Commodity Inattention

Alexander Chernyakov

First Draft: April 2014

This Draft: March 2016

Chernyakov is at Harvard University Department of Economics and Harvard Business School. Email: cherny@fas.harvard.edu. I am deeply grateful for in-depth conversations regarding this research with John Campbell, Lauren Cohen, Robin Greenwood, and Adi Sunderam. I am also thankful for useful comments and suggestions from Malcolm Baker, Johnny Kang, Sam Kruger, Owen Lamont, Jamie Lee, Derek Schaeffer, Akhil Shashidhar, David Smalling, Eric Zwick and Harvard Finance Seminar participants.

Abstract

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.

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 and 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 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, 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 un-

[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.

derreact 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 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 and 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 and 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 and 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 and 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 and 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 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 3 examines several channels of friction that could prevent information about commodity returns from being efficiently incorporated into equity prices. Section 4 presents robustness to other effects and choice of methodologies. Finally, Section 5 concludes.

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

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

Metal. Table 1 lists the commodities used in the study by commodity sector; this set covers the most widely studied commodities in the literature. Futures contract prices and specifications are obtained from Bloomberg as in Koijen et al. (2013). Equity data is obtained from CRSP and Compustat.

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.

$$R^{fut}_{\tau,d,t} \equiv \frac{F_{\tau,d,t} - F_{\tau+1,d,t-1}}{F_{\tau+1,d,t-1}} \tag{2.1}$$

$$R_{d,t} \equiv \frac{1}{N_d} \sum_{k=1}^{N_d} R^{fut}_{k,d,t} \tag{2.2}$$

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 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.

$$R_{c,t} \equiv \frac{1}{M_c} \sum_{d=1}^{M_c} R_{d,t} \tag{2.3}$$

where M_c is the number of commodities in sector c.

[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.

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:

$$IR_{i,t} = a + \beta_{i,m,t}R_{m,t} + \beta_{i,E,t}R_{E,t} + \beta_{i,A,t}R_{A,t} + \beta_{i,Me,t}R_{Me,t} + \varepsilon_{i,t} \tag{2.4}$$

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 (2.4) 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 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 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 (2.4).

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 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 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

[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.

$|\beta|$) 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.

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 (2.4) characterizes firms based on contemporaneous relationships with the commodity sectors. I define commodity news for stock i, $R_{i,c,t}$, in equation (2.5) 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 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:

$$R_{i,c,t} \equiv \boldsymbol{\beta'_{j,c,t}} \boldsymbol{R_{c,t}} \tag{2.5}$$

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

[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.

(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 4 and 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 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 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 2 plots the results of this experiment. The results are consistent with underreaction to information: both small and large stocks 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.

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 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 7 provides summary statistics on the funds whose holdings are used to construct inattention measures.

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

[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.

i held by fund f at time t, I define the amount (in dollars) held of that stock by f as:

$$H_{i,f,t} = SHARES_{i,f,t} P_{i,t} \tag{3.1}$$

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:

$$\begin{aligned} \Delta H_{i,f,t+1} &= SHARES_{i,f,t} P_{i,t+1} - SHARES_{i,f,t} P_{i,t} \\ &= SHARES_{i,f,t} \Delta P_{i,t+1} \end{aligned} \tag{3.2}$$

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:

$$\sigma^2(\Delta H_{i,f,t}) = \frac{1}{T-1} \sum_{t=1}^{T} \left(\Delta H_{i,f,t} - \overline{\Delta H}_{i,f,t} \right)^2 \tag{3.3}$$

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:

$$RAWATTN^s_{i,f,t} = \frac{\sigma^2(\Delta H_{i,f,t})}{\sum_{j=1}^{K_f} \sigma^2(\Delta H_{j,f,t})} \tag{3.4}$$

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:

$$RAWATTN^s_{i,t} = \sum_{f \in F_i} RAWATTN^s_{i,f,t} \tag{3.5}$$

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

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

run monthly Fama-MacBeth regressions of the form:

$$\begin{aligned} RAWATTN^s_{i,t} &= \theta_{0,t} + \theta_{bm,t} log(BM_{i,t}) + \theta_{me,t} log(ME_{i,t}) + \theta_{io,t} IO_{i,t} + \\ &+ \theta_{br,t} BREADTH_{i,t} + \theta_{mb,t} \beta_{m,i,t} + \theta_{r,t} R_{i,t} + \theta_{mom,t} R_{i,t-12\rightarrow t-1} + \\ &+ \theta_{iv,t} IV_{i,t} + \varepsilon^s_{i,t} \end{aligned} \tag{3.6}$$

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

$$ATTN^s_{i,t} \equiv \varepsilon^s_{i,t} \tag{3.7}$$

This particular metric captures the effects that I would like to demonstrate while controlling for already known factors affecting anomalies. Table 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 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 and 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 9 and 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 9 and 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.

Results presented in Table 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 (3.6). 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.

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:

$$\beta_{f,c,t} \equiv \sum_{i=1}^{K_f} w_{i,f,t}\beta_{i,c,t} \tag{3.8}$$

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 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:

$$SRAWATTN^n_{i,c,t} = \frac{1}{F_i}\sum_{f=1}^{F_i} |\beta_{f,c,t}| \tag{3.9}$$

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

$$RAWATTN^n_{i,t} \equiv \sum_{c=1}^{C_i} w_{i,c,t} SRAWATTN_{i,c,t} \tag{3.10}$$

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 (3.6):

$$\begin{aligned} RAWATTN^n_{i,t} &= \theta_{0,t} + \theta_{bm,t}log(BM)_{i,t} + \theta_{me,t}log(ME_{i,t}) + \theta_{io,t}IO_{i,t} + \\ &+ \theta_{br,t}BREADTH_{i,t} + \theta_{mb,t}\beta_{m,i,t} + \theta_{r,t}R_{i,t} + \theta_{mom,t}R_{i,t-12\to t-1} + \\ &+ \theta_{iv,t}IV_{i,t} + \varepsilon^n_{i,t} \end{aligned} \tag{3.11}$$

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

$$ATTN^n_{i,t} \equiv \varepsilon^n_{i,t} \tag{3.12}$$

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 11. Similarly to $RAWATTN^s$ it has a negative loading on BM and a positive loading on IV. The R^2 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 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 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 13. The $5-1$ portfolio in the

"Low" attention category generates an alpha of 1.4% per month (N-W 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.

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 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 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 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

$$conews_{i,t} \equiv log(1 + CONEWS_{i,t}) \tag{3.13}$$

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

$$cmdtyconews_{i,c,t} \equiv log(1 + CMDTYCONEWS_{i,c,t}) \tag{3.14}$$

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:

$$cs_{i,c,t} \equiv \frac{cmdtyconews_{i,c,t}}{conews_{i,t}} \tag{3.15}$$

Companies may have multiple commodity sectors associated with them, as explained in Section 2.2, so I follow the same procedure as in Section 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 $cs_{i,c,t}$:

$$RAWCS_{i,t} \equiv \sum_{c=1}^{C} w_{i,c,t} cs_{i,c,t} \tag{3.16}$$

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 the form:

$$\begin{aligned} RAWCS_{i,t} &= \theta_{0,t} + \theta_{bm,t} log(BM)_{i,t} + \theta_{me,t} log(ME_{i,t}) + \theta_{io,t} IO_{i,t} + \\ &+ \theta_{br,t} BREADTH_{i,t} + \theta_{mb,t} \beta_{m,i,t} + \theta_{r,t} R_{i,t} + \theta_{mom,t} R_{i,t-12\rightarrow t-1} + \\ &+ \theta_{iv,t} IV_{i,t} + \varepsilon^{cs}_{i,t} \end{aligned} \tag{3.17}$$

and defining the residualized version of commodity salience as

$$CS_{i,t} \equiv \varepsilon^{cs}_{i,t} \tag{3.18}$$

Table 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 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 15 and 16 provide the results of this experiment.

Examining Table 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 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 and Lou (2012), or because the connection to the information source may be nuanced.

3.4 Time Varying Cognitive Burden

Section 3.1 and 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,

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

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)).

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:

$$XD_t \equiv \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(R_{i,t}-\overline{R_t}\right)^2} \tag{3.19}$$

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 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 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 17 and 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.

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 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:

$$
\begin{aligned}
R^e_{i,t+1} &= a_{t+1} + \left(\beta_0 + \beta_1 \mathbf{1}_{HIB,t} + \beta_2 RAWCS_{i,t} + \beta_3 RAWATTN^n_{i,t} + \beta_4 RAWATTN^s_{i,t}\right) R_{i,c,t} + \qquad (3.20)\\
&\quad + \eta_1 \mathbf{1}_{HIB,t} + \eta_2 RAWCS_{i,t} + \eta_3 RAWATTN^n_{i,t} + \eta_4 RAWATTN^s_{i,t} + \boldsymbol{\theta}' \boldsymbol{X_{i,t}} + \varepsilon_{i,t+1}\\
&= a_{t+1} + \boldsymbol{\beta}' \left[R_{i,c,t}\ \boldsymbol{\Gamma_{i,t}} R_{i,c,t}\right]^T + \boldsymbol{\eta}' \boldsymbol{\Gamma_{i,t}} + \boldsymbol{\theta}' \boldsymbol{X_{i,t}} + \varepsilon_{i,t+1} \qquad (3.21)
\end{aligned}
$$

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 3.4; $RAWCS_{i,t}$ is the commodity salience associated with stock i at time t from Section 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 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 3.1, and $\boldsymbol{X_{i,t}}$ is a vector of control variables:

$$\boldsymbol{X_{i,t}} = [log(BM_{i,t}),\ log(ME_{i,t}),\ IO_{i,t},\ BREADTH_{i,t},\ \beta_{m,i,t},\ R_{i,t-1},\ R_{i,t-13\to t-2},\ IV_{i,t}] \quad (3.22)$$

as in Eq 3.6. Regression (3.20) 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 (3.20), 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):

$$\overline{R^e}_{t+1} = a_{t+1} + \boldsymbol{\beta'} \left[\overline{R}_{c,t}\ \overline{\boldsymbol{\Gamma_{i,t}} R_{i,c,t}}\right]^T + \boldsymbol{\eta'}\overline{\boldsymbol{\Gamma_{i,t}}} + \boldsymbol{\theta'}\overline{\boldsymbol{X_{i,t}}} + \overline{\varepsilon}_{t+1} \quad (3.23)$$

Subtracting (3.23) from (3.20) yields:

$$\widetilde{R^e}_{i,t+1} = \boldsymbol{\beta'} \left[\widetilde{R}_{i,c,t}\ \widetilde{\boldsymbol{\Gamma_{i,t}} R_{i,c,t}}\right]^T + \boldsymbol{\eta'}\widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{i,t}} + \boldsymbol{\theta'}\widetilde{\boldsymbol{X}}_{\boldsymbol{i,t}} + \widetilde{\varepsilon}_{i,t+1} \quad (3.24)$$

where the tilde over the variable indicates the cross-sectional demeaned variable for that time period, for example $\widetilde{\boldsymbol{X}}_{\boldsymbol{i,t}} = \boldsymbol{X_{i,t}} - \overline{\boldsymbol{X}}_{\boldsymbol{t}}$. Equation (3.24) 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 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 thus far. The magnitude is not significantly affected by any of the presented controls (more controls will be added in Section 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.

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).

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 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 and 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 and Grinblatt (1999) and intra-industry large to small stock information diffusion of Hou (2007). The industry momentum 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 20. The controls do not decrease the significance of the result that I have shown in previous sections: stocks underreact to commodity news.

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 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 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 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.

4.3 Individual Commodities and Alternative Methodology

While the stock universe selection mechanism described in Section 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 and 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 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 and 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 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:

$$\min_{a,\beta} \frac{1}{T} \sum_{t=1}^{T} \left(IR_{j,t} - a_{j,t} - \boldsymbol{\beta'_{j,t} R_{d,t}} \right)^2 + \lambda \left[(1-\alpha) \|\boldsymbol{\beta_{j,t}}\|_2^2 / 2 + \alpha \|\boldsymbol{\beta_{j,t}}\| \right] \tag{4.1}$$

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 and 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.

$$RZ^{alt}_{i,c,t} = \frac{R^{alt}_{i,c,t} - \overline{R^{alt}_{c,t}}}{\sigma_t(R^{alt}_{i,c,t})} \tag{4.2}$$

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 23 and 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 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.

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.

A Appendix

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)
```

References

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

Barber, B. M. and 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.

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

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

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

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

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

Hirshleifer, D., Lim, S. S. and 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. and Kennard, R. W. (1970), 'Ridge regression: Biased estimation for nonorthogonal problems', *Technometrics* **12**(1), 55–67.

Hong, H., Lim, T. and 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. and 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. and 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. and Vrugt, E. B. (2013), 'Carry', *Working Paper* .

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

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

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. and 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.

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

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.

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

Zou, H. and 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.

Table 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		

Table 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 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

Table 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

Table 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 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 $R_{i,c,t} \equiv \boldsymbol{\beta}'_{\boldsymbol{j,c,t}} \boldsymbol{R_{c,t}}$ (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	-0.059	-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 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]**							

Table 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

Table 8: Stock Inattention Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (3.6). 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 (3.7), 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]**		

Table 9: Individual Stock Attention - $RAW\,ATTN^s$

At the end of each month stocks are sorted into low and high attention stocks using $RAW\,ATTN^s_{i,t}$ 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.

$RAW\,ATTN^s$	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}
	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
Low			[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]**							
	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]							
High	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 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]**							

Table 11: News Inattention Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (3.11). 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 (3.12), 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.

	$log(BM_{i,t})$	$log(ME_{i,t})$	$IO_{i,t}$	$BREADTH_{i,t}$	$\beta_{i,m,t}$	$R^e_{i,t}$	$R^e_{i,t-12\rightarrow t-1}$	$IV_{i,t}$	(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]		

Table 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]**							

Table 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 (3.11)) 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}
	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
Low			[-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]**							
	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
High			[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]**							

Table 14: Commodity Salience Metric Fama-MacBeth Residualizing Regressions

Results of Fama-MacBeth residualizing regressions, equation (3.17). 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]		

Table 15: Commodity Salience Using $RAWCS$

Stocks are sorted into low and high salience categories based on $RAWCS_{i,t}$ (3.16). 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 16: Commodity Salience Using CS

Stocks are sorted into low and high salience categories based on $CS_{i,t}$ (3.18). 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}
Low	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
			[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]							
High	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
			[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]**							

Table 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 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	-0.937	**-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
			[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]**							
Low	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
			[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 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 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]**							

Table 19: Panel Regressions of Inattention Channels

To verify the uniqueness of each inattention channel, I run a pooled panel regression, eq (3.24), 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]**		

Table 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.

	$R_{i,c,t}$	$R^e_{i,t}$	$log(ME_{i,t})$	$R^e_{i,t-12\rightarrow t-1}$	$bm_{i,t}$	$IR^e_{i,t-11\rightarrow t}$	$IR^e_{i,t-12\rightarrow t-1}$	$IR^e_{i,t}$	$IR^e_{i,t-1}$	$LIR^e_{i,t}$	(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]		

Table 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

Table 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]**							

Table 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 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 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 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]**							

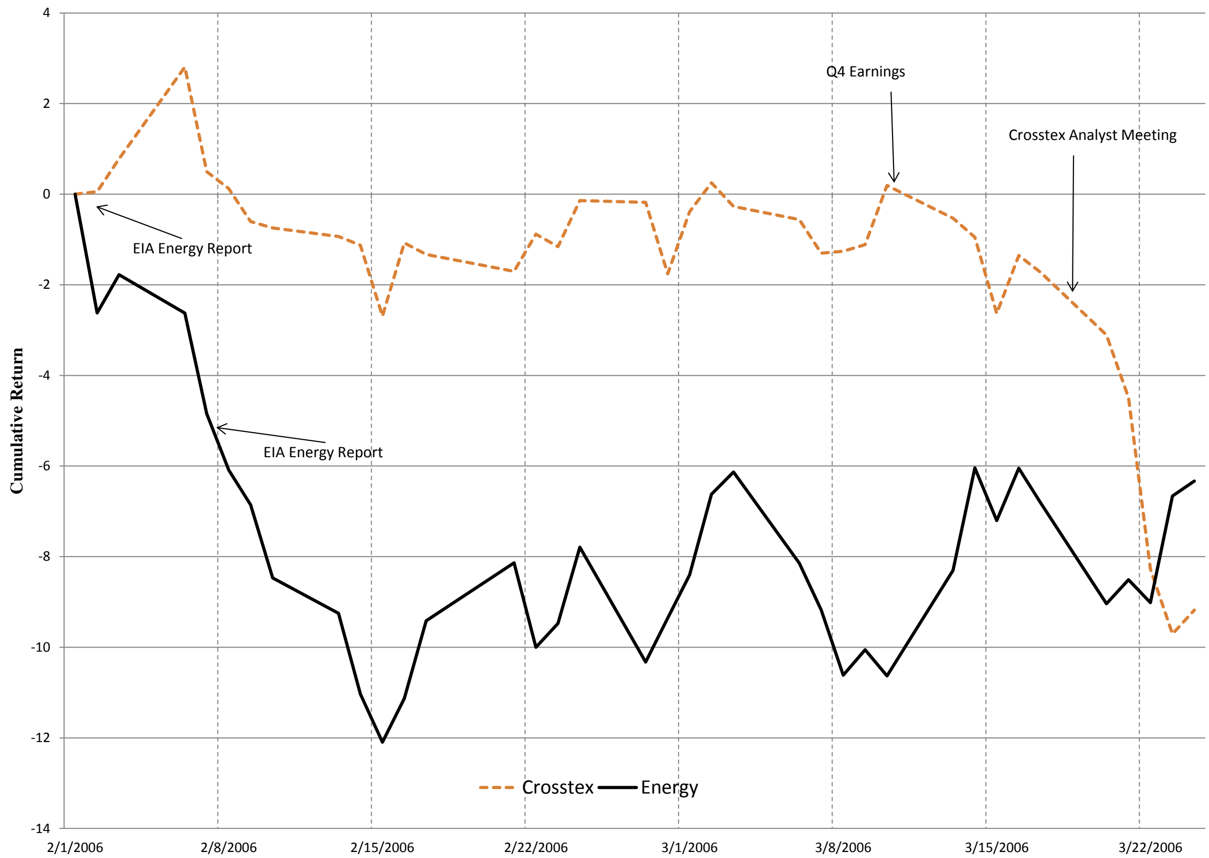

Figure 1: XTXI and Energy Cumulative Returns

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.

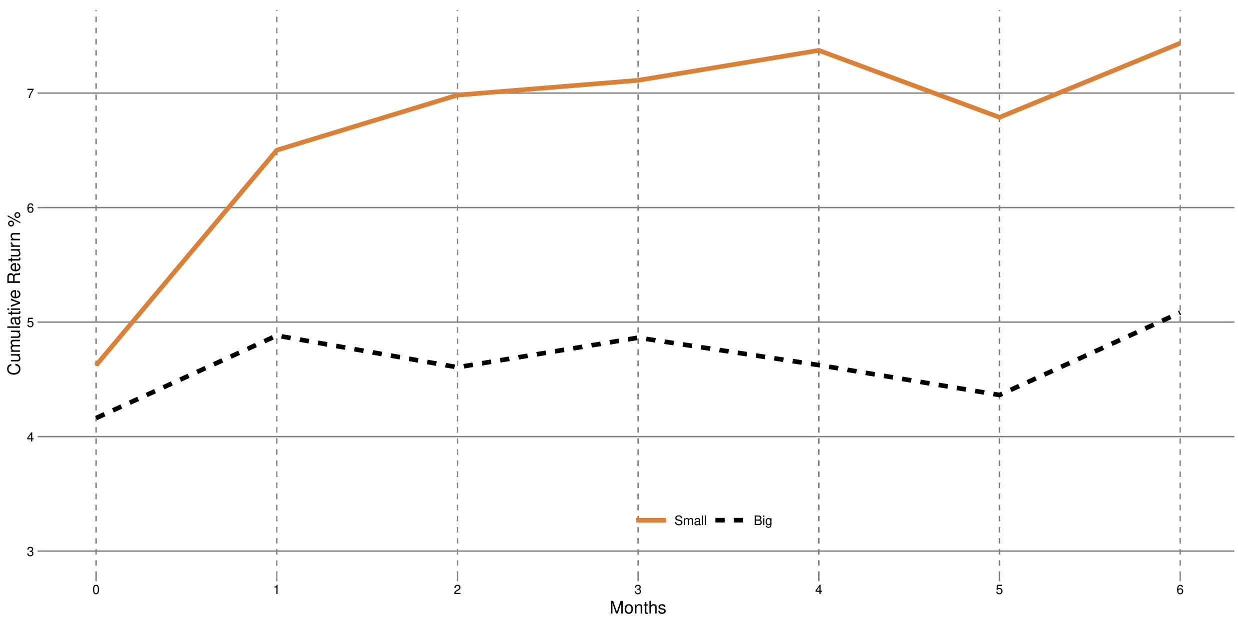

Figure 2: Inattention Horizon

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