns to $R_{hml,t}$. To construct the strategy we run a forecasting regression of $R_{hml,t-1} = a + \beta_s SKEW_{t-2} + \epsilon_{t-1}$ using 36 months of lagged data (this is the burn in period). Based on this model we compute the forecast for next month's HML return, $\hat{R}_{hml,t}$. To determine how much to invest each month we assume an endowment of $\mathcal{W}_{t-1}$ and Build a portfolio by putting $w_{hml,t-1}$ of it into HML and $(1 - |w_{hml,t-1}|)$ into the risk free rate where $w_{hml,t-1} = \tanh\left(\frac{\hat{R}_{hml,t}}{\sqrt{\frac{1}{J}\sum_{j=1}^{t}(\hat{R}_{hml,j} - \overline{\hat{R}}_{hml,t})^2}}\right)$. This is the hyperbolic tangent of the forecasted HML return scaled by the standard deviation of previous HML forecasts. The hyperbolic tangent is applied so that $w_{hml,t-1} \in [-1, 1]$. The return to this portfolio is $1 + R_{p,t} = 1 + w_{hml,t-1}R_{hml,t} + (1 - |w_{hml,t-1}|)R_{free,t}$. The mechanism described avoids look ahead bias in values and parameters. This cumulative return is plotted in the figure and labeled Active HML. Passive HML corresponds to monthly rebalancing strategy that invests equal weights into being long "H" and short "L".

the run-up in tech stocks of the late 90's. It is also short HML during/post the 2008 financial crisis. These two periods account for a substantial portion of the profit generated by this strategy. This is to be expected: we are attempting to pick up states of the world when the price investors are willing to pay for portfolio insurance spikes and these states should not be a frequent occurrence.

While we have been fortunate that our time sample includes diversity in business cycle conditions, we are still constrained by the availability of options data from OptionMetrics. To verify that our results work in other time samples - a truly out of sample test - we attempt to impute the value of $SKEW$ based on quantities that $SKEW$ should be picking up. To extend the sample in a principled way, we use the LASSO variable selection method of Tibshirani (1996). This methodology is an ℓ^1 penalized regression that selects variables among a candidate set that best capture the true relationship and kicks out all irrelevant ones. The set of possible variables that we include is $I_t \equiv \{\overline{\kappa}_3, R_{hml,t}, R^e_{m,t}, dy_t, tp_t, dp_t\}$ and I_{t-1}. $\overline{\kappa}_3$ is the physical third cumulant of market returns computed over the past 3 months of daily data, $R_{hml,t}$ is the US HML return, $R^e_{m,t}$ is the market return, dy_t is the dividend yield, tp_t is the term premium, dp_t is the default premium. The logic for including these variables is simple: $\overline{\kappa}_3$ could capture the portion of $SKEW$ corresponding to the physical distribution (though based on Pan (2002) we know that it will be a very small effect), and dy_t, tp_t, and dp_t could capture portions of the jump risk premium.

To operationalize this technology, we use five-fold cross validation using the 1996-2012 sample, to fit the LASSO model to the $SKEW$ using I_t and I_{t-1}. The cross validation is needed to select the shrinkage parameter that LASSO uses to determine how aggressive it should be in shrinking regression coefficients. For a particular shrinkage parameter, we divide the sample (1996-2012) into 5 sections, pick a section to leave out, fit LASSO over the remaining four sections and compute the root-mean-square error of the model in forecasting the level of $SKEW$ in the left out section; we then leave a different section out and repeat the process. The average root-mean-square error for this particular value of the shrinkage parameter is recorded. A shrinkage parameter is selected that creates the lowest forecasting

Table 2.5: Extending the Sample: 1963 - 1996

We extend the sample to 1963 by replicating $SKEW$ using other variables, termed $\widehat{SKEW}$. This table reports the results of forecasting returns on HML using innovations in $\widehat{SKEW}$ and the market using $\widehat{SKEW}$ from 1963 - 1996.

Dependent Variable	$\hat{\varepsilon}_{skew,t-1}$	$R_{m,t}$	$\widehat{SKEW}_{t-1} \cdot R_{m,t}$	$\widehat{SKEW}_{t-1}$	$R_{m,t-1}$	(Intercept)	R^2	N
$R_{hml,t}$	**-0.519**					**0.461**	1.49%	396
	[-2.472]					**[3.078]**		
$R_{hml,t}$	**-0.413**	**-0.198**				**0.551**	12.70%	396
	[-2.274]	**[-5.239]**				**[3.936]**		
$R_{hml,t}$	**-0.423**	**-0.217**	0.006			**0.548**	12.50%	396
	[-2.069]	**[-2.400]**	[0.228]			**[3.977]**		
$R_{m,t}$				**0.443**		-0.781	0.85%	396
				[2.107]		[-1.277]		
$R_{m,t}$				**0.479**	0.071	-0.914	1.09%	396
				[2.360]	[1.364]	[-1.512]		

error. This shrinkage parameter corresponds to a particular set of variables out of the available set. A simple OLS model is then fit from 1996-2012 using the selected set of variables and is used to compute a fitted value of $SKEW$, termed $\widehat{SKEW}$, going back to 1963.

We validate that our results hold over this significantly longer non-overlapping sample period: we show that $\widehat{SKEW}$ has forecasting power for R_m^e and R_{hml} is slow to respond to innovations in $\widehat{SKEW}$, termed $\hat{\varepsilon}_{skew}$. Using the sample from 1963 - 1996, the first part of the Table 2.5 shows that R_{hml} responds slowly to innovations in $\hat{\varepsilon}_{skew}$: a one volatility point increase in $\widehat{SKEW}$ corresponds to 40 - 50 basis points poorer performance in HML the following month. The second part of the table shows that $\widehat{SKEW}$ is capable of forecasting the market return the following month as we saw in the 1996 - 2012 sample using $SKEW$. The R^2 in these regressions is significantly lower as one would expect: we are attempting to capture characteristics of a variable that is best reflected through option prices and thus our ability to do this is limited. However, it is reassuring that we can capture a relevant portion of it to validate our results.

2.3.2 Corporate Bonds

We have shown that HML responds to innovations in $SKEW$ slowly in three regions and over a lengthy time sample; these results led us directly to an implementable trading strategy. Do these results hold in other asset classes? We examine corporate bonds to try and answer this question. At the same time, it provides another out of sample test of our results regarding slow investor reaction to changes in crash risk. Corporate bonds are a natural asset class to examine because they contain a large cross-section (like equities) and have assets that are considered risky in absolute terms (junk bonds) by investors and those that are considered safe (investment grade bonds). Bank of America/ML provides total return indices by rating category (AAA through CCC) which we use as basis assets.

We first examine if these assets have differential exposure to innovations in $SKEW$ in Table 2.6. We see clearly that AAA bonds enjoy a positive return while CCC bonds have a negative return when $SKEW$ increases contemporaneously. Thus the long CCC, short AAA trade performs poorly when crash risk increases; this is the same type of result that we saw with value and growth stocks.

We next ask the question: do these securities incorporate changes in risk premia into their prices efficiently? We saw that the cross-section of equities does not and thus HML incorporates changes in $SKEW$ with a delay. To answer this question we run the analysis in equation (2.7) replacing $R_{hml,t}$ on the left hand side of the regression with $R_{CCC-AAA,t}$: the return of CCC bonds minus the return on AAA bonds. Figure 2.5 presents the results of this analysis. At lag zero we see the contemporaneous results presented in Table 2.6: the long junk short investment grade (CCC-AAA) trade performs poorly when there is a positive innovation in $SKEW$. However, we also see that it proceeds to perform poorly the following month as well (until reversing in month two). That is, these securities are also slow to fully incorporate all available information into their prices though they are more expedient than the cross-section of equities (which took 4 months). One can speculate regarding the reason for this: one story might be that corporate bonds have a more sophisticated investor base since fewer retail investors participate actively in the bond market. Another reason might be that

Table 2.6: Corporate Bond Returns and ε_{skew}

We document that AAA bonds perform well when $SKEW$ increases while CCC bonds perform poorly. Thus the trade that goes long CCC and short AAA bonds behaves like HML with respect to innovations in $SKEW$.

Dependent Variable	$\varepsilon_{skew,t}$	$\varepsilon_{vix,t}$	$R_{m,t}$	(Intercept)	R^2	N
$R_{aaa,t}$	0.228	-0.049	-0.006	**0.271**	1.08%	201
	[1.699]	[-0.532]	[-0.143]	**[2.741]**		
$R_{aa,t}$	0.154	-0.066	0.006	**0.277**	2.65%	201
	[1.527]	[-0.876]	[0.142]	**[2.734]**		
$R_{a,t}$	0.167	-0.113	0.021	**0.298**	8.28%	201
	[1.565]	[-1.053]	[0.419]	**[2.478]**		
$R_{bbb,t}$	-0.117	-0.055	0.072	**0.338**	11.48%	201
	[-0.798]	[-0.688]	[1.506]	**[2.547]**		
$R_{bb,t}$	-0.296	-0.117	**0.166**	**0.392**	39.26%	190
	[-1.454]	[-1.789]	**[2.948]**	**[2.384]**		
$R_{b,t}$	**-0.427**	**-0.115**	**0.259**	0.258	46.08%	190
	[-2.167]	**[-2.216]**	**[5.063]**	[1.549]		
$R_{ccc,t}$	**-0.720**	-0.121	**0.454**	0.337	46.06%	190
	[-2.259]	[-1.532]	**[3.696]**	[1.247]		
$R_{ccc,t} - R_{aaa,t}$	**-0.945**	-0.070	**0.465**	-0.181	44.34%	190
	[-2.413]	[-0.854]	**[3.810]**	[-0.650]		

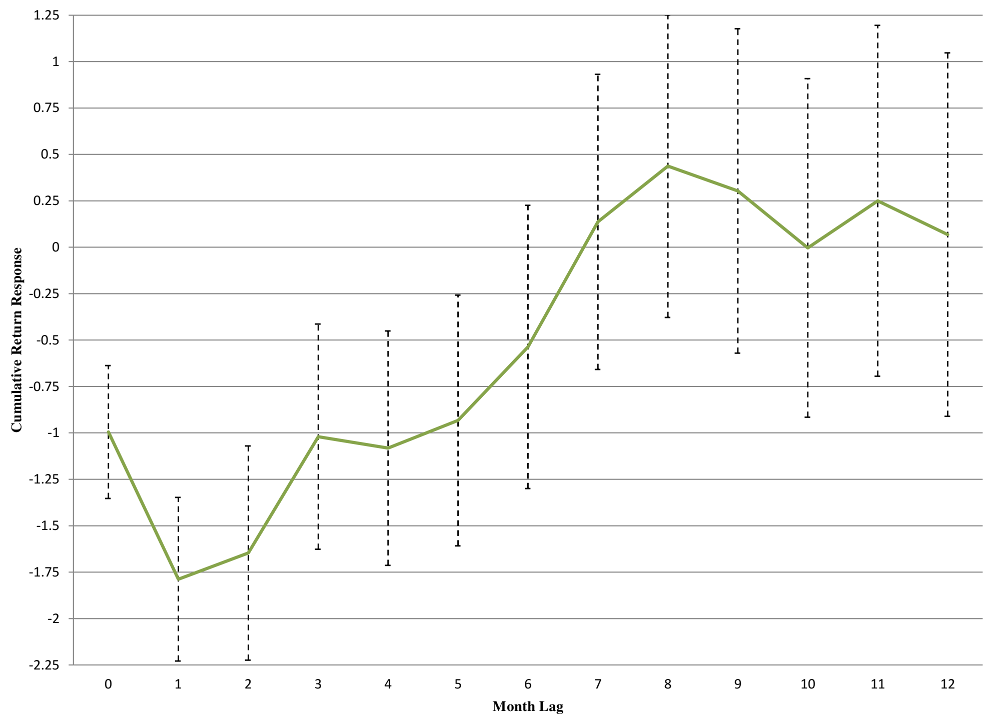


Figure 2.5: Cumulative CCC-AAA Response to ε_{skew}

Plot of the cumulative response to innovations in $SKEW$ of a trade that goes long CCC and short AAA bonds. Corporate bonds, like equities, are also slow to fully incorporate all information from $SKEW$.

jump risk is highly relevant for credit investors since they are concerned about bankruptcy. A significantly negative jump in the market (and thus in equity valuations) could drastically alter the probability of bankruptcy. Therefore, credit investors may be more sensitive to this particular information than equity investors.

These results can be presented in a regression framework using equation (2.11) replacing R_{hml} on the left hand side with returns on each bond rating category and using one lag of $SKEW$ innovations as opposed to four based on Figure 2.5. Table 2.7 presents the results of these regressions. We see that a one point increase in $SKEW$ predicts a -77 basis point return to $R_{CCC-AAA}$ the following month. These results confirm that underreaction to changes in crash risk is prevalent in the financial markets.

Table 2.7: Bond Return Forecasts Using $\varepsilon_{skew,t-1}$

This table shows that lagged innovations in $SKEW$ are able to forecast corporate bond returns highlighting that this asset class is also slow to incorporate information from the options market into prices.

Dependent Variable	$\varepsilon_{skew,t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
$R_{aaa,t}$	**0.397**	0.034	-0.006	**0.276**	7.45%	200
	[2.374]	[0.493]	[-0.494]	**[2.636]**		
$R_{aa,t}$	**0.318**	0.051	-0.004	**0.271**	6.49%	200
	[2.165]	[0.761]	[-0.401]	**[2.653]**		
$R_{a,t}$	**0.365**	0.146	-0.014	**0.276**	10.04%	200
	[2.176]	[1.365]	[-0.954]	**[2.127]**		
$R_{bbb,t}$	0.233	0.283	-0.032	**0.330**	16.27%	200
	[1.856]	[1.946]	[-1.438]	**[2.327]**		
$R_{bb,t}$	0.073	**0.445**	-0.033	**0.334**	36.43%	190
	[0.693]	**[2.537]**	[-1.271]	**[2.030]**		
$R_{b,t}$	-0.262	**0.500**	-0.024	0.198	43.02%	190
	[-1.670]	**[3.386]**	[-1.063]	[1.090]		
$R_{ccc,t}$	-0.376	**0.569**	0.009	0.226	43.44%	190
	[-1.504]	**[2.036]**	[0.201]	[0.793]		
$R_{ccc,t} - R_{aaa,t}$	**-0.775**	**0.545**	0.014	-0.277	43.70%	190
	[-2.897]	**[2.090]**	[0.337]	[-0.914]		

2.4 Investor Trading Behavior

We have shown that prices are slow to fully incorporate information from $SKEW$ into HML. A large behavioral/frictional literature has documented significant delays in price reactions, discussed nicely by Hong & Stein (1999), Duffie (2010). Ang et al. (2012) show that options on individual securities have relevant information for future returns of those securities not captured by the standard risk factors. They propose that sophisticated investors express their views in the options market and this information is reflected in individual stock returns slowly.

Our work finds results that are similar: innovations in $SKEW$ contain information about future HML returns. While there is a clear contemporaneous reaction as shown in Figure 2.3, there is also a significant delay. We confirm that investors are in fact slow to react to this information by examining mutual fund flows. Just like HML returns, we find that flows into value and growth mutual funds are predictable using past innovations in $SKEW$. A positive innovation in $SKEW$ causes investors to withdraw money from value funds while not withdrawing money from growth funds.

We first classify funds by value style using their four factor exposure

$$R_{f,t} = a + \beta_{m,t}R_{m,t} + \beta_{smb,t}R_{smb,t} + \beta_{hml,t}R_{hml,t} + \beta_{umd,t}R_{umd,t} + \epsilon_{f,t} \tag{2.14}$$

using 36 months rolling regression as in Chan et al. (2002) among others. Funds are arranged into quintiles each month based on last month's HML exposure: $\beta_{hml,t-1}$. We further define

$$FLOW_{f,t} \equiv \frac{TNA_{f,t} - TNA_{f,t-1}(1 + R_{f,t})}{TNA_{f,t-1}} \tag{2.15}$$

for each fund f and month t where TNA is the total net assets and $R_{f,t}$ is the return of fund f. This is the percentage increase/decrease in the assets of the fund due to contributions/withdrawals by investors. Additionally as Chevalier & Ellison (1997) show, flows into mutual funds are highly dependent on the funds' past performance; we are sure to condition on this in our analysis so that any effect in investor behavior we find is due to information in $SKEW$ as opposed to past fund returns. Therefore, we also compute the 1 year rolling cumulative return of each fund: $R_{f,t-12\to t-1}$. Then, for each value style bucket (1 - 5) a TNA_{t-1} weighted average of $FLOW_{f,t}$ (termed $FLOW_{b,t}$ for $b \in \{1,2,..,5\}$) and $R_{f,t-12\to t-1}$ (termed $R_{b,t-12\to t-1}$) is taken. We regress:

$$FLOW_{b,t} = a + \beta_{b,skew}\varepsilon_{skew,t-1} + \beta_{b,ret}R_{b,t-12\to t-1} + \nu_t \tag{2.16}$$

Results in Table 2.8 show that $\beta_{b,skew}$ is highly significant and negative for value stocks but roughly zero or positive for growth stocks. Investors pull money out of value funds in response to an increase in market crash risk with a lag. The magnitudes are significant: one volatility point increase in $SKEW$ causes a .2% outflow from value funds relative to growth funds the following month. This underreaction by investors to the information expressed by option market participants drives the predictability of HML returns.

One may be concerned that mutual funds pre-position their portfolios in anticipation of flows and thus dampen the effect of lagged information on stocks. Consider a savvy value mutual fund manager who received an inflow of money this month: he may decide that there is a good chance that he will also receive an inflow of money the next month since flows are persistent. Knowing this fact, he uses his cash position (or takes a loan from a bank) to purchase value stocks this month anticipating to return his cash position (or pay the loan) to equilibrium the next month. This would serve to dampen the predictability of value stocks due to investor underreaction. To account for this fact we extract innovations from flows into each quintile, $\varepsilon^{flow}_{b,t}$, and treat these as unexpected flows to the fund manager. We then regress these unexpected flows on $\varepsilon_{skew,t-1}$ and past fund returns, $R_{b,t-12\to t-1}$:

$$\varepsilon^{flow}_{b,t} = a + \beta_{b,skew}\varepsilon_{skew,t-1} + \beta_{b,ret}R_{b,t-12\to t-1} + \nu_t \tag{2.17}$$

The results of this regression are presented in Table 2.9. This, however, does not alter our conclusions: innovations in $SKEW$ predict unexpected flows into growth and value stocks. Investors withdraw .1% from value funds relative to growth funds in response to one volatility point increase in $SKEW$ the previous month.

Table 2.8: Flows Into Value and Growth Funds

Mutual funds are sorted into style quintiles according to $\beta_{hml,t-1}$ in a regression of $R_{f,t} = a + \beta_{m,t}R^e_{m,t} + \beta_{smb,t}R_{smb,t} + \beta_{hml,t}R_{hml,t} + \beta_{umd,t}R_{umd,t}$ - where $R_{f,t}$ is the return of fund f at time t - as in Chan et al. (2002) among others. $FLOW_{f,t} \equiv \frac{TNA_{f,t}-TNA_{f,t-1}(1+R_{f,t})}{TNA_{f,t-1}}$ is computed for each fund f and month t. To control for the flow performance relationship cumulative returns over the past year are also computed and denoted by $R_{f,t-12\to t-1}$. Then for each quintile an asset (TNA_{t-1}) weighted average is taken across $FLOW_{b,t} \equiv \sum_{f\in b} w_{f,t-1}FLOW_{f,t}$ and $R_{b,t-12\to t-1} \equiv \sum_{f\in b} w_{f,t-1}R_{f,t-12\to t-1}$. Then for each quintile we regress $FLOW_t = a + \beta_{skew}\varepsilon_{skew,t-1} + \beta_{ret}R_{t-12\to t-1} + \nu_t$. The table shows that an increase in $SKEW_t$ causes investors to pull money away from value funds.

Value Style Quintile	$\varepsilon_{skew,t-1}$	$R_{t-12\to t-1}$	(Intercept)	R^2	N
1	0.007		-0.182	-0.50%	199
	[0.164]		[-1.036]		
	0.035	**0.016**	-0.362	16.88%	199
	[0.821]	**[2.678]**	[-3.111]		
2	-0.099		0.089	1.31%	199
	[-1.988]		[1.002]		
	-0.082	**0.010**	-0.019	5.22%	199
	[-1.592]	**[1.597]**	[-0.168]		
3	-0.027		0.007	-0.14%	199
	[-0.632]		[0.167]		
	-0.012	**0.008**	-0.066	4.30%	199
	[-0.258]	**[2.120]**	[-1.072]		
4	**-0.137**		-0.055	4.25%	199
	[-3.826]		[-0.518]		
	-0.117	**0.009**	-0.147	7.82%	199
	[-2.507]	**[2.192]**	[-1.535]		
5	**-0.208**		0.057	4.73%	199
	[-2.581]		[0.377]		
	-0.164	**0.017**	-0.139	12.76%	199
	[-2.039]	**[3.289]**	[-0.792]		
5-1	**-0.215**		0.239	1.66%	199
	[-2.076]		[0.675]		
	-0.163	**0.057**	0.215	53.75%	199
	[-1.980]	**[3.742]**	[1.387]		

Table 2.9: Unexpected Flows Into Value and Growth Funds

This table reports the results of equation (2.17) showing that innovations in $FLOW$ (relative to past values of $FLOW$) are predictable using lagged innovations in $SKEW$.

Value Style Quintile	$\varepsilon_{skew,t-1}$	$R_{t-12\to t-1}$	(Intercept)	R^2	N
1	**0.045**		-0.038	0.04%	199
	[2.108]		[-0.930]		
	0.042	-0.002	-0.020	-0.06%	199
	[1.208]	[-0.685]	[-0.412]		
2	-0.032		-0.019	-0.17%	199
	[-0.784]		[-0.421]		
	-0.035	-0.002	0.000	-0.44%	199
	[-0.868]	[-0.558]	[0.001]		
3	-0.017		-0.027	-0.36%	199
	[-0.364]		[-0.808]		
	-0.022	-0.003	-0.001	-0.17%	199
	[-0.479]	[-0.793]	[-0.011]		
4	-0.051		-0.025	0.62%	199
	[-1.600]		[-0.694]		
	-0.059	-0.003	0.009	1.02%	199
	[-1.490]	[-1.328]	[0.178]		
5	-0.050		0.007	0.32%	199
	[-0.961]		[0.175]		
	-0.047	0.001	-0.006	-0.09%	199
	[-0.895]	[0.826]	[-0.144]		
5-1	**-0.102**		0.013	1.40%	199
	[-2.328]		[0.238]		
	-0.104	-0.002	0.014	1.10%	199
	[-2.383]	[-0.372]	[0.259]		

2.5 Discussion and Conclusion

We have shown that returns to HML (as well as corporate bonds) have a significant amount of predictability. Utilizing this predictability, we perform an "out-of-sample" test of performance: active HML significantly improves the returns to an investor relative to an investment in passive HML. A dollar invested in passive HML in 1999 grows to approximately $1.50 by the end of 2012. On the other hand, that same investment grows to roughly $2.50 in active HML. Furthermore, this predictability is easy to extract: it does not require complex computation just the implied volatility skew on the S&P 500. Using mutual fund flow data, we show that this predictability is due to delayed reaction by investors.

Our results have deep implications for theories attempting to explain high returns on HML: theories must now consider explaining returns to active HML (a much more difficult thing to do given the favorable return profile). More generally, our results relate to a large literature on slow moving capital and segmented markets. We show how a large, heavily examined factor can have a significant amount of return predictability due to the slow rotation into and out of value/growth stocks by investors. One may regard slight economic frictions that prevent small stocks from re-pricing perfectly as unimportant. However, the return predictability that we identify here is on an aggregate, economically meaningful level. A significant portion of this predictability comes from periods when the market experiences stress: during the tech bubble and during the financial crisis of 2008. We highlight that these periods can generate significant mispricing for large aggregates.

Chapter 3

Is Real Interest Rate Risk Priced? Theory and Empirical Evidence[1]

3.1 Introduction

Are expected returns related to covariance with shocks to the real risk free interest rate? Put differently, is the real risk free rate a priced state variable? Since Fama (1970), financial economists have understood that state variables can be priced if they are correlated with changes to (1) investor preferences or (2) the consumption-investment opportunity set.[2] Because the risk free rate is an equilibrium outcome that is sensitive to preferences and consumption-investment opportunities, it is a prime candidate to be a priced state variable.

Previous research primarily focuses on shocks to consumption-investment opportunities. For example, Merton (1973) Intertemporal Capital Asset Pricing Model (ICAPM) considers changing investment opportunities while holding preferences constant. Campbell (1993) follows the same approach to derive ICAPM pricing as a function of changes to expected returns. More recently, Bansal & Yaron (2004) initiated a literature on long-run consumption

[1] This chapter is co-authored jointly with Samuel Kruger

[2] Fama (1970) considered consumption and investment opportunities separately. In practice, these two opportunity sets are typically collapsed by considering a single homogeneous consumption good.

growth shocks in which expectations about future consumption growth are priced. In these frameworks, positive interest rate shocks are generally good news, which makes long-duration assets valuable hedges, reducing their risk premia.

In contrast, Albuquerque et al. (2016) (AELR) present an interesting model that considers preference shocks to investor patience. In their framework, positive interest rate shocks stem from impatience and are generally bad news, making long-duration assets more risky and increasing their risk premia. We examine a generalized version of the AELR model with both consumption-investment and preference shocks. Expected consumption growth and time preferences both impact interest rates, and covariance with these shocks is priced relative to the Capital Asset Pricing Model (CAPM) and the Consumption CAPM (CCAPM). However, the two types of interest rate risk carry different prices. Relative to both the CAPM and CCAPM, the price of interest rate risk associated with time preference shocks differs from the price of consumption growth interest rate risk by a factor of $\frac{-1}{\psi-1}$, where ψ is elasticity of intertemporal substitution. For $\psi > 1$, this means the two different interest rate risk premia have opposite signs. The current AELR model specification is undefined as ψ approaches 1 and time preference risk premia are very large when ψ is close to 1. This is an important point as we try to understand the true nature of this parameter. For example Hall (1988) argues that $\psi \approx 0$ while Guvenen (2006) notes that much of the macro literature has concluded that$\psi \approx 1$. Unless there is a well micro-founded reason to explicitly exclude the possibility of $\psi = 1$, we suggest it is beneficial to allow this parameter to take on the full range of values and be dictated by data.

Empirically, we estimate real interest rate shocks based on a vector autoregression (VAR) model of nominal interest rates, CPI inflation rates, and other state variables. When sorted based on interest rate exposure, stocks with high exposure have slightly lower expected returns, both on an absolute basis and relative to CAPM and Fama & French (1993) three factor model predictions. This evidence is consistent with risk premia required for time preference shocks and at odds with risk premia demanded for consumption-investment shocks. That said, the effects are modest, and the return differences are not statistically significant.

Moreover, the overall stock market appears to have very little exposure to interest rate risk. The market's interest rate news beta is an insignificant 0.11, which would carry a risk premium of -8 bps based on our cross-sectional pricing results. This evidence contradicts the conclusions of AELR, who claim that interest rate risk (valuation risk) explains the equity premium puzzle. The main difference between our empirical work and theirs is that we directly estimate covariance between excess returns and real interest rate shocks, whereas AELR do not estimate this moment in their GMM analysis. AELR's benchmark estimates imply that excess equity returns have a correlation of approximately -0.94 with interest rate shocks while we estimate this correlation as 0.05 in the data.

The rest of the paper is organized as follows: Section 3.2 presents the theoretical underpinnings of the generalized AELR model and shows that preferences with this specification are undefined for $\psi = 1$. We then derive the ICAPM that is consistent with $\psi \neq 1$ to show that risk premia take on large values in the neighborhood of 1. Section 3.3 presents our empirical analysis and findings that the real risk free rate is essentially uncorrelated with the stock market and isn't a priced state variable in the cross-section of stocks and bonds. Finally, Section 3.4 concludes.

3.2 Theory

We consider a model with shocks to consumption growth and time preferences. Thus, the model violates both of Fama (1970) assumptions. Interest rate shocks are priced relative to the CAPM and the CCAPM. The model essentially nests the long-run risk consumption growth shocks of Bansal & Yaron (2004) with the valuation shocks of AELR. The main result is that consumption growth interest rate risk has a different price than time preference interest rate risk, and the two risk premia have opposite signs when elasticity of intertemporal substitution is greater than one. Our main results are presented and discussed below; detailed derivations are in the appendix.

3.2.1 Setup and General Pricing Equations

Following AELR, we consider a representative agent with recursive utility function:

$$U_t = \left[\lambda_t C_t^{1-1/\psi} + \delta \left(U_{t+1}^{*}\right)^{1-1/\psi}\right]^{1/(1-1/\psi)} \tag{3.1}$$

where C_t is consumption at time t, δ is a positive scalar capturing time discounting, ψ is elasticity of intertemporal substitution, and $U_{t+1}^{*} = \left\{E_t\left[U_{t+1}^{1-\gamma}\right]\right\}^{1/(1-\gamma)}$ is the certainty equivalent of future utility with relative risk aversion of γ. The function is defined for $\psi \neq 1$ and $\gamma \neq 1$. This utility function represents standard Epstein-Zin (EZ) preferences of Epstein & Zin (1991) and Weil (1989) except that time preferences are allowed to vary over time instead of being constant[3]. Time preferences are affected by $\frac{\lambda_{t+1}}{\lambda_t}$, which is assumed known at time t. These preferences relax the traditional restraint of recursive preferences that the aggregator function is independent of time and state. Specifically, the preferences imply dropping assumption A3 of Skiadas (2009). This utility function is not defined for $\psi = 1$. We consider alternative preferences that are defined for $\psi = 1$ at the end of this section and in the appendix.

If we restrict $\psi \neq 1$, then using standard techniques for working with EZ preferences, AELR show that equation (3.1) implies a log stochastic discount factor of:

$$m_{t+1} = \theta \log\left(\delta \frac{\lambda_{t+1}}{\lambda_t}\right) - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)\, r_{w,t+1} \tag{3.2}$$

where

$$\theta = \frac{1-\gamma}{1-1/\psi} \tag{3.3}$$

Lower case letters signify logs. Δc_{t+1} is log consumption growth from period t to period $t+1$. $r_{w,t+1}$ is the log return on the overall wealth portfolio. This stochastic discount factor is standard for EZ preferences except that time discounting (δ) is augmented by $\frac{\lambda_{t+1}}{\lambda_t}$.

We assume that innovations to consumption and expected future consumption are jointly lognormal and homoskedastic. Similarly, innovations to time preferences and expected time

[3]Other authors have looked at allowing time preferences to vary, for example: Maurer (2012), Normandin et al. (1998)

preferences are jointly lognormal and homoskedastic. Formally,

$$E_t\left[c_{t+a}\right] = E_{t-1}\left[c_{t+a}\right] + \varepsilon^c_{a,t} \tag{3.4}$$

$$E_t\left[\lambda_{t+1+b}\right] = E_{t-1}\left[\lambda_{t+1+b}\right] + \varepsilon^\lambda_{b,t} \tag{3.5}$$

with $\left[\left\{\varepsilon^c_{a,t}\right\}_{a>0}, \left\{\varepsilon^\lambda_{b,t}\right\}_{b>0}\right]$ distributed jointly normally with constant variance (i.e., $cov_t\left(\varepsilon^c_{a,t}, \varepsilon^\lambda_{b,t+1}\right) = V$ for all t).[4] This implies that excess returns on the wealth portfolio are lognormal and homoskedastic. For simplicity, we assume that all other excess returns are lognormal as well. Lognormality and homoscedasticity simplify the model and ensure that risk premia are constant over time, focusing attention on interest rate shocks. In their benchmark model, AELR specify a more restrictive stochastic process for λ_{t+1} and assume that expected consumption growth is constant over time. Similarly, Bansal & Yaron (2004) specify a more restrictive consumption growth process in their fluctuating growth rates model.

The stochastic discount factor of equation (3.2) can be used to price all assets. In particular, it implies a risk free rate of[5]:

$$r_{f,t+1} = -\log\left(\delta\frac{\lambda_{t+1}}{\lambda_t}\right) + \frac{1}{\psi}E_t\left[\Delta c_{t+1}\right] - \frac{1-\theta}{2}\sigma^2_w - \frac{\theta}{2\psi^2}\sigma^2_c \tag{3.6}$$

and risk premia of:

$$E_t\left[r_{i,t+1}\right] - r_{f,t+1} + \frac{1}{2}\sigma^2_i = \frac{\theta}{\psi}\sigma_{ic} + (1-\theta)\,\sigma_{iw} \tag{3.7}$$

σ^2_w is the variance of excess returns to the wealth portfolio. $\sigma^2_c = var_t\left(\varepsilon^c_{0,t+1}\right)$ is consumption variance relative to expectations last period. σ_{ic} is covariance of asset i's return with current consumption shocks. σ_{iw} is covariance of asset i's return with wealth portfolio returns. $\frac{1}{2}\sigma^2_i$ is a Jensen's inequality correction for expected log returns using variance of asset i's return. From equations (3.6) and (3.7), it is clear that the real risk free interest rate changes over time in response to time preferences ($\frac{\lambda_{t+1}}{\lambda_t}$) and expected consumption growth ($E_t\left[\Delta c_{t+1}\right]$)

[4]Note that λ_{t+1} is known one period in advance so time t shocks to λ expectations start with λ_{t+1}.

[5]We work with real variables in our analysis though AELR also provide a version of their model that specifies an inflation process.

and that risk premia are constant over time.

3.2.2 Substituting out Consumption (ICAPM)

Following Campbell (1993), we log-linearize the representative agent's budget constraint ($W_{t+1} = R_{w,t+1}(W_t - C_t)$) to yield:

$$r_{w,t+1} - E_t[r_{w,t+1}] = (E_{t+1} - E_t)\sum_{j=0}^{\infty}\rho^j \Delta c_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{w,t+1+j} \qquad (3.8)$$

where ρ is a log-linearization constant.[6] Because risk premia are constant over time, $News_{h,t+1} \equiv (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{w,t+1+j}$ depends solely on changes to expected interest rates, which change over time in response to time preferences and expected consumption growth as described by equation (3.6).[7] We use the identity (3.8) and the risk free rate decomposition, equation (3.6), to substitute out current consumption covariance from the risk premia in equation (3.7).

These substitutions yield the following ICAPM:

$$E_t[r_{i,t+1}] - r_{f,t+1} + \frac{1}{2}\sigma_i^2 = \gamma\sigma_{iw} + (\gamma - 1)\sigma_{ih(c)} - \frac{\gamma - 1}{\psi - 1}\sigma_{ih(\lambda)} \qquad (3.9)$$

Risk premia are determined by covariance with the market and covariance with state variables related to future interest rates. $\sigma_{ih(c)}$ is covariance with consumption growth shocks to future interest rates. $\sigma_{ih(\lambda)}$ is covariance with time preference shocks to future interest rates. Together, they add up to covariance with overall interest rate news:

$$\begin{aligned}\sigma_{ih} &\equiv cov_t\left(r_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}\right) \\ &= \sigma_{ih(c)} + \sigma_{ih(\lambda)} \qquad (3.10)\end{aligned}$$

The risk prices in equation (3.9) are revealing. Market return risk (σ_{iw}) is priced by relative risk aversion (γ) as in other ICAPM models. Also consistent with other ICAPM

[6]Specifically, $\rho = 1 - \exp\left(\overline{c - w}\right)$ where $\overline{c - w}$ is the average log consumption-wealth ratio. We use a monthly coefficient value of $\rho = 0.996$ in our analysis.

[7]The h subscript follows the notation of Campbell (1993) to indicate hedging of future interest rates.

models, state variable covariance ($\sigma_{ih(c)}$ and $\sigma_{ih(\lambda)}$) is priced only if $\gamma \neq 1$. Yet, the two components of interest rate risk have different prices. Whereas $\sigma_{ih(c)}$ is priced by $\gamma - 1$, $\sigma_{ih(\lambda)}$ is priced by $-\frac{\gamma-1}{\psi-1}$. When $\psi > 1$, the prices have opposite signs, and if ψ is close to 1, time-preference risk is amplified relative to consumption growth risk. The key distinction between equation (3.9) and previous ICAPM models like Campbell (1993) is that we consider shocks to both consumption growth and time preferences. Because Campbell assumes constant preferences, he omits $\sigma_{ih(\lambda)}$ and treats σ_{ih} as equivalent to $\sigma_{ih(c)}$.

3.2.3 Substituting out Wealth Returns (CCAPM)

The budget constraint (equation 3.8) can also be used to substitute out covariance with wealth portfolio returns to express risk premia in terms of a generalized CCAPM along the lines of Bansal & Yaron (2004) long run risk model. The resulting pricing equation is:

$$E_t\left[r_{i,t+1}\right] - r_{f,t+1} + \frac{1}{2}\sigma_i^2 = \gamma\sigma_{ic} + \left(\gamma\psi - 1\right)\sigma_{ih(c)} - \frac{\gamma\psi - 1}{\psi - 1}\sigma_{ih(\lambda)} \tag{3.11}$$

Consumption risk (σ_{ic}) is priced by relative risk aversion (γ) as in the standard CCAPM. Consistent with Bansal & Yaron (2004), interest rate risk is only priced if $\gamma \neq 1/\psi$.[8] That is, interest rate risk is priced under general EZ preferences, but not under power utility. As in our ICAPM, the most striking thing about equation (3.11) is that the two types of interest rate risk are priced differently. Once again, time preference interest rate risk differs from consumption growth interest rate risk by a factor of $\frac{-1}{\psi-1}$.

Our ICAPM, equation (3.9), and generalized CCAPM, equation (3.11), are at odds with traditional reasoning about interest rate risk. If one considers only consumption growth shocks, positive interest rate shocks are good news for investors under typical parameter assumptions ($\gamma > 1$ for the ICAPM and $\gamma > 1/\psi$ for the CCAPM). Thus, assets that positively covary with interest rate shocks are risky and require extra risk premia relative to CAPM and CCAPM pricing. Campbell & Viceira (2002) use this logic to argue that

[8] Bansal & Yaron (2004) express their version of equation (3.11) in terms of future consumption growth. This is just a different way of describing the same relationship.

long term bonds are valuable hedges against interest rate decreases. If $\psi > 1$ and $\frac{1}{\psi-1}\sigma_{ih(\lambda)}$ dominates $\sigma_{ih(c)}$, the logic actually goes the opposite way: investors want to hedge against interest rate *increases*, making long term assets (including bonds) risky investments.

3.2.4 $\psi = 1$ and Disciplining Parameter Values

The typical strategy for working with EZ preferences to examine the $\psi = 1$ case is to take the limit of the utility function: this yields Cobb-Douglas style preferences. Specifically, if one defines the usual EZ value function[9]

$$V_t = \left[(1-\delta)C_t^{1-1/\psi} + \delta\left(V_{t+1}^*\right)^{1-1/\psi}\right]^{1/(1-1/\psi)}$$

then taking the limit as $\psi \to 1$ can be done by L'Hopital's rule (detailed derivation in the appendix):

$$\lim_{\psi\to 1} V_t = C_t^{(1-\delta)}\left(V_{t+1}^*\right)^{\delta}$$

However, this procedure cannot be performed with the AELR specification because the limit diverges. Specifically

$$\lim_{\psi\to 1} lnU_t = \infty$$

as long as $\lambda_t + \delta \neq 1$, a condition that certainly holds as λ_t is random. A simple rescaling, on the other hand, also does not solve the problem because it creates fundamentally different preferences. Consider, for example, re-specifying the value function by dividing the taste modifier (λ_t) and the time discount factor (δ) by $\lambda_t + \delta$ in the hopes of making the above limit converge:

$$V_t = \left[\frac{\lambda_t}{\lambda_t+\delta}C_t^{1-1/\psi} + \frac{\delta}{\lambda_t+\delta}\left(V_{t+1}^*\right)^{1-1/\psi}\right]^{1/(1-1/\psi)} \tag{3.12}$$

[9]Note that the multiplication of C_t by $(1-\delta)$ is simply a rescaling in this context because $(1-\delta)$ is a constant. This is not the same thing as multiplication by random variable.

In this specification, the taste shocks will accumulate. Specifically, the coefficient next to C_{t+1} will be a function of λ_t as well as λ_{t+1} which can be seen by simply iterating the value function forward one period. Consider a simple example: the agent lives for three periods, chooses consumption in period 1 and 2 and then simply consumes the remainder of wealth, W, in the third period. There is no uncertainty in this case as wealth is not random and assume all values of λ are known. Then the optimization solved by the agent is:

$$V_0 = \max_{\{C_0, C_1\}} \left[(1-\beta_0)\, C_0^{1-\frac{1}{\psi}} + \beta_0\,(1-\beta_1)\, C_1^{1-\frac{1}{\psi}} + \beta_0\beta_1(1-\beta_2) C_2^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{3.13}$$

where $C_2 = W - C_0 - C_1$ and $\beta_t = \frac{\delta}{\lambda_t + \delta}$. The first order conditions for this problem imply

$$\frac{1}{\delta}\frac{\lambda_0}{\lambda_1}(\lambda_1 + \delta)\left(\frac{C_0}{C_1}\right)^{-\frac{1}{\psi}} = 1 \tag{3.14}$$

while the same setup with the AELR specification yields first order conditions that imply

$$\frac{1}{\delta}\frac{\lambda_0}{\lambda_1}\left(\frac{C_0}{C_1}\right)^{-\frac{1}{\psi}} = 1 \tag{3.15}$$

The tradeoff between consuming in the first and second period is fundamentally different in the two sets of preferences. Therefore, a simple rescaling does not solve the problem and the AELR preferences as currently specified do not admit a parameter value of $\psi = 1$. It is our view that leaving the model free to select values in this region is important in context of the aforementioned debate in the macro literature.

More generally, the insight of AELR preferences is they are able to insert a wedge between observed consumption and asset returns which have a very low correlation in the data and thus are problematic for many asset pricing models that predict a high correlation. Moreover, the state variable that breaks this correlation is observable via the risk free rate. One can take these ideas and attempt to create a preference relation that allows for all values of ψ and retains these properties. Imagine specifying the value function as

$$V_t = \left[(1-\delta) H_t(C_t)^{1-\frac{1}{\psi}} + \delta \left(E_t V_{t+1}^{1-\gamma} \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{3.16}$$

where H_t is a time varying function of consumption. It is simply another way to introduce the idea of taste shocks that transform the flow of utility in each period to create a wedge between asset prices and consumption. The limit of this value function as $\psi \rightarrow 1$ is well defined and is simply the Cobb-Douglas representation

$$\lim_{\psi \rightarrow 1} V_t = H_t(C_t)^{1-\delta} \left(E_t V_{t+1}^{1-\gamma} \right)^{\frac{\delta}{1-\gamma}} \tag{3.17}$$

yielding an SDF of

$$M_{t+1} = \delta \frac{V_{t+1}^{1-\gamma}}{\left(E_t V_{t+1}^{1-\gamma} \right)} \left(\frac{H_{t+1}(C_{t+1})}{H_t(C_t)} \right)^{-1} \left(\frac{H'_{t+1}(C_{t+1})}{H'_t(C_t)} \right) \tag{3.18}$$

where $H'(C) = \frac{dH(C)}{dC}$. In order for the time varying parameters in H to be observable in the risk free rate it must be the case that they appear in M_{t+1}. For example specifying $H_t(C_t) = \Lambda_t^* C_t$ would result in a model where the state variable would be entirely unobservable (this can be easily seen by the fact that Λ^* cancels in the above equation) and results in a model that has an extra degree of freedom to fit asset prices (namely, the covariance with a completely unobservable state variable). Obviously a different specification or more structure needs to be placed on $H_t(C_t)$ in order to have a model that isn't vacuous. We solve the $H_t(C_t) = \Lambda_t^* C_t$ case explicitly for a simple consumption and Λ^* process to highlight this fact. Defining $H_t(C_t) = \Lambda_t^* C_t$ implies

$$V_t = (\Lambda_t^*)^{1-\delta} C_t^{1-\delta} \left(E_t V_{t+1}^{1-\gamma} \right)^{\frac{\delta}{1-\gamma}} \tag{3.19}$$

$$M_{t+1} \quad = \quad \delta \frac{V_{t+1}^{1-\gamma}}{\left(E_t V_{t+1}^{1-\gamma} \right)} \left(\frac{C_{t+1}}{C_t} \right)^{-1}$$

which is the usual SDF specified in terms of the value function. To substitute out the value function, one can assume a process for log taste shock growth and log consumption growth and then guess-verify the value function. Assume (though this can be generalized using Wold's theorem to any time series process) that

$$\Delta c_{t+1} \quad = \quad \mu_c + \varepsilon_{t+1}^c \tag{3.20}$$

$$\Delta \lambda_{t+1}^* \quad = \quad \mu_{\lambda^*} + \varepsilon_{t+1}^{\lambda^*} \tag{3.21}$$

with $\varepsilon_t^i \sim N(0, \sigma_i^2)$, $\varepsilon_t^c \perp \varepsilon_t^{\lambda^*}$, $\Delta c_{t+1} \equiv \log(\frac{C_{t+1}}{C_t})$, $\Delta\lambda_{t+1}^* \equiv \log(\frac{\Lambda_{t+1}^*}{\Lambda_t^*})$. Furthermore, as in AELR, assume that Λ_{t+1}^* is known at time t. The log SDF is

$$m_{t+1} = \log\delta + (1-\gamma)v_{t+1} - \log\left(E_t V_{t+1}^{1-\gamma}\right) - \Delta c_{t+1} \tag{3.22}$$

Solving this using the method of undetermined coefficients and substituting these results into the log SDF, equation (3.22) yields the log SDF:

$$\begin{aligned} m_{t+1} &= \log\delta - \gamma\Delta c_{t+1} + \delta(1-\gamma)\Delta\lambda_{t+2}^* - (1-\gamma)(\mu_c + \delta\mu_\lambda) - \\ &- \frac{(1-\gamma)^2}{2}(\sigma_c^2 + \delta^2\sigma_{\lambda^*}^2) \end{aligned} \tag{3.23}$$

The risk free rate can be derived as usual and does not depend on λ^*:

$$r_{f,t+1} = -\log\delta + \mu_c + \frac{(1-2\gamma)}{2}\sigma_c^2 \tag{3.24}$$

On the other hand, the risk premia do depend on λ^* :

$$E_t r_{i,t+1} + \frac{1}{2}\sigma_i^2 - r_{f,t+1} = \gamma Cov(r_{i,t+1}, \Delta c_{t+1}) + \delta(\gamma-1)Cov(r_{i,t+1}, \Delta\lambda_{t+2}^*) \tag{3.25}$$

Therefore, expected excess asset returns are a function of log consumption growth as well as log taste growth. In this model, λ^* serves as a free parameter because it is never observed but influences risk premia. In other words, what matters for risk premia is λ^* while λ is what matters for the risk free rate.

While the case of $\psi = 1$ is not defined with the AELR specification, we can introspect about what are reasonable values for ψ and γ along the lines of Epstein et al. (2014). To generate better intuition for how close ψ can be to 1, we propose a thought experiment with simple consumption and time preference processes. Specifically, consider a three period economy with constant perishable consumption endowments of $C_0 = C_1 = C_2 = C$ in each period. Time preferences are known in advance for periods 0 and 1. For simplicity we assume $\lambda_0 = \lambda_1 = 1$ and we also assume $\delta = 1$. The only uncertainty in the economy is period 2 time preferences, which are revealed at time 1. λ_2 takes on two possible values, λ_H or λ_L with probabilities π_H and π_L, respectively. We want to know how the representative agent values wealth in state L relative to state H.

In the appendix, we derive Arrow-Debreu state prices for the two states and find that their ratio is:

$$\frac{P_L}{P_H} = \frac{\pi_L}{\pi_H}\left(\frac{1+\lambda_L}{1+\lambda_H}\right)^{-\frac{\gamma-1/\psi}{1-1/\psi}} \tag{3.26}$$

Note that these are prices at time 0 for state-contingent payoffs at time 1. Under power utility with $\gamma = 1/\psi$, the price ratio is simply the probability ratio. This is exactly what we should expect. With power utility, marginal utility of wealth is pinned down by consumption and current time preferences, which is constant across states. By contrast, state prices are highly sensitive to future time preferences when $1/\psi$ differs from γ and is close to 1. We do not have great intuition for whether $-\frac{\gamma-1/\psi}{1-1/\psi}$ should be positive or negative, but we believe its magnitude should be small.

To be more concrete, assume $\pi_L = \pi_H = 0.5$, $\lambda_H = 1$, and $\lambda_L = 0.9$. Table 3.1 presents the equation (3.26) state price ratio for these parameters at various values of γ and ψ. Parameterizations with $\gamma > 1$ are in Panel A. Parameterizations with $\gamma < 1$ are in Panel B. The upward sloping diagonals of 1's in both panels represent power utility with $\gamma = 1/\psi$.

What are reasonable values for $\frac{P_L}{P_H}$? The thought experiment is what you would pay for an extra dollar in a state in which time preferences will soon fall versus an extra dollar in a state in which time preferences will remain constant, keeping in mind that current and future consumption are the same in both states. As a starting point, we propose that it is difficult to rationalize state price ratios larger in magnitude than the ratio of the time preference shock itself. In Table 3.1, ratios between 0.95 and 1.05 are in bold italics, and ratios between 0.9 and 1.1 are highlighted in italics.[10] As expected, ratios in these ranges require $1/\psi$ to be close to γ or far from 1. For example, if γ is 5, ψ must be less than 0.44. With lower relative risk aversion, ψ can be closer to one without posing a problem.

[10]The broader range requires that $\frac{P_L\pi_H}{P_H\pi_L}$ falls between $\left(\frac{\lambda_L}{\lambda_H}\right)$ and $\left(\frac{\lambda_L}{\lambda_H}\right)^{-1}$. The narrower range requires that $\frac{P_L\pi_H}{P_H\pi_L}$ falls between $\left(\frac{1+\lambda_L}{1+\lambda_H}\right)$ and $\left(\frac{1+\lambda_L}{1+\lambda_H}\right)^{-1}$, which is equivalent to the condition that $|\gamma - 1/\psi| \leq |1 - 1/\psi|$.

Table 3.1: State Price Ratios

This table displays state price ratios from equation (3.26) at different values of relative risk aversion (RRA) and elasticity of intertemporal substitution (EIS).

A. RRA > 1

	Relative Risk Aversion								
EIS	**1.01**	**1.10**	**1.25**	**1.5**	**2**	**3**	**5**	**10**	**25**
0.04	***1.05***	***1.05***	***1.05***	***1.05***	***1.05***	***1.05***	***1.04***	***1.03***	***1.00***
0.10	***1.05***	***1.05***	***1.05***	***1.05***	***1.05***	***1.04***	***1.03***	***1.00***	*0.92*
0.20	***1.05***	***1.05***	***1.05***	***1.05***	***1.04***	***1.03***	***1.00***	*0.94*	0.77
0.33	***1.05***	***1.05***	***1.05***	***1.04***	***1.03***	***1.00***	***0.95***	0.84	0.57
0.50	***1.05***	***1.05***	***1.04***	***1.03***	***1.00***	***0.95***	0.86	0.66	0.31
0.67	***1.05***	***1.04***	***1.03***	***1.00***	***0.95***	0.86	0.70	0.42	0.09
0.80	***1.05***	***1.03***	***1.00***	***0.95***	0.86	0.70	0.46	0.17	<.01
0.91	***1.05***	***1.00***	*0.93*	0.81	0.63	0.38	0.14	0.01	<.01
0.99	***1.00***	0.63	0.29	0.08	<.01	<.01	<.01	<.01	<.01
1.01	*1.11*	1.77	3.84	14.04	>100	>100	>100	>100	>100
1.10	*1.06*	1.11	1.21	1.40	1.85	3.25	10.06	>100	>100
1.25	*1.06*	*1.08*	1.12	1.20	1.36	1.76	2.94	10.59	>100
1.5	*1.05*	*1.07*	*1.09*	1.14	1.23	1.43	1.95	4.20	42.28
2	*1.05*	*1.06*	*1.08*	*1.11*	1.17	1.29	1.59	2.65	12.35
3	*1.05*	*1.06*	*1.07*	*1.09*	1.14	1.23	1.43	2.10	6.67
5	*1.05*	*1.06*	*1.07*	*1.09*	1.12	1.20	1.36	1.87	4.90
10	*1.05*	*1.06*	*1.07*	*1.08*	1.11	1.18	1.32	1.76	4.13
25	*1.05*	*1.06*	*1.07*	*1.08*	*1.11*	1.17	1.30	1.70	3.79

B. RRA < 1

	Relative Risk Aversion								
EIS	**0.04**	**0.10**	**0.20**	**0.33**	**0.50**	**0.67**	**0.80**	**0.91**	**0.99**
0.04	*1.05*	*1.05*	*1.05*	*1.05*	*1.05*	*1.05*	*1.05*	*1.05*	*1.05*
0.10	*1.06*	*1.06*	*1.06*	*1.06*	*1.06*	*1.05*	*1.05*	*1.05*	*1.05*
0.20	*1.07*	*1.06*	*1.06*	*1.06*	*1.06*	*1.06*	*1.06*	*1.05*	*1.05*
0.33	*1.08*	*1.08*	*1.07*	*1.07*	*1.07*	*1.06*	*1.06*	*1.06*	*1.05*
0.50	*1.11*	*1.10*	*1.10*	*1.09*	*1.08*	*1.07*	*1.06*	*1.06*	*1.05*
0.67	1.16	1.15	1.14	1.13	*1.11*	*1.09*	*1.07*	*1.06*	*1.05*
0.80	1.28	1.27	1.24	1.21	1.17	1.13	*1.10*	*1.07*	*1.05*
0.91	1.72	1.67	1.59	1.48	1.36	1.25	1.17	*1.10*	*1.06*
0.99	>100	>100	63.74	32.16	13.68	5.82	2.94	1.68	*1.11*
1.01	<.01	<.01	0.02	0.03	0.08	0.19	0.37	0.66	***1.00***
1.10	0.61	0.63	0.67	0.72	0.79	0.87	*0.94*	***1.00***	***1.05***
1.25	0.82	0.84	0.86	0.89	*0.93*	***0.97***	***1.00***	***1.03***	***1.05***
1.5	*0.91*	*0.92*	*0.93*	***0.95***	***0.97***	***1.00***	***1.02***	***1.04***	***1.05***
2	***0.95***	***0.96***	***0.97***	***0.98***	***1.00***	***1.02***	***1.03***	***1.04***	***1.05***
3	***0.98***	***0.98***	***0.99***	***1.00***	***1.01***	***1.03***	***1.04***	***1.05***	***1.05***
5	***0.99***	***0.99***	***1.00***	***1.01***	***1.02***	***1.03***	***1.04***	***1.05***	***1.05***
10	***1.00***	***1.00***	***1.01***	***1.01***	***1.02***	***1.03***	***1.04***	***1.05***	***1.05***
25	***1.00***	***1.00***	***1.01***	***1.02***	***1.02***	***1.03***	***1.04***	***1.05***	***1.05***

3.3 Empirical Analysis

Our empirical focus is not to test the model discussed in the previous section but rather to directly address the question of whether real interest rate risk is priced. This question is actually a bit at odds with the model in that it implies a single type of interest rate risk whereas the model shows that their are two different interest rate factors with different risk prices. Ideally, we would like to separately measure consumption growth and time preference interest rate risk. Given the unobservability of time preferences and the imprecise and low-frequency nature of consumption data, measuring aggregate interest rate risk is probably the best we can do. Moreover, aggregate interest rate risk is of direct interest because interest rates are highly visible and economically important. Even though we don't directly test it, the model does inform how we think about and measure interest rate risk. Perhaps most significantly, the model predicts that investors care about shocks to both current and expected future risk free interest rates. Thus, instead of considering just $cov_t\left(r_{i,t+1}, r_{f,t+2} - E_t\left[r_{f,t+2}\right]\right)$, we focus on $\sigma_{ih} \equiv cov_t\left(r_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{w,t+1+j}\right)$.

Our empirical work faces two primary challenges. First, our focus is on real interest rates. This is the risk free rate in our model, and it is the relevant quantity for actual economic decisions. Unfortunately, real interest rates are not directly observed. We overcome this problem by modeling expected Consumer Price Index (CPI) inflation and estimating monthly real interest rates as the difference between nominal 1-month Treasury bill interest rates and expected inflation over the next month. For our baseline estimates, we focus on the 1983 to 2012 time period because monetary policy has been more consistent and inflation has been less volatile during the Greenspan and Bernanke Federal Reserve chairmanships than in previous periods.

Our second empirical challenge is that interest rate risk involves shocks to expectations. Thus, we need to estimate interest rate expectations. We do this with a vector autoregression (VAR) of interest rates, inflation, and other state variables. From the VAR, we extract an estimate for the time series of $(E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{f,t+1+j}$ innovations, which we in turn use to estimate σ_{ih} for various assets.

3.3.1 Vector Autoregression

Our VAR model is:

$$Y_t = AY_{t-1} + \omega_t \tag{3.27}$$

Y_t is a $k \times 1$ vector with the nominal 1-month treasury bill log yield and seasonally adjusted log CPI inflation over the past month as its first two elements. The remaining elements of Y_t are state variables useful for forecasting these two variables. The assumption that the VAR model has only one lag is not restrictive because lagged variables can be included in Y_t. We demean Y_t before estimating the VAR to avoid the need for a constant in equation (3.27).

We define vector ei to be the ith column of a $k \times k$ identity matrix. Using this notation we can extract expectations and shocks to current and future expectations from Y_t, A, and ω_t. Our interest is in the real risk free interest rate, which we estimate as the nominal 1-month treasury bill yield less expected inflation:

$$\widehat{r_{f,t+1}} = (e1' - e2'A)\, Y_t \tag{3.28}$$

Similarly, expected future risk free rates are:

$$E_t\left[\widehat{r_{f,t+j}}\right] = (e1' - e2'A)\, A^{j-1} Y_t \tag{3.29}$$

Shocks to current and expected risk free rates are:

$$(E_{t+1} - E_t)\, \widehat{r_{f,t+1+j}} = (e1' - e2'A)\, A^{j-1} \omega_{t+1} \tag{3.30}$$

Most importantly, total interest rate news is:

$$\begin{aligned} News_{h,t+1} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \widehat{r_{f,t+1+j}} \\ &= (e1' - e2'A) \sum_{j=1}^{\infty} \rho^j A^{j-1} \omega_{t+1} \\ &= (e1' - e2'A)\, \rho\, (I - \rho A)^{-1}\, \omega_{t+1} \end{aligned} \tag{3.31}$$

where I is the identity matrix.

All that remains is to choose state variables for Y_t and estimate equation (3.27). Following

Campbell (1996), we include the relative treasury bill rate, defined as the difference between the current one-month treasury bill yield and the average one-month treasury bill yield over the previous 12 months. Similarly, we include the relative monthly CPI inflation rate, defined the same way. Next, we include the yield spread between 10-year treasury bonds and 3-month treasury bonds because the slope of the yield curve is known to predict interest rate changes. Finally, we include the CRSP value-weighted market return and the log dividend-price ratio (defined as dividends over the past year divided by current price), which is known to predict market returns. These variables are useful to the extent that equity returns are related to expected future interest rates. We considered including additional lags of these variables by re-estimating equation (3.27) with multiple lags of Y_t. The Bayesian Information Criteria is insensitive to adding lags so we do not include lagged variables in Y_t.

Table 3.2 shows coefficient estimates and standard errors for the elements of A related to predicting nominal interest rates and inflation. Columns (1) and (2) report results for the 1983 to 2012 time period, which is our primary focus. Nominal interest rate shocks are highly persistent with lag coefficient of 0.96. Inflation shocks are much less persistent and only have a lag coefficient of 0.07. Inflation is increasing in lagged nominal yields. The VAR explains 95% of the variation in nominal yields over time. Inflation changes are less predictable with an R-squared of 0.24.

Because our main interest is in the risk free rate, we plot $\widehat{r_{f,t+1}}$ in Figure 3.1. Along with our estimated real risk free rate, we also plot the nominal one-month treasury bill yield and the Federal Reserve Bank of Cleveland's real risk free rate estimate.[11] As we would expect in a stable inflation environment, real interest rates generally follow the same pattern as nominal interest rates. Nonetheless, inflation expectations do change over time, particularly over the past few years. Our real risk free rate estimate closely tracks the Federal Reserve Bank of Cleveland's estimate, which increases our confidence in our methodology.

As a robustness check, we also estimate real risk free rates and real risk free rate news

[11]The Federal Reserve Bank of Cleveland's real risk free rate estimates are described by Haubrich et al. (2008, 2012).

Table 3.2: VAR Results

y1 is the nominal log yield on a one-month treasury bill. Inflation is one-month log inflation. Relative y1 and relative inflation are the difference between current yields and inflation and average values over the past twelve months. y120 - y3 is the yield spread between 10-year and 3-month treasury bonds. rmrf is the excess return of the CRSP value weighted market return over the risk free rate. d - p is the log dividend-price ratio, calculated for the CRSP value-weighted market index using current prices and average dividends over the past twelve months. Results are for a 1-lag VAR of demeaned y1, inflation, relative y1, relative inflation, rmrf, and d-p. Coefficients for dependent variables y1 and inflation are reported. The other dependent variables are omitted for brevity. Bootstrapped standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	1983-2012		**1927-2012**	
	(1)	**(2)**	**(3)**	**(4)**
	y1	**inflation**	**y1**	**inflation**
Lagged Variables				
y1	0.9639***	0.1939*	0.9741***	0.0631
	(0.0202)	(0.1003)	(0.0116)	(0.0773)
inflation	0.0314	0.0737	0.0102*	0.7762***
	(0.0297)	(0.1734)	(0.0062)	(0.0709)
relative	-0.0976**	0.1295	-0.1752***	0.5909***
y1	(0.0457)	(0.1585)	(0.0407)	(0.1599)
relative	-0.0136	0.3268*	-0.003	-0.4554***
inflation	(0.0281)	(0.1767)	(0.0056)	(0.0837)
y120 - y3	-0.0032	-0.002	-0.0062**	0.0014
	(0.0036)	(0.0155)	(0.0024)	(0.0122)
rmrf	0.0013*	0.0083*	0.0008**	0.0061*
	(0.0007)	(0.0042)	(0.0004)	(0.0034)
d - p	0.0001	0.0002	0.0000	-0.0002
	(0.0001)	(0.0005)	(0.0000)	(0.0003)
R-Squared	0.95	0.24	0.95	0.32

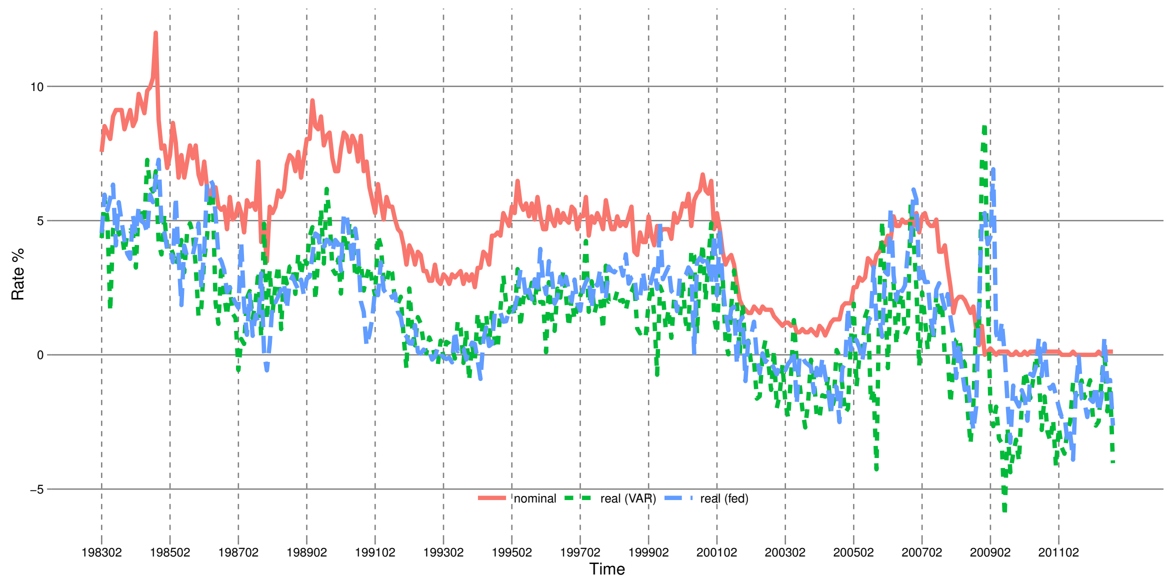


Figure 3.1: Risk Free Rates 1983-2012

The nominal risk free rate is the yield on a one-month nominal treasury bill. The real risk free rate is estimated using our VAR analysis. We also report the real risk free rate estimated by the Federal Reserve Bank of Cleveland.

over a longer time period, starting in 1927. Our methodology for the longer time period is the same as before except that we use the unadjusted CPI because the seasonally adjusted CPI is only available starting in 1947. Columns (3) and (4) of Table 3.2 report the VAR results. In the extended time sample, inflation shocks are more persistent (inflation's lagged coefficient is 0.78, compared to 0.07 before). The results are otherwise similar to the original VAR. Figure 3.2 plots nominal and estimated real interest rates from 1927 to 2012. Expected inflation varies more in the extended sample than it does after 1983. Thus, the real and nominal interest rates do not track each other as closely. Expected inflation is particularly high in the 1930's, 1940's, and 1970's, and deflation caused real interest rates to exceed nominal interest rates in the 1920's.

3.3.2 Cross-Sectional Equity Pricing

If real interest rate risk is priced and stocks vary in their exposure to real interest rate risk, real interest rate risk should be priced in the cross section of stock returns. This is not the first paper to connect time series interest rate changes with cross-sectional stock

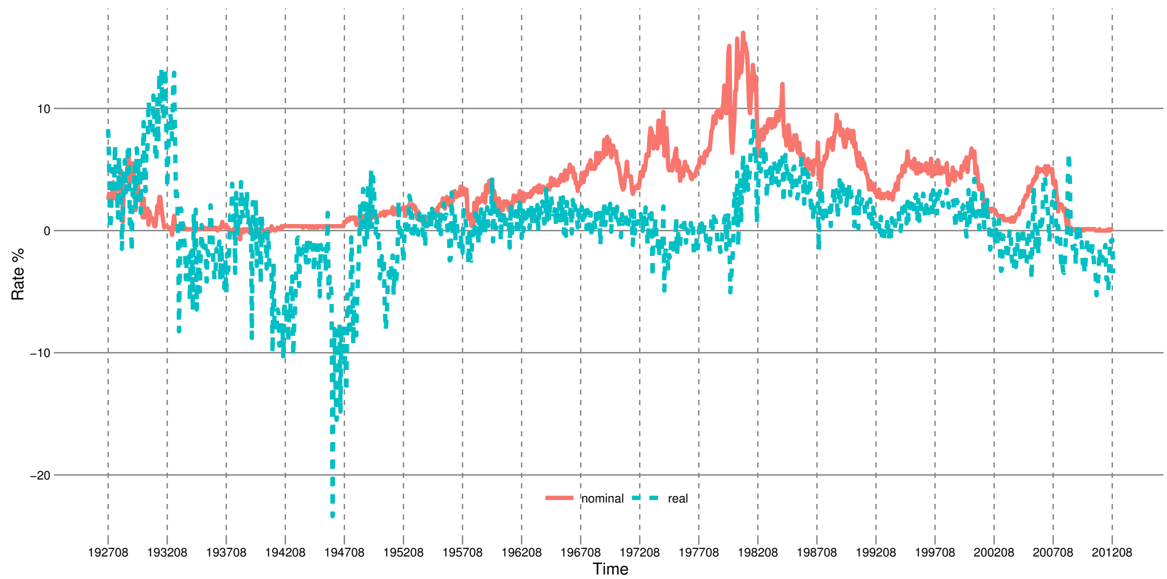


Figure 3.2: Risk Free Rates 1927-2012

The nominal risk free rate is the yield on a one-month nominal treasury bill. The real risk free rate is estimated using our VAR analysis.

returns. For example, Fama & French (1993) find comovement between excess stock returns and excess returns on long term bonds but conclude that bond factors have little impact on cross sectional stock prices. Petkova (2006) finds that innovations to term spreads and one month nominal interest rates are correlated with and partially explain size and value returns. Nieuwerburgh et al. (2012) find that high returns to value stocks relative to growth stocks are explained by covariance with shocks to nominal bond risk premia whereas returns to treasury bond portfolios of different maturities are largely explained by differential exposure to the level of interest rates. Our empirical analysis differs from previous studies because we focus specifically on stock exposure to real interest rate innovations. Moreover, we sort stocks based on this exposure instead of focusing on established size and value returns.

To test whether interest rate risk is priced we sort stocks into portfolios according to covariance with interest rate news ($News_{h,t+1}$). Specifically, we estimate $\sigma_{ih} = cov_t\left(r_{i,t+1}, News_{h,t+1}\right)$ on a rolling basis for all NYSE, AMEX, and NASDAQ common stocks using returns and VAR $News_h$ estimates over the past three years, with the requirement that included stocks must have at least two years of historical data. Value-weighted decile portfolios are formed monthly by sorting stocks according to those estimates.

Table 3.3 reports market capitalization, average excess returns, and $\beta_{ih} = \frac{\sigma_{ih}}{\sigma_h^2}$ estimates for each portfolio. The table also reports pricing errors (alphas) relative to the CAPM and Fama & French (1993) three factor model and factor loadings (betas) for the three factor model. Panel A reports results for our baseline 1985-2012 time period.[12] Risk free rate news betas increase across the portfolios, and decile 10's news beta is a significant 0.58 higher than decile 1's news beta. Monthly excess returns are 42 bps lower in the 10th decile than in the 1st decile, but this return difference is not statistically significant, and there is no clear pattern to excess returns across the decile portfolios other than a drop in returns in decile 10. CAPM and 3 Factor alphas follow the same basic pattern. Factor loadings are also similar across the portfolios. The one exception is that decile 10 has a large negative loading on the value factor (HML). The bottom line is that there is no evidence that interest rate risk is priced in the cross section of equities.

Results are similar in the extended 1929-2012 sample, reported in Panel B. Once again, average excess returns and alpha estimates decrease with interest rate news exposure, but the differences are not significant. The most striking difference between Panel A and Panel B is that β_{ih} differences across the portfolios are not significant in the extended sample. This suggests that stock-level interest rate risk was not stable over time early in the sample, undercutting our ability to form interest rate risk portfolios. This problem appears to be concentrated in the first few decades of the sample when inflation and interest rates were most volatile. In later analysis, we examine a 1952 to 2012 sample and find significant β_{ih} differences between the decile portfolios. As in the other samples, these β_{ih} differences are not accompanied by significant return differences.

[12]We form the portfolios based on at least two years of historical data, which causes the sample to start in 1985 instead of 1983.

Table 3.3: Real Risk Free Rate News Covariance Deciles

Value-weighted decile portfolios are formed at the end of each month by sorting stocks based on covariance with risk free rate news over the past three years. The table reports betas with respect to risk free rate news, average size, and average excess returns for each portfolio. The table also reports results for time series regressions of excess returns on excess market returns (the CAPM regression) and excess market returns, the Fama-French size factor (smb), and the Fama-French value factor (hml) (the 3 Factor regression). Standard errors for the 10-1 portfolio difference are reported in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. The sample is NYSE, AMEX, and NASDAQ common stocks.

A. 1985-2012

Decile	**Rf News**	**Market Cap**	**Excess**	**CAPM**	**3 Factor**	**Factor Loadings (Betas)**		
	Beta	**($B)**	**Return**	**Alpha**	**Alpha**	**rmrf**	**smb**	**hml**
1	-0.17	0.72	0.63%	-0.19%	-0.16%	1.27	0.61	-0.06
2	0.07	1.36	0.94%	0.24%	0.30%	1.10	0.22	-0.15
3	-0.04	1.94	0.87%	0.25%	0.23%	1.04	0.07	0.04
4	0.13	2.42	0.65%	0.06%	0.03%	1.00	-0.04	0.09
5	0.00	2.74	0.51%	-0.03%	-0.05%	0.94	-0.10	0.03
6	0.02	2.76	0.48%	-0.06%	-0.08%	0.93	-0.14	0.05
7	0.03	2.58	0.54%	-0.02%	-0.04%	0.97	-0.11	0.03
8	0.15	2.21	0.68%	0.06%	0.08%	1.04	-0.13	-0.07
9	0.14	1.69	0.61%	-0.06%	-0.04%	1.10	0.01	-0.06
10	0.41	0.85	0.21%	-0.62%	-0.44%	1.21	0.55	-0.47
10-1	0.58**	0.13**	-0.42%	-0.42%	-0.27%	-0.06	-0.07	-0.41***
	(0.23)	(0.06)	(0.33%)	(0.34%)	(0.34%)	(0.08)	(0.11)	(0.12)

B. 1929-2012

Decile	**Rf News**	**Market Cap**	**Excess**	**CAPM**	**3 Factor**	**Factor Loadings (Betas)**		
	Beta	**($B)**	**Return**	**Alpha**	**Alpha**	**rmrf**	**smb**	**hml**
1	-0.01	0.17	0.66%	-0.05%	-0.12%	1.15	0.52	-0.03
2	0.00	0.48	0.66%	0.04%	0.03%	1.04	0.20	-0.06
3	0.03	0.69	0.70%	0.13%	0.12%	0.99	0.08	-0.01
4	0.06	0.86	0.71%	0.15%	0.15%	0.96	0.02	0.00
5	0.01	0.98	0.60%	0.04%	0.02%	0.97	-0.03	0.06
6	0.03	1.05	0.56%	-0.01%	-0.03%	0.98	-0.03	0.09
7	0.06	1.08	0.58%	-0.01%	-0.02%	1.03	-0.08	0.08
8	0.06	1.05	0.56%	-0.07%	-0.10%	1.08	0.00	0.11
9	0.10	0.83	0.61%	-0.07%	-0.12%	1.15	0.04	0.17
10	0.11	0.38	0.58%	-0.18%	-0.27%	1.23	0.50	0.03
10-1	0.13	0.21***	-0.09%	-0.13%	-0.14%	0.07**	-0.02	0.05
	(0.09)	(0.02)	(0.18%)	(0.18%)	(0.18%)	(0.03)	(0.06)	(0.05)

3.3.3 Equity Premium

Because the market portfolio is a claim to future dividends, it may be exposed to interest rate risk. Thus, interest rate risk may affect expected equity returns and could explain part of the equity premium puzzle. The magnitude and direction of this effect depend on the market return's covariance with interest rate news and the price of interest rate risk.

AELR imply that interest rate risk explains virtually all of the equity premium. In their benchmark model, assets are priced based on covariance with consumption growth shocks and time preference shocks, which map directly into interest rate shocks. Consistent with previous studies, they estimate that equity returns are essentially uncorrelated with consumption growth. Thus, their explanation of the equity premium is almost entirely based on interest rate risk. Equities are risky because they have a long duration and are sensitive to persistent real interest rate shocks. Duration simultaneously explains the upward sloping yield curve and the equity premium. In the AELR benchmark model, equity returns are highly sensitive to interest rate shocks, with a correlation of approximately -0.94.

Using our estimates of interest rate news, we can directly measure these two moments. Panel A of Table 3.4 shows results for the 1985 to 2012 time period. Excess market returns ($rmrf$) have a correlation of 0.05 and a beta of 0.11 with respect to interest rate news. These estimates are close to zero, suggesting that equity returns have little exposure to interest rate risk. According to the point estimate, the market return is positively correlated with interest rate shocks, consistent with long run consumption growth shocks and in contrast to AELR's time preference shocks.

Table 3.4 also reports interest rate correlations and betas for the long-short decile 10 minus decile 1 interest rate risk portfolio and for 1 to 2 year and 5 to 10 year bonds.[13] By construction, the long-short interest rate risk portfolio has a positive beta. The bond portfolios have negative exposures to interest rate news. However, these exposures are small. Interest rate betas are -0.04 for both portfolios, and the beta is only significantly different from zero for the short-term bonds.

[13]Bond return data is from CRSP.

The final rows of Table 3.4 report average excess returns and average excess returns divided by interest rate news beta. If interest rate news is the primary risk factor investors care about, this ratio (the implied price of beta) should be consistent across assets. The point estimates clearly differ. In particular, the bond returns and cross-sectional interest rate risk portfolio imply a negative price of interest rate risk whereas market returns imply a positive price. Unfortunately, betas and average returns are measured too imprecisely to definitively rule out consistent interest rate risk pricing across the assets. Panel B of Table 3.4 reports the same statistics for a longer sample period, starting in 1952 when CRSP bond return data starts. The basic results are all the same.

Our findings suggest that interest rate risk is unlikely to explain the equity premium. Certainly, there is no evidence in favor of the hypothesis that equities face significant interest rate risk. How can this be reconciled with AELR's empirical findings? The main difference between our analysis and AELR's is that AELR do not estimate real interest rate innovations. Their GMM includes the unconditional correlation between equity returns and the real risk free rate at an annual frequency but omits the more important correlation of interest rate *news* with *excess* equity returns. Our analysis estimates this moment and finds that it is essentially zero.

3.4 Conclusion

Is real interest rate risk priced? Theoretically, it could be priced in either direction. Empirically, there is little evidence that real interest rate risk is priced at all.

Our interest rate risk model has two theoretical implications. First, it matters where interest rate shocks comes from. Interest rate increases stemming from news about future consumption growth are generally good news to investors whereas interest rate increases stemming from time preference shocks are generally bad news. Thus, long-run consumption risk logic implies that long-duration assets are relatively safe whereas time preference risk logic implies that long-duration assets are relatively risky. A more general lesson is the importance of thinking in general equilibrium terms. Because interest rates are endogenous,

Table 3.4: Equity Market and Bond Real Interest Rate Risk

rmrf is the excess return on the CRSP value weighted market portfolio. Decile 10-1 is returns to long-short portfolio representing the difference between the 10th and first riskfree rate news covariance portfolios, described in Table 2. 1-2 and 5-10 year bonds represent excess returns to treasury bonds of those durations, as calculated by CRSP. Correlations and betas with respect to riskfree rate news and average returns are reported for each return series. The price of beta is defined as average returns divided by beta. Standard errors are reported in parentheses. Standard errors for the price of beta are calculated using the delta method. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

A. 1985-2012

	rmrf	**Decile 10-1**	**1-2 Year Bonds**	**5-10 Year Bonds**
Rf News	0.04	0.14**	-0.14***	-0.03
Correlation	(0.05)	(0.05)	(0.05)	(0.05)
Rf News	0.11	0.58**	-0.04***	-0.04
Beta	(0.17)	(0.23)	(0.02)	(0.06)
Average	0.60%**	-0.42%	0.12%***	0.34%***
Excess Returns	(0.25%)	(0.33%)	(0.02%)	(0.09%)
Price of	5.35%	-0.72%**	-3.14%***	-9.70%
Beta	(10.57%)	(0.30%)	(0.64%)	(13.81%)

B. 1952-2012

	rmrf	**Decile 10-1**	**1-2 Year Bonds**	**5-10 Year Bonds**
Rf News	0.05	0.12***	-0.40***	-0.12***
Correlation	(0.04)	(0.04)	(0.03)	(0.04)
Rf News	0.10	0.30***	-0.12***	-0.10***
Beta	(0.08)	(0.09)	(0.01)	(0.03)
Average	0.55%***	-0.16%	0.09%***	0.16%***
Excess Returns	(0.16%)	(0.19%)	(0.02%)	(0.06%)
Price of	5.43%	-0.54%	-0.72%***	-1.57%***
Beta	(5.91%)	(0.46%)	-(0.12%)	-(0.13%)

interest rate risk is not a meaningful concept without specifying what is driving interest rate shocks.

The second theoretical implication of our model is that AELR preferences with ψ close to 1 and significantly different from $1/\gamma$ imply implausible aversion to future time preference shocks.

Empirically, stocks sorted on interest rate risk have only small, statistically insignificant return differences. Moreover, the market return and treasury bond returns have low covariance with interest rate news. Thus, interest rate risk is unlikely to explain much of equity or bond return premia even if it is priced to some extent in the cross section. Overall, our results suggest that interest rate risk is not a major concern to investors.

References

Albuquerque, R., Eichenbaum, M., Luo, V. & Rebelo, S. (2016), 'Valuation risk and asset pricing', *Journal of Finance* .

Ang, A., Bali, T. & Cakici, N. (2012), 'The joint cross section of stocks and options', *Working Paper* .

Ang, A., Hodrick, R. J., Xing, Y. & Zhang, X. (2006), 'The cross-section of volatility and expected returns', *The Journal of Finance* **61**(1), 259–299.

Avramov, D., Chordia, T. & Goyal, A. (2006), 'Liquidity and autocorrelations in individual stock returns', *The Journal of Finance* **61**(5), 2365–2394.

Bakshi, G., Kapadia, N. & Madan, D. (2003), 'Stock return characteristics, skew laws, and the differential pricing of individual equity options', *The Review of Financial Studies* **16**(1), pp. 101–143.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–1509.

Barber, B. M. & Odean, T. (2008), 'All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors', *Review of Financial Studies* **21**(2), 785–818.

Bollen, N. P. B. & Whaley, R. E. (2004), 'Does net buying pressure affect the shape of implied volatility functions?', *The Journal of Finance* **59**(2), pp. 711–753.

Brunnermeier, M. K., Nagel, S. & Pedersen, L. H. (2008), 'Carry trades and currency crashes', *Working Paper* .

Campbell, J. & Vuolteenaho, T. (2004), 'Bad beta, good beta', *American Economic Review* **94**(5), 1249–1275.

Campbell, J. Y. (1993), 'Intertemporal asset pricing without consumption data', *The American Economic Review* pp. 487–512.

Campbell, J. Y. (1996), 'Understanding risk and return', *Journal of Political Economy* pp. 298–345.

Campbell, J. Y. & Viceira, L. M. (2002), *Strategic asset allocation: portfolio choice for long-term investors*, Oxford University Press, USA.

Chan, L. K., Chen, H.-L. & Lakonishok, J. (2002), 'On mutual fund investment styles', *Review of Financial Studies* **15**(5), 1407–1437.

Chang, B. Y., Christoffersen, P. & Jacobs, K. (2013), 'Market skewness risk and the cross section of stock returns', *Journal of Financial Economics* **107**, 46 – 68.

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

Chevalier, J. & Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of Political Economy* **105**(6), 1167–1200.

Cochrane, J. H. (2011), 'Presidential address: Discount rates', *The Journal of Finance* **66**(4), 1047–1108.

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

Cohen, L. & Lou, D. (2012), 'Complicated firms', *Journal of Financial Economics* **104**(2), 383–400.

Conrad, J., Dittmar, R. F. & Ghysels, E. (2013), 'Ex ante skewness and expected stock returns', *The Journal of Finance* **68**(1), 85–124.

Daniel, K. & Titman, S. (1997), 'Evidence on the characteristics of cross sectional variation in stock returns', *The Journal of Finance* **52**(1), 1–33.

Duffie, D. (2010), 'Presidential address: Asset price dynamics with slow-moving capital', *The Journal of Finance* **65**(4), 1237–1267.

Epstein, L. G., Farhi, E. & Strzalecki, T. (2014), 'How much would you pay to resolve long-run risk?', *The American Economic Review* **104**(9), 2680–2697.

Epstein, L. G. & Zin, S. E. (1991), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis', *Journal of political Economy* pp. 263–286.

Fama, E. F. (1970), 'Multiperiod consumption-investment decisions', *The American Economic Review* **60**(1), 163–174.

Fama, E. F. & French, K. R. (1987), 'Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage', *Journal of Business* pp. 55–73.

Fama, E. F. & French, K. R. (1992), 'The cross-section of expected stock returns', *The Journal of Finance* **47**(2), 427–465.

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

Fama, E. F. & French, K. R. (1996), 'Multifactor explanations of asset pricing anomalies', *The Journal of Finance* **51**(1), 55–84.

Fama, E. F. & French, K. R. (1998), 'Value versus growth: The international evidence', *The Journal of Finance* **53**(6), 1975–1999.

Guvenen, F. (2006), 'Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective', *Journal of Monetary Economics* **53**(7), 1451–1472.

Hall, R. E. (1988), 'Intertemporal substitution in consumption', *Journal of Political Economy* **96**(2), 339–357.

Haubrich, J. G., Pennacchi, G. & Ritchken, P. H. (2008), 'Estimating real and nominal term structures using treasury yields, inflation, inflation forecasts, and inflation swap rates', *Working Paper* .

Haubrich, J., Pennacchi, G. & Ritchken, P. (2012), 'Inflation expectations, real rates, and risk premia: evidence from inflation swaps', *Review of Financial Studies* **25**(5), 1588–1629.

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

Hoerl, A. E. & Kennard, R. W. (1970), 'Ridge regression: Biased estimation for nonorthogonal problems', *Technometrics* **12**(1), 55–67.

Hong, H., Lim, T. & Stein, J. C. (2000), 'Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies', *The Journal of Finance* **55**(1), 265–295.

Hong, H. & Stein, J. C. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *The Journal of Finance* **54**(6), 2143–2184.

Hou, K. (2007), 'Industry information diffusion and the lead-lag effect in stock returns', *Review of Financial Studies* **20**(4), 1113–1138.

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

Kahneman, D. (1973), *Attention and Effort*, Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Koijen, R. S. J., Moskowitz, T. J., Pedersen, L. H. & Vrugt, E. B. (2013), 'Carry', *Working Paper* .

Maurer, T. (2012), 'Is consumption growth merely a sideshow in asset pricing?', *Working Paper* .

Merton, R. C. (1973), 'An intertemporal capital asset pricing model', *Econometrica: Journal of the Econometric Society* pp. 867–887.

Moskowitz, T. J. & Grinblatt, M. (1999), 'Do industries explain momentum?', *The Journal of Finance* **54**(4), 1249–1290.

Nieuwerburgh, S. V., Lustig, H. N. & Koijen, R. S. (2012), 'The cross-section and time-series of stock and bond returns', *Working Paper* .

Normandin, M., St-Amour, P. et al. (1998), ‘Substitution, risk aversion, taste shocks and equity premia’, *Journal of Applied Econometrics* **13**(3), 265–281.

Pan, J. (2002), ‘The jump-risk premia implicit in options: Evidence from an integrated time-series study’, *Journal of Financial Economics* **63**(1), 3–50.

Park, J. & Ratti, R. A. (2008), ‘Oil price shocks and stock markets in the u.s. and 13 european countries’, *Energy Economics* **30**, 2587 – 2608.

Petkova, R. (2006), ‘Do the fama–french factors proxy for innovations in predictive variables?’, *The Journal of Finance* **61**(2), 581–612.

Petkova, R. & Zhang, L. (2005), ‘Is value riskier than growth?’, *Journal of Financial Economics* **78**(1), 187–202.

Pindyck, R. S. (2001), ‘The dynamics of commodity spot and futures markets: A primer’, *Working Paper* .

Pollet, J. M. (2005), ‘Predicting asset returns with expected oil price changes’, *Working Paper* .

Porta, R. L., Lakonishok, J., Shleifer, A. & Vishny, R. (1997), ‘Good news for value stocks: Further evidence on market efficiency’, *The Journal of Finance* **52**(2), 859–874.

Rizova, S. (2010), ‘Predictable trade flows and returns of trade-linked countries’, *Working Paper* .

Schwartz, E. S. (1997), ‘The stochastic behavior of commodity prices: Implications for valuation and hedging’, *The Journal of Finance* **52**(3), 923–973.

Sharpe, W. F. (1994), ‘The sharpe ratio’, *Journal of Portfolio Management* .

Sims, C. A. (2003), ‘Implications of rational inattention’, *Journal of Monetary Economics* **50**, 665–690.

Skiadas, C. (2009), *Asset pricing theory*, Princeton University Press.

Smalling, D. (2012), ‘Forgotten portfolios’, *Working Paper* .

Tibshirani, R. (1996), ‘Regression shrinkage and selection via the lasso’, *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 267–288.

Weil, P. (1989), ‘The equity premium puzzle and the risk-free rate puzzle’, *Journal of Monetary Economics* **24**(3), 401–421.

Xing, Y., Zhang, X. & Zhao, R. (2010), ‘What does the individual option volatility smirk tell us about future equity returns?’, *Journal of Financial and Quantitative Analysis* **45**, 641–662.

Yang, F. (2013), ‘Investment shocks and the commodity basis spread’, *Journal of Financial Economics* **110**(1), 164 – 184.

Zou, H. & Hastie, T. (2005), ‘Regularization and variable selection via the elastic net’, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **67**(2), pp. 301–320.

Appendix A

Appendix To Commodity Inattention

There are several steps to parsing company names from CRSP in order to convert them to viable Factiva searches; I detail the process I followed in this section. First, any companies who's names are acronyms, for example the Chicago Board of Exchange (CBOE) and Home Box Office (HBO) are often listed in CRSP as "C B O E" and "H B O" while being cited as "CBOE" and "HBO" in news stories so I remove spaces in such instances. Second, companies of the form ABC.com and XYZ.com are represented in CRSP as "ABC COM" and "XYZ COM". I reinsert the "." between the company name and "COM". Third, CRSP abbreviates some words in company names: for example rather than ABC Holdings they may write "ABC HLDGS" or "ABC HLDS" or "ABC HLDNGS" (there are many more variations). I standardize all of these abbreviations to their full word. Fourth, CRSP includes state names at the end of company names occasionally but these are generally not listed in news articles so I remove these from company names (ex: CA, NJ, FL, etc.) if a state abbreviation appears as the last word of a company name. Finally, CRSP includes abbreviations for corporate entity identification like "CO", "CORP", "LLC" at the end of company names that are not present in news articles so I remove these identifications if the company name is more than one word (in the event it is only one word I leave it in because some companies are not identifiable with such a short name).

The commodity sector keywords associated with each commodity sector are as follows: Energy keywords are "Brent", "Crude Oil", "Gasoil", "Heating Oil", "Oil", "Natural Gas", "Gasoline", "Gas", "WTI", "West Texas Intermediate"; Ag keywords are "Cocoa", "Coffee", "Corn", "Cotton", "Kansas Wheat", "Soybeans", "Sugar", "Wheat"; Metal keywords are "Aluminum", "Copper", "Gold", "Lead", "Nickel", "Silver", "Zinc". The search string for news articles without commodity names is

> CompanyName and date from StartDate to EndDate and (rst=FTFT or rst=J or rst=NYTF)

while the search string with commodity names is

> CompanyName and (Keyword1 or Keyword2 or Keyword3 or ...) date from StartDate to EndDate and (rst=FTFT or rst=J or rst=NYTF)

Appendix B

Appendix To Market Crash Risk and Slow Moving Capital

B.1 Robustness

We perform several robustness checks. First we verify that our results are not solely driven by the 2008 financial crisis. We split the time period in half (by years): pre and post 2004 and test equation our forecasting equation in US, JPY, EUR. Results are robust in both sub-samples. It is difficult to split it up much more finely than this due to the low number of observations already in the sample (constraint is the options data). Table B.1 reports the results.

The second set of robustness checks we run is to verify that the particular volatility maturity that we use (1 year) is not the only source of the results. To do this we run predictive regressions using other volatility maturities (from 1 month to 1 year): our results are highly robust to this. Table B.2 reports the results.

B.2 Extracting Risk Neutral Moments

We follow a similar data cleaning methodology to Chang et al. (2013) to remove options with potentially erroneous quotes. In particular each day we remove options with prices of

Table B.1: Robustness to Sub-Samples

We show that our results are robust to sub-samples by splitting the time-series in two. A finer split is difficult due to the lack of S&P 500 options data going back further in time.

(a) 1996-2004

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.793**			**3.799**	6.80%	95
	[-2.872]			**[3.162]**		
US	**-0.663**	**-0.473**		**3.487**	36.48%	95
	[-2.982]	**[-5.790]**		**[4.020]**		
US	**-0.685**	**-0.535**	0.014	**3.571**	35.82%	95
	[-2.996]	**[-2.010]**	[0.276]	**[3.788]**		
JPY	-0.399			**2.234**	1.50%	95
	[-1.187]			**[1.739]**		
JPY	-0.327	**-0.264**		**2.060**	12.72%	95
	[-1.097]	**[-3.280]**		**[1.867]**		
JPY	-0.383	**-0.423**	0.036	**2.280**	12.16%	95
	[-1.408]	**[-1.964]**	[0.651]	**[2.268]**		
EUR	**-0.650**			**3.777**	9.71%	95
	[-2.468]			**[3.688]**		
EUR	**-0.600**	**-0.182**		**3.657**	18.02%	95
	[-2.383]	**[-2.251]**		**[4.002]**		
EUR	**-0.643**	-0.305	0.028	**3.826**	17.50%	95
	[-2.330]	[-1.240]	[0.539]	**[3.705]**		

(b) 2004-2012

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.360**			**1.918**	6.34%	106
	[-2.712]			**[2.783]**		
US	**-0.412**	**0.189**		**2.095**	18.60%	106
	[-2.948]	**[2.190]**		**[2.808]**		
US	**-0.394**	0.003	0.033	**1.989**	20.35%	106
	[-3.230]	[0.029]	[1.368]	**[3.159]**		
JPY	**-0.218**			**1.470**	2.34%	106
	[-2.351]			**[2.843]**		
JPY	**-0.193**	-0.090		**1.386**	5.06%	106
	[-2.213]	[-1.851]		**[2.837]**		
JPY	**-0.190**	**-0.124**	0.006	**1.367**	4.24%	106
	[-2.227]	**[-2.062]**	[0.725]	**[2.841]**		
EUR	-0.136			**0.700**	0.43%	106
	[-1.654]			**[1.960]**		
EUR	**-0.204**	**0.245**		**0.930**	28.91%	106
	[-2.367]	**[5.274]**		**[2.722]**		
EUR	-0.193	0.135	0.020	**0.867**	29.40%	106
	[-1.948]	[1.829]	[1.255]	**[2.012]**		

Table B.2: Robustness: Other S&P 500 Implied Volatility Maturities

To demonstrate that our results are robust to using other maturity points on the S&P 500 volatility surface, we run two regressions: $R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s \varepsilon_{skew,t} + \epsilon_t$ where $\varepsilon_{skew,t}$ is the innovation in $SKEW_t$ from an AR(1) model (this validates the contemporaneous relationship). The other regression is the forecasting regression of $R_{hml,t} = a + \beta_s SKEW_{t-1} + \epsilon_t$ to show that one can forecast HML with other maturities as well.

Vol Tenure	$R_{m,t}$	$\varepsilon_{skew,t}$	$SKEW_{t-1}$	(Intercept)	R^2	N
1 month	**-0.246**	**-0.422**		0.398	10.56%	200
	[-1.978]	**[-3.496]**		[1.375]		
			-0.252	**1.306**	2.34%	200
			[-2.521]	**[3.036]**		
3 months	-0.274	**-0.623**		0.416	10.57%	200
	[-1.928]	**[-3.349]**		[1.430]		
			-0.442	**2.266**	5.65%	200
			[-3.924]	**[4.349]**		
6 months	-0.264	**-0.717**		0.413	9.85%	200
	[-1.807]	**[-3.026]**		[1.398]		
			-0.544	**2.834**	7.39%	200
			[-4.188]	**[4.507]**		
9 months	-0.276	**-0.876**		0.422	10.72%	200
	[-1.905]	**[-3.230]**		[1.423]		
			-0.539	**2.803**	6.69%	200
			[-3.934]	**[4.326]**		
12 months	-0.265	**-0.817**		0.417	9.43%	200
	[-1.849]	**[-3.001]**		[1.385]		
			-0.552	**2.849**	6.86%	200
			[-3.922]	**[4.347]**		

less than \$.375, options that are in the money, days where there are less than 10 options quoted, options that have less than 10 days to maturity and options that violate arbitrage conditions. After this filtering, we extract implied volatilities from the remaining options. On each day, for each maturity we generate a volatility surface by interpolating the implied volatilities using a penalized cubic regression spline with a grid of 1000 evenly spaced strike to spot ratios (the strike to spot ratio is the ratio of the option strike to the index spot price) in $[.0001, 3]$. For values outside of the range available in the market we simply set the implied volatility to the nearest market available quote. This is to prevent the cubic spline from generating extreme (implausible) volatilities that would depend on the slope of the fit near potentially noisy tail options. Once we have a daily implied volatility surface, we compute the value of risk neutral skewness as follows. As in Bakshi et al. (2003), define:

$$
\begin{aligned}
V(t,\tau) &\equiv \int_{S_t}^{\infty} \frac{2(1-\ln\frac{K}{S_t})}{K^2} C(t,\tau;K)dK \\
&\quad + \int_0^{S_t} \frac{2(1+\ln\frac{S_t}{K})}{K^2} P(t,\tau;K)dK && \text{(B.1)}\\
W(t,\tau) &\equiv \int_{S_t}^{\infty} \frac{6\ln\frac{K}{S_t} - 3(\ln\frac{K}{S_t})^2}{K^2} C(t,\tau;K)dK \\
&\quad - \int_0^{S_t} \frac{6\ln\frac{S_t}{K} + 3(\ln\frac{S_t}{K})^2}{K^2} P(t,\tau;K)dK && \text{(B.2)}\\
\mu(t,\tau) &\equiv e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau) && \text{(B.3)}\\
RNSKEW(t,\tau) &= \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{(e^{r\tau}V(t,\tau) - \mu(t,\tau)^2)^{(3/2)}} && \text{(B.4)}
\end{aligned}
$$

where $C(t,\tau;K)$ is the price of a call at time t with strike K and τ years until maturity, $P(t,\tau;K)$ is the price of a put at time t with strike K and τ years until maturity, and r is the risk free rate. $RNSKEW(t,\tau)$ is the value of risk neutral skewness as used by Chang et al. (2013). To be consistent in maturity with the VIX and previous studies of volatility such as Ang et al. (2006) we use the value of $\tau = 30/365$ and the VIX as a proxy for volatility. To build the constant maturity measures we interpolate using linear regression between measures computed using available maturities. The measures are then annualized (as the VIX is reported in annualized units).

Appendix C

Appendix To Is Real Interest Rate Risk Priced? Theory and Empirical Evidence

C.1 Setup and General Pricing Equations

C.1.1 $\psi \neq 1$

The representative agent has the augmented Epstein-Zin preferences described by equation (3.1):

$$U_t = \max_{C_t} \left[\lambda_t C_t^{1-1/\psi} + \delta \left(U_{t+1}^{*} \right)^{1-1/\psi} \right]^{1/(1-1/\psi)}$$

where $U_{t+1}^{*} = \left\{ E_t \left[U_{t+1}^{1-\gamma} \right] \right\}^{1/(1-\gamma)}$ is the certainty equivalent of future utility. Optimization is subject to budget constraint:

$$W_{t+1} = R_{w,t+1} \left(W_t - C_t \right) \tag{C.1}$$

where W_t is wealth at time t and $R_{w,t+1}$ is the return on the overall wealth portfolio, which is a claim to all future consumption.

AELR use standard techniques from the Epstein-Zin preference literature to show that the preferences represented by equation (3.1) imply the log stochastic discount factor (sdf):

$$m_{t+1} = \theta \log\left(\delta \frac{\lambda_{t+1}}{\lambda_t}\right) - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)\, r_{w,t+1}$$

This sdf should not be surprising. It is just the standard Epstein-Zin sdf with time-varying time discounting (i.e., $\delta \frac{\lambda_{t+1}}{\lambda_t}$ instead of δ).

Using $0 = E_t\left[m_{t+1} + r_{i,t+1}\right] + \frac{1}{2}\left(\sigma_m^2 + \sigma_i^2 + 2\sigma_{mi}\right)$ (the log version of $1 = E_t\left[M_{t+1}R_{i,t+1}\right]$), we calculate the expected return for any asset as:

$$\begin{aligned} E_t\left[r_{i,t+1}\right] + \frac{1}{2}\sigma_i^2 \quad = \quad & -\theta \log\left(\delta \frac{\lambda_{t+1}}{\lambda_t}\right) + \frac{\theta}{\psi} E_t\left[\Delta c_{t+1}\right] + (1-\theta)\, E_t\left[r_{w,t+1}\right] \\ & -\frac{1}{2}\left(\frac{\theta}{\psi}\right)^2 \sigma_c^2 - \frac{1}{2}(1-\theta)^2 \sigma_w^2 + \frac{\theta}{\psi}(\theta - 1)\,\sigma_{wc} \\ & +\frac{\theta}{\psi}\sigma_{ic} + (1-\theta)\,\sigma_{iw} \end{aligned} \tag{C.2}$$

The $\frac{1}{2}\sigma_i^2$ on the left hand side of equation (C.2) is a Jensen's inequality correction for log returns.

The risk free rate is of particular interest:

$$\begin{aligned} r_{f,t+1} \quad = \quad & -\theta \log\left(\delta \frac{\lambda_{t+1}}{\lambda_t}\right) + \frac{\theta}{\psi} E_t\left[\Delta c_{t+1}\right] + (1-\theta)\, E_t\left[r_{w,t+1}\right] \\ & -\frac{1}{2}\left(\frac{\theta}{\psi}\right)^2 \sigma_c^2 - \frac{1}{2}(1-\theta)^2 \sigma_w^2 + \frac{\theta}{\psi}(\theta - 1)\,\sigma_{wc} \end{aligned} \tag{C.3}$$

Differencing equations (C.2) and (C.3) yields the risk premia of equation (3.7):

$$E_t\left[r_{i,t+1}\right] - r_{f,t+1} + \frac{1}{2}\sigma_i^2 = \frac{\theta}{\psi}\sigma_{ic} + (1-\theta)\,\sigma_{iw}$$

which is exactly the same expression as in standard Epstein-Zin models. Substituting $E_t\left[r_{w,t+1}\right]$ into equation (C.3), yields equation (3.6):

$$r_{f,t+1} = -\log\left(\delta \frac{\lambda_{t+1}}{\lambda_t}\right) + \frac{1}{\psi} E_t\left[\Delta c_{t+1}\right] - \frac{1-\theta}{2}\sigma_w^2 - \frac{\theta}{2\psi^2}\sigma_c^2$$

which is the same as standard Epstein-Zin models except that δ is replaced by $\delta \frac{\lambda_{t+1}}{\lambda_t}$.

C.1.2 $\psi = 1$

The limit of the value function as $\psi \to 1$ under AELR preferences does not exist. In the case of vanilla EZ preferences, one can find the limit of the value function by using L'Hopital's rule:

$$V_t = \left[(1-\delta)C_t^{1-1/\psi} + \delta\left(V_{t+1}^*\right)^{1-1/\psi}\right]^{1/(1-1/\psi)}$$

$$\begin{aligned}
lnV_t &= \frac{1}{1-\frac{1}{\psi}} ln\left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(V_{t+1}^*\right)^{1-\frac{1}{\psi}}\right] \\
\lim_{\psi\to 1} lnV_t &= \lim_{\psi\to 1} \frac{\frac{1}{(1-\delta)C_t^{1-\frac{1}{\psi}}+\delta\left(V_{t+1}^*\right)^{1-\frac{1}{\psi}}}\left(\frac{1}{\psi^2}(1-\delta)C_t^{1-\frac{1}{\psi}} lnC_t + \frac{1}{\psi^2}\delta\left(V_{t+1}^*\right)^{1-\frac{1}{\psi}} ln\left(V_{t+1}^*\right)\right)}{\frac{1}{\psi^2}} \\
&= (1-\delta)lnC_t + \delta ln\left(V_{t+1}^*\right) \\
V_t &= C_t^{(1-\delta)}\left(V_{t+1}^*\right)^{\delta}
\end{aligned}$$

However, this procedure cannot be performed with the AELR specification because the limit diverges. Specifically

$$\begin{aligned}
\lim_{\psi\to 1} lnU_t &= \lim_{\psi\to 1} \frac{ln\left[\lambda_t C_t^{1-1/\psi} + \delta\left(U_{t+1}^*\right)^{1-1/\psi}\right]}{1-1/\psi} \\
&= \frac{ln(\lambda_t+\delta)}{0} \to \infty
\end{aligned}$$

We also examine alternative preferences that use a generalized form of consumption in EZ preferences that are defined for all ψ. Specifying the value function as

$$V_t = \left[(1-\delta)H_t(C_t)^{1-\frac{1}{\psi}} + \delta\left(E_t V_{t+1}^{1-\gamma}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

Defining $H_t(C_t) = \Lambda_t^* C_t$ implies

$$V_t = \left(\Lambda_t^*\right)^{1-\delta} C_t^{1-\delta}\left(E_t V_{t+1}^{1-\gamma}\right)^{\frac{\delta}{1-\gamma}}$$

$$M_{t+1} = \delta\frac{V_{t+1}^{1-\gamma}}{\left(E_t V_{t+1}^{1-\gamma}\right)}\left(\frac{\Lambda_t^* C_t}{\Lambda_{t+1}^* C_{t+1}}\right)\left(\frac{\Lambda_{t+1}^*}{\Lambda_t^*}\right)$$

$$= \quad \delta \frac{V_{t+1}^{1-\gamma}}{\left(E_t V_{t+1}^{1-\gamma}\right)} \left(\frac{C_{t+1}}{C_t}\right)^{-1} \tag{C.4}$$

which is the usual SDF specified in terms of the value function. To substitute out the value function, one can assume a process for log taste shock growth and log consumption growth and then guess-verify the value function. Assume:

$$\Delta c_{t+1} \quad = \quad \mu_c + \varepsilon_{t+1}^c \tag{C.5}$$

$$\Delta \lambda_{t+1}^* \quad = \quad \mu_{\lambda^*} + \varepsilon_{t+1}^{\lambda^*} \tag{C.6}$$

with $\varepsilon_t^i \sim N(0, \sigma_i^2)$, $\varepsilon_t^c \perp \varepsilon_t^{\lambda^*}$, $\Delta c_{t+1} \equiv \log(\frac{C_{t+1}}{C_t})$, $\Delta \lambda_{t+1}^* \equiv \log(\frac{\Lambda_{t+1}^*}{\Lambda_t^*})$, and that Λ_{t+1}^* is known at time t. The log SDF is

$$m_{t+1} = \log \delta + (1-\gamma) v_{t+1} - \log\left(E_t V_{t+1}^{1-\gamma}\right) - \Delta c_{t+1} \tag{C.7}$$

Guess that the log value function is

$$v_t = A_0 + A_1 c_t + A_2 \lambda_t^* + A_3 \lambda_{t+1}^* \tag{C.8}$$

Then

$$\begin{aligned} \log\left(E_t V_{t+1}^{1-\gamma}\right) \quad &= \quad \log E_t \left[\exp\left\{(1-\gamma)(A_0 + A_1 c_{t+1} + A_2 \lambda_{t+1}^* + A_3 \lambda_{t+2}^*\right\}\right] \\ &= \quad \log \left[\exp\left\{(1-\gamma)(A_0 + A_1 c_t + (A_2 + A_3)\lambda_{t+1}^*)\right\} \times \right. \\ &\times \quad \left. E_t \left(\exp\left\{(1-\gamma)(A_1 \Delta c_{t+1} + A_3 \Delta \lambda_{t+2}^*)\right\}\right)\right] \\ &= \quad (1-\gamma)(A_0 + (A_2 + A_3)\lambda_{t+1}^* + A_1 c_t + A_1 \mu_c + A_3 \mu_{\lambda^*}) + \\ &+ \quad \frac{(1-\gamma)^2}{2} \left[A_1^2 \sigma_c^2 + A_3^2 \sigma_{\lambda^*}^2\right] \end{aligned} \tag{C.9}$$

Using the usual method of undetermined coefficients:

$$\begin{aligned} A_0 + A_1 c_t + A_2 \lambda_t^* + A_3 \lambda_{t+1}^* \quad &= \quad (1-\delta)\lambda_t^* + (1-\delta) c_t + \\ &+ \quad \frac{\delta}{1-\gamma} \left[(1-\gamma)(A_0 + (A_2 + A_3)\lambda_{t+1}^* + A_1 c_t + \right. \\ &+ \quad \left. A_1 \mu_c + A_3 \mu_{\lambda^*})\right] + \frac{\delta}{1-\gamma} \left[\frac{(1-\gamma)^2}{2} \left[A_1^2 \sigma_c^2 + A_3^2 \sigma_{\lambda^*}^2\right]\right] \end{aligned}$$

$$\begin{aligned}
A_1 &= 1 && \text{(C.10)}\\
A_2 &= (1-\delta) && \text{(C.11)}\\
A_3 &= \delta && \text{(C.12)}\\
A_0 &= \frac{\delta}{1-\delta}(\mu_c + \mu_{\lambda^*} + \frac{(1-\gamma)}{2}(\sigma_c^2 + \delta^2\sigma_{\lambda^*}^2)) && \text{(C.13)}
\end{aligned}$$

Substituting these results into the log SDF yields

$$\begin{aligned}
m_{t+1} &= \log\delta + (1-\gamma)(A_0 + c_{t+1} + \lambda^*_{t+1} + \delta\Delta\lambda^*_{t+2}) \\
&\quad - \left[(1-\gamma)(A_0 + \lambda^*_{t+1} + c_t + \mu_c + \delta\mu_{\lambda^*}) + \frac{(1-\gamma)^2}{2}\left[\sigma_c^2 + \delta^2\sigma_{\lambda^*}^2\right]\right] - \Delta c_{t+1} && \text{(C.14)}\\
&= \log\delta - \gamma\Delta c_{t+1} + \delta(1-\gamma)\Delta\lambda^*_{t+2} - (1-\gamma)(\mu_c + \delta\mu_{\lambda^*}) - \frac{(1-\gamma)^2}{2}(\sigma_c^2 + \delta^2\sigma_{\lambda^*}^2) && \text{(C.15)}
\end{aligned}$$

C.2 Substituting out Consumption (ICAPM)

Following Campbell (1993) we log linearize the budget constraint to yield equation (3.8):

$$r_{w,t+1} - E_t\left[r_{w,t+1}\right] = (E_{t+1} - E_t)\sum_{j=0}^{\infty}\rho^j\Delta c_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{w,t+1+j}$$

where $\rho = 1 - \exp\left(\overline{c-w}\right)$ is a log-linearization constant ($\overline{c-w}$ is the average log consumption-wealth ratio). Rearranging, we can express current consumption shocks as:

$$\begin{aligned}
\Delta c_{t+1} - E_t\left[\Delta c_{t+1}\right] &= r_{w,t+1} - E_t\left[r_{w,t+1}\right] \\
&\quad + (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{w,t+1+j} \\
&\quad - (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j\Delta c_{t+1+j} && \text{(C.16)}
\end{aligned}$$

So far, we have only made use of modified Epstein-Zin preferences and the budget constraint. We now use assumptions about consumption and time preference innovations for the first time. Due to our homoscedasticity assumption, risk premia do not change over time, and the risk free rate only changes in response to time preference and consumption growth innovations.

Thus, innovations to expected returns can be decomposed as:

$$\begin{aligned}(E_{t+1}-E_t)\, r_{w,t+1+j} &= (E_{t+1}-E_t)\, r_{f,t+1+j} \\ &= (E_{t+1}-E_t)\log\left(\frac{\lambda_{t+j}}{\lambda_{t+j+1}}\right)+\frac{1}{\psi}(E_{t+1}-E_t)\left[\Delta c_{t+j+1}\right] \quad \text{(C.17)}\end{aligned}$$

for $j \geq 1$. Substituting equation (C.17) into equation (C.16) yields:

$$\begin{aligned}\Delta c_{t+1}-E_t\left[\Delta c_{t+1}\right] &= r_{w,t+1}-E_t\left[r_{w,t+1}\right] \\ &\quad -\left(1-\frac{1}{\psi}\right)(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\Delta c_{t+1+j} \\ &\quad +(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\log\left(\frac{\lambda_{t+j}}{\lambda_{t+j+1}}\right) \quad \text{(C.18)}\end{aligned}$$

Substituting out consumption shock covariance (σ_{ic}) from equation (3.7) yields risk premia as a function of covariances with market returns and innovations to future time preferences and consumption growth:

$$\begin{aligned}E_t\left[r_{i,t+1}\right]-r_{f,t+1}+\frac{1}{2}\sigma_i^2 &= \gamma\sigma_{iw} \\ &\quad +(\gamma-1)\frac{1}{\psi}cov_t\left(r_{i,t+1},(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\Delta c_{t+1+j}\right) \\ &\quad +\frac{\theta}{\psi}cov_t\left(r_{i,t+1},(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\log\left(\frac{\lambda_{t+j}}{\lambda_{t+j+1}}\right)\right) \quad \text{(C.19)}\end{aligned}$$

Equation (3.9) expresses this as:

$$E_t\left[r_{i,t+1}\right]-r_{f,t+1}+\frac{1}{2}\sigma_i^2=\gamma\sigma_{iw}-\frac{\gamma-1}{\psi-1}\sigma_{ih(\lambda)}+(\gamma-1)\,\sigma_{ih(c)}$$

where

$$\sigma_{ih(\lambda)}=cov_t\left(r_{i,t+1},(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\log\left(\frac{\lambda_{t+j}}{\lambda_{t+j+1}}\right)\right) \quad \text{(C.20)}$$

and

$$\sigma_{ih(c)}=\frac{1}{\psi}cov_t\left(r_{i,t+1},(E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\Delta c_{t+1+j}\right) \quad \text{(C.21)}$$

are the two different types of interest rate news covariance.

Another way to see this is to change notation to consider time preference shocks in the same units as consumption. Specifically, consider augmented consumption, defined as:

$$\widetilde{C}_t \equiv \lambda_t^* C_t \tag{C.22}$$

where

$$\lambda_t^* \equiv \lambda_t^{1/(1-1/\psi)} \tag{C.23}$$

With this notation change, equation (3.1) is transformed into standard Epstein-Zin preferences with respect to augmented consumption. All of Campbell (1993) and Bansal & Yaron (2004) results hold with respect to augmented consumption and returns measured in units of augmented consumption. In particular, the augmented risk free rate is:

$$\widetilde{r}_{f,t+1} = -\log(\delta) + \frac{1}{\psi} E_t\left[\Delta\widetilde{c}_{t+1}\right] - \frac{1-\theta}{2}\sigma_w^2 - \frac{\theta}{2\psi^2}\sigma_c^2 \tag{C.24}$$

and the risk premium for any asset is given by

$$E_t\left[\widetilde{r}_{i,t+1}\right] - \widetilde{r}_{f,t+1} + \frac{1}{2}\sigma_i^2 = \gamma\sigma_{iw} + (\gamma - 1)\,\sigma_{ih(\widetilde{c})} \tag{C.25}$$

where tildes represent augmented consumption and returns. Using the identities $\widetilde{r}_{i,t+1} = r_{i,t+1} + \frac{1}{1-1/\psi}\log\left(\frac{\lambda_{t+1}}{\lambda_t}\right)$ and $\Delta\widetilde{c}_{t+1} = \Delta c_{t+1} + \frac{1}{1-1/\psi}\log\left(\frac{\lambda_{t+1}}{\lambda_t}\right)$, equations (C.24) and (C.25) are equivalent to equations (3.6) and (3.9). The time preference risk premia in equations (3.9) and (3.11) blow up as ψ gets close to 1 because time preferences (λ_t) have an outsized impact on augmented consumption through $\lambda_t^* = \lambda_t^{1/(1-1/\psi)}$.

C.3 Substituting out Wealth Returns (CCAPM)

We can also use the budget constraint to substitute out wealth portfolio return covariance (σ_{iw}) from equation (3.7) by rearranging equation (C.18) and using it to decompose σ_{iw}, thereby yielding equation (3.11):

$$E_t\left[r_{i,t+1}\right] - r_{f,t+1} + \frac{1}{2}\sigma_i^2 = \gamma\sigma_{ic} + (\gamma\psi - 1)\,\sigma_{ih(c)} - \frac{\gamma\psi - 1}{\psi - 1}\sigma_{ih(\lambda)}$$

C.4 Disciplining Parameter Values

In a three period setting with $\lambda_0 = \lambda_1 = \delta = 1$, AELR utility can be expressed as:

$$U_0 = \max_{C_0} \left\{ C_0^{1-1/\psi} + \left(E_0 \left[\max_{C_1, C_2} \left\{ C_1^{1-1/\psi} + \lambda_2 C_2^{1-1/\psi} \right\}^{\frac{1-\gamma}{1-1/\psi}} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{1/(1-1/\psi)} \tag{C.26}$$

The Euler equation for an Arrow-Debreu security that pays off in state s is:

$$\begin{aligned} P_s C_0^{-1/\psi} \quad = \quad & \left[\pi_L \left(C_1^{1-1/\psi} + \lambda_L C_2^{1-1/\psi} \right)^{\frac{1-\gamma}{1-1/\psi}} + \pi_H \left(C_1^{1-1/\psi} + \lambda_H C_2^{1-1/\psi} \right)^{\frac{1-\gamma}{1-1/\psi}} \right]^{\frac{\gamma-1/\psi}{1-\gamma}} \\ & * \pi_s \left(C_1^{1-1/\psi} + \lambda_L C_2^{1-1/\psi} \right)^{\frac{1/\psi-\gamma}{1-1/\psi}} * C_1^{-1/\psi} \end{aligned} \tag{C.27}$$

where P_s is the state price for state s, π_s is the probability of state s, and λ_s is the value of λ_2 in state s.

Under our assumption that $C_0 = C_1 = C_2 = C$, equation (C.27) reduces to:

$$P_s = \pi_s \left(1 + \lambda_s\right)^{\frac{1/\psi-\gamma}{1-1/\psi}} \left[\pi_L \left(1 + \lambda_L\right)^{\frac{1-\gamma}{1-1/\psi}} + \pi_H \left(1 + \lambda_H\right)^{\frac{1-\gamma}{1-1/\psi}} \right]^{\frac{\gamma-1/\psi}{1-\gamma}} \tag{C.28}$$

Equation (C.28) immediately implies the state price ratio given by equation (3.26):

$$\frac{P_L}{P_H} = \frac{\pi_L}{\pi_H} \left(\frac{1 + \lambda_L}{1 + \lambda_H} \right)^{-\frac{\gamma-1/\psi}{1-1/\psi}}$$