Commodity Option Implied Volatilities and the Expected Futures Returns

Lin Gao *

November 12, 2017

Abstract

The detrended implied volatility of commodity options (VOL) forecasts the cross section of the commodity futures returns significantly. A zero-cost strategy that is long in low VOL and short in high VOL commodities yields an annualized return of 12.66% and a Sharpe ratio of 0.69. Notably, the excess returns based on the volatility strategy emanate mainly from its forecasting power for the future spot component, different from the other commodity strategies examined so far in the literature which are all driven by roll returns. This strategy demonstrates low correlations (below 10%) with the other strategies such as momentum or basis and performs especially well in recessions. Our results are robust after controlling for illiquidity, other commodity pricing factors, and exposure to the aggregate commodity market volatility. The VOL measure is associated with hedging pressure on the futures and especially on the options market. News media also helps amplify the uncertainty impact. Variables related to investors' lottery preferences and market frictions are able to explain part of the predictive relationship.

*Lin Gao (lin.gao@uni.lu) is at the Luxembourg School of Finance, University of Luxembourg. Address: Campus Kirchberg (Bloc F), rue Richard Coudenhove-Kalergi 6, L-1359 Luxembourg. E-mail: lin.gao@uni.lu. Tel.: +352 466 644 5482.

1 Introduction

What is the relationship between volatility and the expected returns of commodity futures? The Intertemporal Capital Asset Pricing Model (ICAPM) of Merton (1973) suggests that the excess return of the market portfolio should be positively related to the volatility of the market. Economic theories also postulate a positive relationship between idiosyncratic volatility of individual assets and future expected returns. Merton (1987) argues that in an information-segmented market, investors demand higher returns for firms with larger firm-specific volatilities. This additional risk premium is a result of compensation for holding imperfectly diversified portfolios. Coincidentally, behavioral models such as Barberis and Huang (2001) predict higher expected returns for higher volatility stocks as well.

The goal of this paper is to examine the cross-sectional relationship between option-implied volatilities and expected returns of commodity futures. We look at a sample of 25 commodity futures and their related option implied volatilities in the post-1990 period, where options of most commodities are available. Our findings uncover that commodities with high VOL earn *lower* returns in the following month, and the return difference between the highest and lowest VOL commodities is statistically significant. This evidence contradicts the existing theories which imply a positive volatility–return relationship. Furthermore, this result also departs to certain extent from the empirical findings of the volatility risk premium literature on the equities and FX markets. For example, Bollerslev et al. (2009) and Corte et al. (2016) show that the volatility risk premium, defined as the difference between the implied volatility and realized volatility of the respective market, positively forecast future returns.

Interestingly, this evidence from the commodity market echoes the similar puzzling findings in the equity market. Ang et al. (2006) sort individual stock returns on their respective volatility or idiosyncratic volatility. They find that stocks with high idiosyncratic volatility from last month have low average expected returns and the difference between portfolio with the highest and lowest idiosyncratic volatility stocks is significantly negative.

We form a strategy that sells commodities with the highest VOL and buys commodities with the lowest VOL in month t and earns the zero-cost return in month $t+1$. Our strategy yields an average annualized return of 12.66% and a Sharpe ratio of 0.69. Notably, the strategy return is almost entirely driven by the spot return component. This means that commodities with the most expensive volatility insurance tend to experience price decline in the next month, whereas commodities with the cheapest volatility insurance become more expensive in the subsequent period. We further show that our strategy returns are not mainly driven by illiquid commodity futures.

The fact that the VOL strategy return emanates mainly from the spot return component is of crucial importance for commodity investment and risk management. So far, literature has mainly documented commodity investment strategies based on signals from the term structure (commodity futures contracts of different maturities) or roll return component. This means that an important part of those strategy returns in month $t+1$ come from the "artificial" profit of selling the more expensive nearby contract and buying the less expensive next nearby contract in month t. As a result, these strategy returns including momentum are predominantly attributed to the positive roll return component. However, little knowledge has been acquired regarding the spot component or time series evolution of the commodity prices. Such kind of knowledge is apparently important in face of risks and the potential state variables such as volatility that drive the evolution of prices over time.

Therefore, our findings favor a portfolio diversification with the volatility strategy. Alternatively, combining the volatility strategy with the other strategies in the double portfolio sorts earns much higher returns thanks to the different sources of profitability. In addition, one desirable feature of our strategy is that the volatility strategy performs particularly well during recessions, as the VOL strategy acts as a natural hedge against the extremely volatile periods.

So far, most of the commodity literature deals with the realized volatilities and returns with a few exceptions[1]. However, these studies typically employ volatility of commodities as a characteristic to explain returns based on other state variables or test volatility endogenously as a function of other pricing factors[2], whereas we investigate volatility and its predictability for commodity returns. We find that the option-implied volatility demonstrates superior predictive power to the realized volatility, the variance risk premium, and option implied skewness both in terms of cross-sectional strategy return and time series predictability for the overall commodity market. Furthermore, the predictability of the implied volatility is not solely driven by its good or bad volatility component. Tail risk accounts for a substantial part of the VOL predictive ability, but does not outperform the VOL strategy.

It is essential to understand what drives the volatility of commodity and its strategy returns. In the time series approach, we show that the volatility strategy return is little correlated with the factor returns represented by the other commodity strategies or the Fung and Hsieh (2001) hedge fund factors. Neither are they related to the risk factors from the equity or bond markets, or business cycle factors. While our strategy benefits from increasing commodity market volatility, it is not a good hedge for the equity market volatility. In addition, the VOL

[1]For example, Bakshi et al. (2016), Dhume (2011), Ng and Pirrong (1994), Gorton et al. (2013), or Fama and French (1987) examine realized volatilities. Some papers apply implied volatility or volatiliy risk premium for crude oil. For example, Pan and Kang (2015) find a negative predictive ability of oil variance risk premium for oil futures returns. Christoffersen and Pan (2015) find the change in implied volatility in crude oil can predict the cross-section of stock returns after the financialization period.

[2]For example, Ng and Pirrong (1994) or Fama and French (1987) test commodity volatility as a function of inventory or commodity basis based on the theory of storage. They focus on the relationship between volatility and the lagged variables such as basis or inventory. Gorton et al. (2013) characterize the cross section of commodity portfolio returns with their contemporaneous volatilities and relate the results to inventory. Our study instead tests the predictive ability of volatility at t for commodity returns of the next period t+1.

return is related to the inter-bank borrowing capacity (TED spread) or to the broker-dealer index.

These results prompt us to examine the determinants of VOL and the sources for predictability in more details. Previous literature on commodity futures risk premia focuses mainly on the following three aspects: hedging pressure, systematic risks, and noise trader risks.[3] We examine these explanations by employing traders' positions data, the aggregate commodity market risks, and the media exposure of the individual commodities.

First, the theory of normal backwardation by Keynes (1930) and Hicks (1930) proposes that hedgers pay an additional premium to transfer their risk to speculators, which causes the futures contracts to be traded below the expected future spot prices. As such, hedging pressure generates a positive premium for buying the futures contracts. We show that the volatility factor is negatively linked to the hedging pressure of producers, who are usually on the short side of the futures contracts. In comparison it is positively correlated with net short positions and hence negatively associated with the net long positions of the speculators, especially money managers who are mainly composed of hedge funds, whereas the role of swap dealers (index traders) is not clear. The results become more accentuated when we observe options market alone or combined with the futures markets.

These findings lend support to the limits to arbitrage interpretation for the returns delivered by the volatility strategy, as illustrated in Acharya et al. (2013). For the limits to arbitrage to take effect would require risk-averse hedgers on the one side and capital-constrained arbitrageurs (speculators) on the other side. Brunnermeier and Pedersen (2009) show that when margin-constrained speculators trade in multiple assets across time periods, assets with higher volatility are more sensitive to changes in speculator's wealth. The risk-averse hedgers of commodities (typically producers) prefer to hedge commodities with comparably cheap

[3] For a brief summary of the related literature, see for example Adam and Fernando (2006) for the first two views and Tang and Xiong (2012) or Cheng et al. (2015) for the financialization literature.

hedging costs (low option implied volatility). When the major part of the producers are on the short side of the market, they typically intend to secure the current price level. Therefore, they are willing to give up an additional premium to resolve uncertainty according to the theory of normal backwardation (Hirschleifer, 1990). Thus commodity futures are traded below the expected price. As a result future commodity returns increase. These results are in line with those from the other asset markets such as currencies. The only difference is that the hedgers in these markets are typically on the long side (Corte et al., 2016). Conversely, when speculators are capital constrained and volatility insurance is high, producers are forced to reduce their inventory. Then future spot prices decline.

A second strand of theory relates to the systematic risks resulting from undiversifiable risks in the futures markets (see, for example, Dusak (1973) or Black (1976)). However, the returns generated by the VOL strategy with individual commodity volatilities are not related to the returns generated by the aggregate commodity market volatility, meaning that our strategy returns do not stem from the compensation for the overall commodity market risk.

A third strand of the literature introduces noise trader risks, especially in face of the financialization of the commodity futures markets. As a proxy for noise trader risks and investors' attention we employ the media coverage measure. Interestingly, media coverage also generates similar portfolio return patterns to VOL: high media coverage leads to lower returns in the next period. However, the two types of returns are not highly correlated. Media uncertainty cannot explain the option implied volatility of commodities in the panel, suggesting that these two types of returns are driven by difference sources.

In addition to the explanations discussed above, we explore candidates related to the idiosyncratic volatility literature: variables on investors' lottery preferences and market frictions. Following the approach in Hou and Loh (2016), we conduct a four-stage procedure and decompose the coefficient of VOL in the Fama-MacBeth predictive regressions into one component explained by the candidate variables and one orthogonal residual part. Our results show that the predictive ability of VOL does come to some extent from these sources. For example, the maximum return from last month, return reversal or illiquidity measure show promise in explaining the predictability. However, there still remains a large significant part unexplained by these variables.

Our paper provides a broad overview of the strategies examined so far in the literature. The basis predictor has been addressed as a proxy for fundamentals such as inventory (Fama and French, 1987). Yang (2013) rationalizes the basis factor with investment shocks. Szymanowska et al. (2014) provide evidence that the basis factor explains the cross section of the nearby commodity returns, whereas two additional portfolios explain the spreading returns. Hong and Yogo (2012) follow a time series approach to show the predictive ability of open interest for future commodity returns. We find evidence that open interest growth also predicts commodity returns cross-sectionally. Bakshi et al. (2016) argue that an average commodity factor, a carry factor (basis), and a momentum factor are capable of describing the cross-sectional variation of commodity returns. Boons and Prado (2016) add the basis-momentum factor into the pricing literature, where the average of 12-month return differences at two different points on the futures curve predict the cross section of commodity futures returns. De Roon et al. (2000) discovers that producers' hedging pressure predicts future commodity returns.

We demonstrate that the model incorporating the volatility factor is better aligned with the data compared to the models with the existing factors studied by Szymanowska et al. (2014), Bakshi et al. (2016) or Boons and Prado (2016). The volatility factor premium is large and statistically highly significant both in the Fama-MacBeth cross-sectional settings as well as in the GMM approach in a stochastic discount factor. Thus, it significantly improves the model fit. The Kan and Robotti (2009) test of the null hypothesis that the difference between the Hansen-Jagannathan distance of our model and the other models is rejected, suggesting that the option implied volatility serves as a pricing factor for the cross-section of commodity returns. Applying a new approach developed by Giglio and Xiu (2017), we yield similar robust results for the VOL factor.

Beyond the research papers mentioned above, our study complements a growing body of literature studying the general behavior of the broad commodity market over time such as Erb and Harvey (2006), Gorton et al. (2013), or Gorton and Rouwenhorst (2006). Our paper is also partially related to the financialization literature, in which the role of speculators and hedgers are intensively discussed (see, for example, Cheng et al. (2015) or Tang and Xiong (2012)). Furthermore, our paper connects to the financial intermediary literature (Adrian et al., 2014) as well as the effect of financial insititutions (Henderson et al., 2014). Etula (2013) documents that the speculator risk tolerance – the broker dealer measure – strongly predicts commodity futures returns with a negative sign. Our paper differs from the literature on liquidity risk in the commodity markets such as the rollover risk (see, for example, Mou (2011), Marshall et al. (2012), or Neuhierl and Thompson (2016)), in that we study the cross-sectional asset pricing evidence, instead of the impact of some particular event dates.

2 Data Description

Commodity Futures and Options Data. We obtain the commodity futures and options data from Commodity Research Bureau (CRB). Table A.1 provides an overview of the commodities examined in the sample, the exchanges on which they are traded, as well as the option expiration months. As a proxy for the option's liquidity, we present the statistical properties of the dollar value open interest of the commodity options in Table A.2. The dollar value of open interest is backed out as the difference of total open interest between the futures & options combined and the futures only category from the Commitment of Traders Report from CFTC. As the commodity options are of American style, we convert the option prices to European options following the approach by Barone-Adesi and Whaley (1987). Out-of-money call and put options are applied to compute the option implied moments. To ensure sufficient liquidity, we keep options with at least 15 days and at most 8 months to expiry. We further exclude options violating standard no-arbitrage conditions. We exclude options with a price which is below or equal to five times the minimum tick value. We truncate the upper and lower strike at $K_t = F_{t,T} \cdot exp\{\pm 6\sigma(T-t)\}$. Jiang and Tian (2005) find that the truncation error can be ignored if the truncation points are more than two standard deviations away from the forward price. We also try with the alternative truncation points at 10σ. The difference is negligible. Following Bakshi et al. (2003), we compute the model-free option implied volatility and skewness. We use monthly data by sampling end-of-month implied volatilities from January 1990 to October 2014.

We compare the daily implied volatility of crude oil (IVOil) and gold (IVGold) with their counterparts traded on the CME exchange: the 30-day CBOE gold volatility index (GVZ) and the crude oil volatility index (OVX). OVX and GVZ are based on options on the United States Oil Fund and the SPDR Gold Shares. It is evident in Figure A.1 that IVOil and IVGold track their respective counterpart very closely. The correlation coefficient is 99.1% for oil and 98.8% for gold.

In order to obtain a continuous time series of commodity futures returns, we take the nearest-to-maturity contract with maturity month T and roll to the next contract at the end of month $T-2$ to avoid illiquidity issues near expiration. We then splice the daily log returns of the chosen contracts into a continuous time series for each commodity before aggregating the daily log returns to the monthly level.

Other Data. We further employ data on dollar value of commodity open interest, which is computed as the number of futures contracts outstanding multiplied by the size of the respective contract and the price of the nearest-to-maturity futures contract, as well as long and short positions of commercial traders (hedgers) for each futures contract. These data are available from Commodity Futures Trading Commission. The data are available at weekly frequency. We convert them to end-of-month data. Trading volume data are provided by CRB.

3 Strategies

Detrended Implied Volatility (VOL). The commodities examined in our sample represents a wide variety regarding the production region, seasonality, industrial or consumption usage, perishability etc. Such distinct characteristics are also reflected in the diverse implied volatility patterns. As an illustration, Figure 1 presents the implied volatilities of crude oil, copper, corn, and gold. For the ease of comparison, we take the VIX index from the equity market as a benchmark. The commodity volatilities in the figure are very different from each other both in terms of the average level and the time-varying patterns. For example, crude oil IV is on average larger than VIX, whereas gold IV is approximately comparable to VIX. Corn IV exhibits very salient seasonality patterns. Copper's IV peaks around 2006 or 2009 and is much higher than the IV in the 1990s.

All these facts suggest that we should not directly compare the implied volatilities among the commodities due to the different unconditional mean of volatilities as well as the time-varying nature of them. Instead, we should adopt a measure that takes both the level as well as the time-varying properties into account. Gorton et al. (2013), for example, use the difference between the volatility and the unconditional mean of volatility to characterize volatility innovations in commodity futures. However, they did not consider the time-varying aspect of volatility and their measure is ex post based on the whole sample information.

In the following, we propose a detrended volatility measure to extract volatility innovations. Specifically, at the end of each month t, we detrend the 30-day implied volatility by the past 12-month average.

$$VOL_{i,t} = IV_{i,t} - \frac{1}{n}\sum_{\tau=1}^{n} IV_{i,t-\tau}, \tag{1}$$

where $n = 12$. $IV_{i,t}$ refers to the 30-day implied volatilies of commodity i at time t. As such, the last part of Equation 1 serves as a proxy for the conditional mean of the volatility. Using yearly average also has the advantage that we can keep the seasonality shocks in the implied volatility in place, as these are also time-varying. Another appealing feature of this approach is that it can be easily implemented in real time and does not rely on any parameter estimation. For comparison and robustness, Figure 2 displays the strategy return and t-stat regarding the choice of different n. It turns out that though the strategy return is large and statistically significant for all the choices, n=12 or 24 is actually a good proxy for the conditional mean of volatility. No detrending (n=0) or using too short (n=6) or too long window (n=36) loses some information in the estimate.

In each month t, we group commodities into four portfolios using the VOL measure. We allocate the top 25% of all commodities with the highest VOL to Portfolio 1, and 25% of all commodities with the lowest VOL to Portfolio 4. Subsequently, we form the strategy that is long in Portfolio 4 and short in Portfolio 1. We then compute the average excess return within each portfolio at $t + 1$ and calculate the returns of the zero-cost portfolio.

Momentum (MOM). At the end of each period t, we form four portfolios based on the commodity futures returns over the previous 12 months from $t - 11$ to t. We assign the 25% of all commodities with the highest lagged returns to Portfolio 1, and the the 25% of all commodities with the lowest lagged returns to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t + 1$. The strategy is long in Portfolio 1 (winner commodities) and short in Portfolio 4 (loser commodities).

Basis. At the end of each period t, we form four portfolios by sorting on the basis of each commodity futures. The basis in our approach is defined as $[log(F_t^{T_2}) - log(F_t^{T_1})]/(T_2 - T_1)$. $F_t^{T_1}$ is the price of the nearby futures contract and $F_t^{T_2}$ refers to the second nearby futures contract. We assign the 25% of all commodities with the highest basis to Portfolio 1, and the the 25% of all commodities with the lowest basis to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t+1$. The strategy is long in Portfolio 4 (lowest basis) and short in Portfolio 1 (highest basis).

Basis-Momentum (BasMom). Following Boons and Prado (2016), we form four portfolios based on the difference between momentum in the first- and second-nearby futures. We assign the 25% of all commodities with the highest basis-momentum to Portfolio 1, and the the 25% of all commodities with the lowest basis-momentum to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t+1$. The strategy is long in Portfolio 1 (highest BasMom) and short in Portfolio 4 (lowest BasMom).

Hedging Pressure. Following De Roon et al. (2000) we define hedging pressure as the difference of open interest between commercial traders that are short in futures contract and those that are long in futures contract, divided by the total open interest of commercial traders. At the end of each period t, we group four portfolios by sorting on the hedging pressure of each commodity futures. We assign the 25% of all commodities with the highest hedging pressure to Portfolio 1, and the the 25% of all commodities with the lowest hedging pressure to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t+1$. The strategy is long in Portfolio 1 (highest hedging pressure) and short in Portfolio 4 (lowest hedging pressure).

Open Interest Growth. Following Hong and Yogo (2012), we define open interest (in dollar) for each commodity as the nearest-to-maturity futures price times the quantity of futures contracts outstanding times the contract size. We then compute the log growth rate of open interest. Because the monthly growth rate of open interest is noisy, we smooth it by taking a 12-month average in the time series. At the end of each period t, we group four portfolios by sorting on the open interest growth of each commodity futures. We assign the 25% of all commodities with the highest hedging pressure to Portfolio 1, and the the 25% of all commodities with the lowest hedging pressure to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t+1$. The strategy is long in Portfolio 1 (highest open interest growth) and short in Portfolio 4 (lowest open interest growth).

4 Empirical Results

4.1 Main Results

Table 1 presents the returns and their summary statistics of portfolios generated by our VOL strategy. First, our strategy that buys commodities with the highest VOL and sells those with the lowest VOL from the last period yields an annualized return of 12.66%, which is statistically significant at the 1% level. The Sharpe ratio is 0.69. Second, the strategy return mainly stems from the short leg (highest VOL portfolio). The portfolio with highest VOL in the last period bears a loss of 10.08%. Third, in general portfolio returns increase from Portfolio High to Low, though the increase is not monotonic. Fourth, in order to see to which part the strategy return can be attributed to, in Panel B and C we decompose the commodity returns into nearby returns and roll returns as in Szymanowska et al. (2014). If we regard the nearest to maturity contract as a proxy for spot prices, then the nearby returns equal approximately the spot returns. The long–short portfolio returns are almost totally driven by the nearby returns, which can be regarded as the log difference of spot prices, whereas the roll return component of the strategy portfolio is negative (-1.8%). This evidence is similar to the findings in the variance risk premia literature that the implied volatility or variance risk premia mainly drive spot returns (See, e.g. Corte et al. (2016)). In addition, in the nearby return part, the long and short leg contribute almost equally to the long-short return. However, the positive return of the "Low" portfolio is canceled out by the negative roll return.

In Panel D, we display the transition matrix. The commodities rotate across the portfolios based on the transition probabilities. The steady state transition probabilities $\overline{\pi}$ are almost equal among all four portfolios. Those of the High and Low Portfolio are relatively stable compared to the middle portfolios. To further show that the long and short leg of our strategy are not composed of the same set of commodities over time, Table A.3 reports the number of occasions when the VOL of the respective commodity has been the highest (lowest), or among the two, three, or four highest (lowest) across all the commodities. Especially in the two middle columns where the four lowest and highest VOL commodities are reported, the number is very balanced among the commodities, suggesting that the long and short portfolios are not driven by certain commodities over time. In the mean time the relatively stable transition probabilities for these two portfolios give a reasonable turnover regarding transaction costs.

Panel E reports the characteristics of the VOL portfolios before and after the portfolio formation. The VOL measure is relatively persistent. The difference between the highest and lowest VOL portfolio is large and statistically significant before and after the formation. The highest VOL portfolio exhibits higher momentum returns, lower basis, and lower hedging pressure relative to the lowest VOL portfolio. Interestingly, the difference between the Dollar open interest of the highest and lowest portfolio is not significantly different, suggesting that the results are unlikely to be mainly driven by liquidity.

It may be of interest to examine how the strategy return changes with regard to different formation periods. Figure 3 displays the average annualized monthly returns at t generated by sorting on VOL at $t-1$, $t-2$,... or up to $t-12$. Panel A displays the strategy returns and their decomposition into nearby and roll returns. The VOL strategy return is to some extent persistent for the first two lags of VOL, with the second lag still yielding a positive excess return of over 10%. The returns decline dramatically from the third lag on and reverse at the end of the testing period. In line with the results in Table 1, the spot returns mainly drive the excess returns, whereas the roll returns change little over the various formation periods. In order to understand the change of the long and short leg portfolio regarding the different the holding month, Panel B presents the long and short leg returns. It is evident that the short leg returns are more persistent than the long leg returns, which turn negative from the second month on before reversing in the 11th or 12th month. Interestingly, when decomposing the short and long leg returns in Panel C and D, the nearby returns of the long leg portfolio is positive, suggesting that low VOL leads to higher nearby returns and hence rising futures prices. The nearby returns for the long leg are large and remain positive until up to the 8th month ahead.

In addition, we test whether the VOL returns emanate from the unconditional volatility difference among the commodities. Though we have shown in Table A.3 that the portfolio composition is not concentrated on certain commodities, we formally confirm this impression by forming portfolios based on the unconditional VOL. The results are reported in Table 3. First, Panel A shows that $\overline{VOL}$ underperforms its conditional counterpart. The average annual excess return of 6.10% is small and statistically insignificant. Moreover, the return is almost totally driven by the roll component, instead of by the nearby return which is the case for the conditional strategy. Third, the correlations with the other strategies are high. For example, the correlation with momentum is 0.67 and that with open interest is 0.49. To conclude, the VOL strategy return depends crucially on the conditional sort of the volatility and is not attributed to the unconditional characteristics of commodities.

Furthermore, we address the potential criticism that our strategy results may be driven by illiquid commodities. We choose those commodities that are both members of the two major commodity indices: the S&P GSCI index (formerly Goldman Sachs Commodity Index) and the Bloomberg Commodity Index (formerly DJ UBS or DJ AIG index). These two indices are the most widely recognized benchmarks and include the most liquid commodity futures. The last three columns of Table A.1 document whether a commodity is component of the two indices respectively as well as whether it is included in our small sample. Table A.5 reports the results. Our strategy that sells the highest VOL and buys the lowest VOL commodities at the end of month t yields a similar annualized return of 12.9% in the small sample, which is significant at the 5% level. Similar to the results with the big sample, the return comes mainly from the nearby return component. This evidence suggests that our strategy is not driven by the illiquid commodity futures and is therefore well tradable.

4.2 Comparing VOL with Other Commodity Strategies

We compare our strategy returns with the returns generated from other strategies employed in the commodity literature to assess the strategy predictability. These strategies are summarized in Section 3.

Panel A of Table 2 displays the excess returns generated by these trading strategies. Consistent with the literature, momentum, basis momentum, and open interest deliver sizable excess returns, whereas the returns obtained by basis and hedging pressure are small and statistically insignificant at the 5% level. The VOL strategy is comparable to that of the momentum strategy. However, our strategy displays a lower standard deviation and hence higher statistical significance for the returns and a higher Sharpe-ratio. Basis-momentum, which exploits the momentum return difference of two nearby contracts, yields the highest returns (16.77%). Open interest also delivers reasonable returns, suggesting that liquidity is an important state variable. In Figure 4, we plot the cumulative returns of the strategies over the sample period. It is evident that the VOL strategy delivers constantly positive returns and performs especially (but not only) well during the crisis periods.

The most interesting part of the comparison is that our strategy is the only one among the examined strategies where a major portion of return is attributed to the predictability for the spot or nearby returns. The other strategy returns including that of momentum emanate from the roll returns. For example, the large positive roll returns of momentum (24.89%) is accompanied by the negative nearby return (-12.63%). Similar cases apply for basis and basis-momentum. Open interest obtains positive returns both in the roll and nearby return part, but the nearby return component is not statistically significant.

As such, the VOL strategy provides a meaningful diversification for a portfolio which generates returns with the other commodity strategies. In Panel D, we compute the correlation matrix among the different strategies. Consistent with the finding that the VOL strategy is driven by different return component from the other strategies, the correlation is close to zero with the other strategies. Notably, the correlation with the open interest strategy is extremely low (-0.02). This evidence suggests that our results are unlikely to be liquidity-driven. The correlation among the other strategies are much higher on average. For example, the momentum strategy is highly correlated with basis (0.45), basis-momentum (0.39) and open interest (0.39).

4.3 Alternative Volatility-Related Measures

Raw implied volatility, realized volatility, variance risk premia, and skew. We compare the VOL strategy which is based on the option implied volatilities with strategies with alternative volatility measures. Specifically, we compare with the raw implied volatility, the realized volatility, the variance risk premia which is calculated as the difference between the realized variance and implied variance, and the option implied skewness. The realized volatility is computed as the square root of the sum of the squared daily returns in the current month. Similar to the implied volatility, we also detrend the realized volatility[4]. Table A.4 reports the results. The portfolio sorts based on raw implied volatility and the realized volatility yield similar results as the VOL measure. High volatility portfolios earn lower returns compared to low volatility portfolios. However, neither of them have as strong predictability as the detrended implied volatility. The portfolio sorts on the variance risk premia and option implied skewness do not yield meaningful results either.

Good and bad volatility. In order to examine whether the predictability is particularly driven the positive or negative component of the risk-neutral return distribution, we explore the two components separately in the analysis. Following Equation (12) in Kilic and Shaliastovich (2017), we decompose the implied variance into good and bad variance components iv_t^g and iv_t^b, which correspond to the prices of the positive and negative payoff components. Specifically

$$\begin{aligned} iv_t^g(\tau) &= \int_{S_t}^{\infty} \frac{2\left(1-\ln\left(\frac{X}{S_t}\right)\right)}{X^2} C_t(\tau, X)dX \\ iv_t^b(\tau) &= \int_{0}^{S_t} \frac{2\left(1+\ln\left(\frac{S_t}{X}\right)\right)}{X^2} P_t(\tau, X)dX \end{aligned}, \tag{2}$$

where $C_t(\tau, X)$ and $P_t(\tau, X)$ denote the call and put prices at time t with time to maturity τ and strike price X. The underlying price is denoted as S_t. Good volatility is constructed by the out-of-money call options that is especially sensitive when the return realization tends to be positive whereas bad volatility is made up of out-of-money put options, which are active when the asset prices drop.

[4]Alternatively, we also compute the predicted realized volatility using the HAR model as in Corsi (2009). The results are also inferior to those of the VOL measure.

Table 4 reports the results. To be consistent with the VOL measure before, we also detrend the good and bad volatility, whereas the undetrended case is similar to the results and are available upon request. The annualized return generated by the good and bad VOL strategies are 6.72% and 8.57% respectively. Though returns are generally high (low) for low (high) VOL portfolios, the magnitude and statistical significance are both inferior to those obtained by the total VOL strategy. These results show that the returns obtained by the VOL strategy is not driven by the good or bad parts and the volatility as a whole matters for the results.

Tail risk. Literature documents the importance of time-varying tail risk across the different asset classes (see, for example, Gao et al. (2017) or Gao et al. (2017)). We therefore explore the tail risk exposure across the commodities and compare the results with those obtained by VOL. Following Gao et al. (2017) we construct the tail risk measure as the difference between the prices of two different option portfolios written on the respective commodity futures. The first replication portfolio contains OTM calls and puts with a weight inversely proportional to their squared strikes.

$$IV_t^{first} = \int_{S_t}^{\infty} \frac{2}{X^2} C_t(\tau, X) dX + \int_0^{S_t} \frac{2}{X^2} P_t(\tau, X) dX \tag{3}$$

The second portfolio assigns a larger weight to the deeper out-of-money put and call options.

$$IV_t^{second} = \int_{S_t}^{\infty} \frac{2\left(1 + \ln\left(\frac{X}{S_t}\right)\right)}{X^2} C_t(\tau, X) dX + \int_0^{S_t} \frac{2\left(1 + \ln\left(\frac{S_t}{X}\right)\right)}{X^2} P_t(\tau, X) dX \tag{4}$$

The tail risk measure, which is associated with extreme movements of the futures prices, is hence computed as $Tail = IV_t^{second} - IV_t^{first}$. Due to similar reasons as with the volatilities, we detrend the tail measure by the previous 12 months' average.

Panel C of Table 4 shows that the annualized average return generated by tail risk is 11.53%, similar to but not as high as that of VOL. The correlation between the returns generated by VOL and the tail measure is 0.83, suggesting that the two types strategies are indeed closely related.

4.4 Interacting with Other State Variables / Strategies

In face of the low correlations of the VOL strategy with the other strategies, it may be interesting to see further whether the strategy performance varies across different economic states and whether combining it with other strategies can achieve better results. Table 5 shows the VOL strategy performance during the NBER recession and expansion periods. Consistent with Figure 4, the VOL strategy return is substantially higher during the NBER recessions (32.94%, t-stat=2.60), while the average return in the expansions still remains high (10.45%, t-stat=2.66). In contrast, the other strategies such as momentum or open interest show sightly lower returns during the recessions. Basis also yields higher returns during the recessions, but its return in expansions is as low as 2.97%. Therefore, it is a desirable feature of the VOL strategy that can generate higher returns in recessions while the returns outside recessions are still high.

Subsequently, we interact the strategies by implementing the independent two-way sorts. We first sort the commodities on one of the other strategies listed in the first column and split them at the median. In parallel, we sort the commodities into quartile portfolios on VOL. We then take the intersection between the two strategies. In this way, we are able to examine whether the combined strategy yields greater return.

Panel B shows that the VOL returns remain positive and statistically significant in most cases when sorting on the other variables. For example, when sorting commodities on momentum in addition, the annualized return from the VOL strategy is 13.18% for the high momentum portfolio, and 12.61% for the low momentum portfolio. The VOL returns in high and low open interest cases are similar to each other. The VOL strategy earns even slightly higher returns for the commodities with high open interest than with low open interest. These results are reassuring and show that the VOL strategy is not driven by illiquid commodities that are in practice difficult to be traded.

Sorting both on VOL and another strategy variable helps increase the portfolio returns. For instance, a strategy that is long in high basis and low VOL and short in high basis and high VOL yields an annualized return of 17.73%. Or even more interestingly, a strategy that is long in low basis and low VOL and short in high basis and high VOL can earn an annualized return of 22.93% (t-stat=3.82). Basis often act as a proxy for inventory (Fama and French, 1987). Ng and Pirrong (1994) document that volatility of the spot/nearby futures return could vary regarding the magnitude of basis. As the unconditional mean of the basis may vary across the commodities, the volatility signal combined with basis sorts out the extent to which the relative inventory is tight regarding the demand. As low basis could serve as a potential stock-out signal, low basis combined with low volatility is more likely to trigger price increase in the next month. In comparison, high basis suggests abundance of inventory and high volatility may be a sign of the selling pressure which leads to price decline in the future.

4.5 Predictive Ability of VOL for Commodity Returns

After showing the cross-sectional predictive evidence of commodity implied volatilities, it is straightforward to investigate the predictive results in the panel regressions. We run the following panel regressions:

$$R_{i,t+1} = \beta' F_{i,t} + \lambda R_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{5}$$

As $F_{i,t}$ we examine VOL, basis (as a proxy for storage conditions), the realized volatility, and the VIX (as a proxy for the general market volatility). In addition, we also control for the lagged commodity returns. In the panel regressions we include both time and commodity fixed effects to account for the fixed common components within each commodity and at each time. However, the residuals of each group at the time and commodity level may still correlate with each other. We therefore follow Petersen (2009) to compute the two-way cluster-robust standard errors at the commodity and time level as well to remove the bias in standard error estimation caused by residual correlation within these two dimensions.

Table 6 reports the results. In isolation, the coefficient of VOL is negative and statistically significant at the 1% level. One unit increase in VOL leads to a decline in return of 7.39 percentage points, which is economically large regarding the unconditional mean of the average commodity futures returns (-0.1%). In comparison, the realized volatility also leads to price decrease. However, the standard errors are relatively large, suggesting that the signal by realized volatility is noisy. Similarly, the coefficient of VIX is negative, but its magnitude and t-statistics are strictly inferior to VOL. Basis enters the equation with a positive coefficient and is not statistically significant. When including all the predictors in the joint model, the coefficient of VOL remains stable both in terms of magnitude and statistical significance.

5 Common Factors and Asset Pricing Tests

Time series tests. To understand what drives the times series returns of our VOL strategy, we regress the long-short portfolio returns on different sets of risk factors proposed in the literature. If the underlying factors can price returns effectively, the results should feature an economically small and statistically insignificant intercept, a statistically significant coefficient, and a high adjusted R^2. First, we employ the strategy factors explored in Table 2 and a global commodity excess return factor COMDT, which is the equal-weighted average returns across all the commodities examined. In Panel A of Table 7, the COMDT factor exhibits a negative significant coefficient (t-stat=-3.33), which shows that the VOL return is conversely related to the average commodity market return. In line with the results in Table 2, all the other strategy factors are not significantly related to the VOL returns.

Second, we adopt the hedge fund factors from Fung and Hsieh (2001). Specifically, these factors entail returns from the trend-following strategies applied to commodity, stock, government bond, interest rate, and foreign exchange markets. In Panel B, none of them are significant with an adjusted R^2 close to zero. Third, we look at the equity and bond market factors: the equity and bond market return, the size spread as well as the change in VIX. In Panel C, our portfolio return is not significantly correlated with equity or bond market return, nor the size spread. Interestingly, the VOL return is negatively correlated with the change in VIX. This evidence shows that our strategy return is not a hedge against the equity market uncertainty and the commodity market volatility distinguishes from the equity market volatility.

In Panel D, we employ the log return of the MSCI Emerging Market Equity Total Return Index, the log difference of the Baltic Dry Index (BDI), credit spread, and the Bloom Economic Policy Uncertainty Index. It would be interesting to examine the correlation between the strategy return and the average commodity option implied volatility. Therefore, we compute the average of the available option implied volatilities ($\overline{CIV}$) at each month. The VOL return tends to be positively related to the emerging market return, BDI index, and credit spread, and negatively related to the Bloom index. However, the magnitude of the coefficients is small and statistically insignificant. Intuitively, our strategy benefits from the growing average commodity market volatility. The coefficient of $\overline{CIV}$ is 19.34 and statistically significant at the 1% level.

As high volatility typically comes along with low liquidity, we further examine the liquidity effect in Panel E. We use shocks to the leverage of securities broker-dealer as in Adrian et al. (2014) and Etula (2013), the Pastor-Stambaugh liquidity factor (Pastor and Stambaugh, 2003), the TED spread, the funding liquidity factor (Fontaine and Garcia, 2012), the betting-against-beta (BaB) factor (Frazzini and Pedersen, 2014), and log return of the FTSE World Bank index. As the first measure is only available at the quarterly frequency, we adapt our return data also to the corresponding frequency. Our strategy return is negatively correlated with the broker-dealer index. When market liquidity is tight, shocks to the leverage of broker-dealer tend to be negative, and vice verse. Hence, the negative loading of our strategy return on this proxy suggests that the strategy benefits from the negative shock to liquidity. Similarly, when TED spread widens, signaling tightening liquidity, the VOL strategy return increases.

Cross-Sectional Tests. Our asset universe contains in total 29 portfolios: the four momentum portfolios, four basis portfolios, four basis-momentum portfolios, four open interest portfolios, four hedging pressure portfolios, four volatility portfolios, and sectoral portfolios including energy, grains, livestock, softs, and metals. We exclude the industrials portfolio because the covariance matrix is not of full rank when adding this portfolio to our asset universe. It is worth mentioning that our asset universe contains portfolios formed on broader strategies than those examined by Boons and Prado (2016), Bakshi et al. (2016), or Szymanowska et al. (2014). Thus, including portfolios sorted on many other variables improves the plausibility of the asset pricing tests (based on the critique of Lewellen et al. (2010), Prescription 1).

We adopt three approaches for the cross-sectional asset pricing tests. In addition to the well-known two-step Fama-MacBeth (FMB) approach and the GMM estimation of stochastic discount factor parameters, we also follow the three-pass approach of Giglio and Xiu (2017) taking into account of missing factors or measurement errors.

First, we employ the traditional two-step FMB approach. In this setting the excess return of commodities rx_{t+1} depends on factor risk premia λ and the individual risk quantities β_i.

$$E[\mathbf{rx}] = \lambda' \boldsymbol{\beta} \tag{6}$$

$\boldsymbol{\beta}$ is a vector of the individual beta loadings β_i. Specifically, in the first step, we estimate β_i for each portfolio by running time series regressions of portfolio excess returns on the risk factors. In the second step, we run a cross-sectional regression of the average returns on the estimated betas without an intercept, in order to obtain the factor risk premia λ. We then estimate the standard errors of λ using both the Newey and West (1987) procedure with automatic lag selection with and without Shanken (1992) correction.

Second, we implement the GMM approach to estimate the parameters for the stochastic discount factor. Following Burnside (2011), Menkhoff et al. (2012), and Bakshi et al. (2016), we denote the stochastic discount factor as m_{t+1}. No arbitrage implies:

$$E(m_{t+1}\mathbf{rx}_{t+1}) = 0, \text{ where } m_{t+1} = 1 - \mathbf{b}'(\mathbf{f}_{t+1} - \boldsymbol{\mu}). \tag{7}$$

$\mathbf{f}_{t+1}$ is a vector of risk factors and $\boldsymbol{\mu}$ is a vector of factor means. Our GMM moment conditions are:

$$g(z_{t+1}, \theta) = \begin{bmatrix} [1 - \mathbf{b}'(\mathbf{f}_{t+1} - \boldsymbol{\mu})] \otimes \mathbf{rx}_{t+1} \\ \mathbf{f}_{t+1} - \boldsymbol{\mu} \\ \mathrm{vec}(\mathbf{f}_{t+1} - \boldsymbol{\mu})(\mathbf{f}_{t+1} - \boldsymbol{\mu})' - \mathrm{vec}(\Sigma_{\mathbf{f}}) \end{bmatrix} \tag{8}$$

$\Sigma_{\mathbf{f}}$ is the factor covariance matrix. This setting incorporates the uncertainty embedded in the estimated means and covariance of $\mathbf{f}_{t+1}$ by taking them into the moment conditions.

As in Cochrane (1996) and Menkhoff et al. (2012), our (first-stage) GMM estimation uses a prespecified weighting matrix W_t, which is composed of a identity matrix for the first N asset pricing moment conditions and a large weight assigned to the additional moment conditions to ensure the precise estimation of factor means and covariance matrix.

The specification of SDF in Equation (7) implies the beta pricing model in Equation (6). The factor risk premia λ can be obtained by $\lambda = \Sigma_{\mathbf{f}}\mathbf{b}$, which connect the two models.

We compare our asset pricing model that incorporates the global commodity factor, the basis factor, the momentum factor, the basis-momentum factor, and the VOL factor with those used in Bakshi et al. (2016) and Boons and Prado (2016). Our SDF is specified as follows:

$$\begin{aligned} m_{t+1} = 1 - b_{Comdt}(Comdt_{t+1} - \mu_{Comdt}) - b_{Basis}(Basis_{t+1} - \mu_{Basis}) - b_{Mom}(Mom_{t+1} - \mu_{Mom}) \\ - b_{BasMom}(BasMom_{t+1} - \mu_{BasMom}) - b_{VOL}(VOL_{t+1} - \mu_{VOL}) \end{aligned} \tag{9}$$

Bakshi et al. (2016) employ only the first three factor whereas Boons and Prado (2016) adopt the first four factors from Equation (9).

We report in Panel A of Table 8 the GMM and FMB results for our pricing model (Model A), whereas in Panel B and C we compare with the model of Boons and Prado (2016) (Model B) and Bakshi et al. (2016) (Model C). In each panel, we report the factor risk premia λ and the t-statistics based on Newey-West standard errors with (in brackets) and without (in parentheses) Shanken correction. We also provide the OLS uncentered R^2 as well as the GLS uncentered R^2 (Lewellen et al. (2010), Prescription 3) from the FMB cross-sectional regressions. An additional advantage of the GLS uncentered R^2 is that it reflects the mean-variance efficiency of a model's factor-mimicking portfolios. In addition, the table displays the χ^2 test statistic and its p-value, as well as the Hansen-Jagannathan distance (HJ-distance). As it is not possible to compare directly whether two competing models have the same HJ-distance solely based on the particular HJ-distance value, and the p-value of it is not a good way to compare models, we apply the Kan and Robotti (2009) test[5] of the null hypothesis that the Hansen-Jagannathan distance of Model A with the VOL factor and Model B without VOL are equal. In our case, the factor used in Model B and C are a subset of our model, so the models are nested. We report the p-value of this χ^2 test.

We are primarily interested in the risk premia of the volatility factor VOL, which is the zero-cost return at $t+1$ from buying the commodity futures with the lowest detrended implied volatility at t and sell those with the highest one. In Panel A, adding VOL factor to the existing factors, the risk premia λ_{VOL} (0.011) is large in magnitude and statistically significant. The estimates of λ_{VOL} obtained by both FMB and GMM are identical. The t-statistics based on Newey-West with and without Shanken correction are similar to each other. Taking into account the fact that the factor mean and covariance are estimated simultaneously in GMM, the estimated t-statistics are aligned with those with FMB. Moreover, the factor loading b_{VOL} is also positive and statistically significant. In a nutshell, the results shows that VOL represents a distinct significant source of risk premia.

The *Comdt* factor, which is the average of all available commodity futures returns, demonstrates negative insignificant risk premia for all three models. This evidence is consistent with findings from the equity and currency market literature[6] and contrast those in Bakshi et al. (2016) and Boons and Prado (2016), which demonstrate positive (and even statistically significant) market risk premia. Remember that our asset universe contains broader strategy portfolios than those in the previous studies. To show that our estimates correspond to the actual risk premia of the factors, we follow Lewellen et al. (2010) Prescription 2 to display the average excess return of the factors in Panel D. If the risk premia are correctly identified, they should correspond to their average excess returns. Panel D confirms that it is exactly

[5]See Equation (36) and (27) in Kan and Robotti (2009).

[6]See, for example, Fama and French (1992), Lustig et al. (2011), Corte et al. (2016), Menkhoff et al. (2012).

the case: the estimated risk premia in Panel A as well as Panel B and C are almost equivalent with the average factor return, meaning that the estimates are in a reasonable range.

In all models, the estimated risk premia of basis, momentum, and basis momentum are positive and correspond to the average factor excess returns, suggesting that portfolios that are positively related to these factors earn an additional positive premium. Once the basis-momentum factor is added in Panel B, the loading on basis b_{Basis} becomes insignificant.

The VOL factor contributes positively to the incremental explanatory power of the pricing model. The uncentered GLS R^2 (uncentered OLS R^2) in Panel A is 0.84 (0.86) and is much higher compared to that in Panel B and C, whose value correspond to 0.49 (0.70) and 0.52 (0.58). Thus the pricing model with VOl stands out after the correction for the correlations among the model residuals and provides a good model fit. The χ^2 test of the null hypothesis that the pricing errors are zero is clearly rejected in Model C (p-value>0.05). In model B the p-values are equal to 0.08 (without Shanken correction) and 0.14 (with Shanken correction), indicating that it is a marginal case not to reject the null hypothesis. Finally, in model C, the p-value is as large as 0.37 or 0.53, suggesting that the null hypothesis of zero pricing error is not rejected. Model A also yields the lowest HJ-distance. Furthermore, the Kan and Robotti (2009) test rejects the null hypothesis that Model A with the VOL factor has the same HJ-distance as Model B.

Third, as a robustness check, we follow the three-pass approach of Giglio and Xiu (2017) to estimate the risk premia of the factors discussed above in case these are measured with error or the true factors are missing. As Giglio and Xiu (2017) argue, the observable factors could be just a subset of the true factors or they could be measured with error. Their approach addresses this problem in a three-pass procedure: In the first pass, principle components are used as latent factors to recover and span the factor space. In a second step, the factor loadings are estimated with OLS to obtain the risk premia of the latent factors. In a third pass, the observable factor premia are backed out based on the projection of the observable factors on the latent factors. We present the estimates of this approach in Table A.6. The number of the latent factors is determined by the information criteria following Bai and Ng (2002) which give an optimal number of six or seven. We present in the table the three cases with six, seven, or eight latent factors. We report the estimated factor premia λ of the observable factors comdt, basis, momentum, basis momentum as well as VOL. In addition, we also report the fraction of the factor variations explained by the latent factors as R_g^2. With seven latent factors, the risk premium estimated for VOL is 0.008, which is statistically significant at the 5% level. Its magnitude is slightly below but very close to the estimated premium in Table 8. The significance and magnitude of the VOL premium is lower with only six latent factors, but this is due to the fact that the first six latent factors provide very limited explanatory power for the VOL factor. The results with eight latent factors are very similar to those with seven latent factors. The other factors such as momentum or basis momentum present also marginally lower premium, whereas basis premium is slightly higher. The premium of commodity market factor *comdt* is not stable with the number of latent factors and is statistically insignificant in Table 8. In a nutshell, the estimated VOL risk premium from the three-pass method is consistent with those of FMB and GMM.

6 Understanding the VOL Measure

Having established the predictability of the VOL strategy, it is crucial to understand the information incorporated in the VOL measure and its sources of predictability. Researchers attribute risk premia in the futures returns to hedging pressure, systematic risks, or noise trader risks as a result from the financialization of commodity markets[7]. In the following

[7]For a brief summary of the related literature, see for example Adam and Fernando (2006) for the first two views and Tang and Xiong (2012) or Cheng et al. (2015) for the financialization literature.

section, we are going to examine these three aspects by employing data of the traders' positions, the aggregate commodity market risks, and the media exposure of the individual commodities. Furthermore, in face of the negative relation between VOL and the next month's return we borrow from the idiosyncratic volatility risk literature and explore the explanations with candidates related to investors' lottery preferences and market frictions.

6.1 Hedging in the Options Market and Limits to Arbitrage

In order to understand the sources of the VOL measure, the first possible explanations are the hedging pressure in the futures and options market (see Keynes (1930), Hicks (1930), or Hirschleifer (1990)) and the limits to arbitrage (Acharya et al., 2013). Keynes's theory of normal backwardation suggests that hedging pressure or the supply of futures contracts by producers tends to drive the futures prices below the expected value of the later spot prices. When hedging becomes more expensive in a certain commodity options market, selling pressure grows. The reason is that hedgers are capital-constrained and are obliged to reduce their inventory to meet the risks induced by decreasing hedging capacity. Conversely, when hedging is cheap and hedgers are better-off to hedge more of their production. They are more willing to give up part of the premium for the early resolution of uncertainty according to the theory of normal backwardation, which results in positive future returns for the long-side investors. Panel E of Table 1 shows that the highest VOL portfolio exhibits lower hedging pressure than the lowest VOL portfolio. This evidence suggests that our strategy returns can be at least partially attributed to the hedging pressure, though the portfolio sorts solely based on hedging pressure do not yield as large returns as the VOL strategy.

Why does the implied volatility of commodity options contain more information than the realized volatility? One reason might be that hedgers or speculators using commodity futures options to hedge their production or investment are better informed, so that the option implied volatility conveys much more information than the realized volatility based on futures prices.

To test this hypothesis, we back out the option traders positions from the Commitment of Traders Report (COT) from the US Commodity Futures Trading Commission (CFTC). The COT reports provide a breakdown of each Tuesday's open interest into long and short positions respectively of commercial and noncommercial traders. These data are available both for the futures only segment and the futures and options combined segment[8]. We are able to compute the options only segment data by subtracting the futures only from the futures and options combined data. We then construct the hedging demand proxy as the difference between the short and long positions scaled by the total open interest. The proxy is available for the three groups: futures only, options and futures combined, and options only. Since 2006, CFTC also publishes the Disaggregated Commitment of Traders report. It classifies traders into finer categories: producers/processors, swap dealers, managed money, etc. This classification allows us to examine much finer types of traders.

We run panel regressions with both time and commodity fixed effects to estimate the hedging pressure impact:

$$VOL_{i,t+1} = \beta \cdot HP_{i,t} + \lambda' Control_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{10}$$

where VOL is the detrended implied volatility of commodity i. HP is the hedging demand defined as the net short position of a certain type of traders divided by the sum of long and short position of that type of traders. The control variables $Control$ include the lagged dependent variable, the lagged basis, the lagged realized volatility, and the lagged VIX. Equation (10) includes both time fixed effects μ_t and commodity fixed effects α_i.

[8]Note that due to spreading (traders go long and short simultaneously both in the futures and options market, the options only positions are not always positive. However, they reflect the "net" results from the futures only and combined positions.

Correspondingly, we employ two-way cluster-robust standard errors following Petersen (2009) for the statistical significance estimates of the coefficients, which allows us to correct for correlations among standard errors both across time and commodities.

Table 9 presents the results. Panel A shows the results of the commercial traders (producers or hedgers typically) from the COT reports. They include the following categories: futures only, futures & options, and options only. The latter is backed out from the first two categories. Hedging pressure of commercial traders (hedgers) enters with a negative coefficient for all three categories. The magnitude and statistical significance of HP increase from futures only (coef=-0.35, t-stat=-0.56) to options only (coef=-1.13, t-stat=-2.38). In particular, the coefficient of -1.13 for the options only category is statistically significant at the 5% level and large in economic sense.

Panel B and C presents the results for the disaggregated COT. Panel B shows the futures only data while Panel C displays the combined data with both futures and options. Similar to the evidence with commercial traders before, producers' hedging demand is negatively related to the future VOL. Money managers, who are typically hedge funds taking the opposite side of the futures or options contracts, are the insurance providers. Their coefficient is positive significant, contrary to that of producers. Swap dealers' behavior is somehow noisy and does not have a clear impact, which is consistent with the findings in Cheng et al. (2015).

These results demonstrate that producers' hedging demand matters for commodity volatility. Higher hedging demand of commodity producers who are on the short side of the commodity markets leads to cheaper insurance costs. If speculators are not liquidity-constrained, uncertainty in the respective commodity market is low. Producers are more willing to transfer part of the premium to speculators to secure the future prices and resolve the uncertainty associated with the future price trajectory. Subsequently future prices increase. On the contrary, when fewer producers are willing to hedge their production on the futures or options markets or speculators' capital is constrained, hedging is more expensive. This evidence becomes even stronger in the options market. In order to resolve uncertainty and reduce storage costs, producers are forced to sell part of their inventory on the spot market to avoid further loss, which causes prices to decline.

6.2 The Aggregate Commodity Market Risk

The VOL strategy return may be related to investor's aversion to the aggregate commodity market risk. To test this hypothesis, we form the average of the detrended implied volatilities across commodities and construct an aggregate volatility index $AVOL$. We then estimate the time-varying beta loadings with rolling windows of 36 months as follows: in each month t we regress the commodity excess returns rx_i on $AVOL$ from $t-35$ to t and obtain the coefficient $\beta_{i,t}$. Subsequently, we sort beta from high to low and compute the average return of portfolios at month $t+1$. The strategy then buys the portfolio with the lowest beta and sells the one with the highest beta.

Table 10 reports the results. Similar to the VOL strategy, high (low) volatility beta portfolios earn low (high) future returns. The return is mainly attributed to the predictability of AVOL for the future spot returns. The beta loading before and after portfolio formation is relatively stable. However, the AVOL strategy earns an excess return of 9.37% annually, which is inferior to the VOL strategy. The t-statistics of 2.12 is significant at 5%.

Moreover, the correlation between the returns from the VOL strategy and the AVOL strategy is as low as 0.055. Therefore, these results do not support the hypothesis that the VOL strategy returns emanate from the compensation for the aggregate commodity market volatility.

6.3 Uncertainty and Media Coverage

Another question we are interested in is: Does the VOL measure reflect uncertainty proxied by the news media? Uncertainty proxied by news media has been addressed by several studies so far. The most well-known among them is Baker, Bloom, and Davis (2016), who construct an economic policy uncertainty index based on news paper coverage frequency. This media-based uncertainty index captures the uncertainty at the macro and micro level. Similarly, Plante (2015) identifies OPEC-related oil price uncertainty based on newspaper article count. They find that positive innovations in article count lead to significant increases in oil price uncertainty. If the VOL variable can be interpreted as the uncertainty embedded in the commodity futures contracts, then media should imposes similar effect or help amplify the uncertainty impact.

We construct each commodity's media exposure proxy based on the Ravenpack 4.0 database. Ravenpack collects and identifies news information on entities such as traded commodities. The underlying sources include the *Dow Jones Financial Wire*, *Barron's*, and *The Wall Street Journal* and web news. We are able to identify 149536 news articles related to the 25 commodities examined in our sample[9]. The available period spans from January 2000 to October 2014.

Commodity's news coverage might be endogenously correlated with several features. We therefore follow Hong et al. (2000) and Hillert et al. (2014) to quantify a residual media coverage (RMC) measure. Our cross-sectional regression to obtain RMC is specified as follows:

$$\#news_{i,t} = \alpha + \beta_1 Open_{i,t} + \beta_2 DJ_{i,t} + \beta_3 GSCI_{i,t} + \beta_3 \#news_{i,t-1} + \varepsilon_{i,t}, \tag{11}$$

[9] Wheat(W-) and Kansas Wheat(KW) in our sample share the same news measure due to similar underlying commodities. Similarly, the same applies to Gasoline Blendstock(RB) and Gasoline unleaded (HU).

where $Open_{i,t}$ is the open interest of the commodity i in month t. DJ is the a dummy variable for the membership in the Bloomberg Commodity Index (former DJ UBS index) and GSCI is a dummy for the membership in the S&P GSCI index. We also control for the news lag.

We run the cross-sectional regressions specified in Equation (11) separately for each month of the formation period (12 months) and then compute the average residuals obtained in the 12 cross-sectional regressions from $t-11$ to t. At the end of each period t, we form four portfolios based on RMC. We assign the 25% of all commodities with the highest RMC to Portfolio 1, and the the 25% of all commodities with the lowest RMC to Portfolio 4. We then compute the equal weighted excess return for each portfolio at $t+1$. The strategy is short in Portfolio 1 (high RMC commodities) and long in Portfolio 4 (low RMC commodities).

Panel A of Table 11 reports the returns and descriptive statistics of the portfolios sorted by RMC and the long-short portfolio that is short in the portfolios with the highest RMC and long in portfolio with the lowest RMC. Interestingly, similar to the portfolios sorted by the VOL measure, the returns of portfolios sorted on RMC increase monotonously with the declining residual media coverage. These results suggest that high uncertainty proxied by residual media coverage leads to future price drop. However, the correlation between the RMC strategy return and the VOL strategy return is as low as 0.03, suggesting the two proxies are of different characteristics.

Next, we formally test the predictability of media-based uncertainty for VOL. Panel B presents the results from panel regressions with both time and commodity fixed effects. The left hand side variable is the VOL variable, whereas the right hand side variables are the lagged RMC, Basis, the realized volatility, the VIX, or all the aforementioned variables. For control we also include the lagged VOL. The cluster robust errors are estimated both at the time and commodity dimension. Consistent with the low correlation between the strategy returns, RMC is only weakly positive related to VOL. Its coefficient is not statistically significant, when we observe it alone or control for the other variables.

Due to the low correlations between the two strategies, it may be interesting to examine whether we can obtain higher returns by combining the two strategies. Panel C displays the results from the independent double sort by both RMC and VOL and the summary statistics of the returns from the strategy that is short in high RMC and high VOl commodities and long in low VOL and low RMC commodities. Combining the two strategies indeed increases returns: the annualized average return of the zero-cost strategy is 20.70%, which is economically large and statistically significant at the 1% level.

6.4 Lottery Preferences and Market Frictions

Besides the candidates examined above, we are further interested in the question whether variables related to lottery preferences of investors or market frictions can explain the predictive ability of VOL for futures commodity returns. As a first group of candidate variables, we consider lottery preferences of investors. Barberis and Huang (2001) argue that investors may prefer assets with high skewness in the last period according to the cumulative prospect theory, which results in overpricing and lower returns in the subsequent period. As proxies for the higher return in the last period, we adopt the option-implied skewness of the respective commodities (Skew), the maximum daily return (Maxret), and the futures basis defined in Section 3 (Basis). Boyer et al. (2010) provide evidence that skewness helps explain the fact that stocks with high idiosyncratic volatility have low expected returns. Maxret is proposed by Bali et al. (2001) and is constructed as the maximum daily return in month $t-1$. As futures basis also serves as a noisy expectation of future price changes, we also utilize this measure.

As variables for market frictions, we examine the one-month return reversal effect and (il)liquidity. Huang et al. (2010) argue that the positive return and idiosyncratic volatility relation can be almost entirely explained by the one-month return reversal effect. We employ respective commodity return in month $t-1$ as a proxy for one-month return reversal (Lagret). The positive return–volatility relation could also be a result of illiquidity. As (il)liquidity measure we use the open interest (open) defined in Section 3 and the Amihud (2002) illiquidity measure (Amihud). The Amihud measure is computed as the average of the daily absolute return divided by the daily dollar trading volume in month $t-1$.

We then decompose the predictive ability of VOL for the commodity futures returns using individual commodity level Fama-Macbeth cross-sectional regressions, following the four-step procedure of Hou and Loh (2016). In the first step, we regress the cross section of individual commodity returns on their month $t-1$ VOL:

$$r_{i,t} = \alpha_t + \beta_t VOL_{i,t-1} + \epsilon_{i,t}. \tag{12}$$

In a second step, we add a candidate variable $Candi_{t-1}$ to the regression to show whether the VOL effect still persists after controlling for the candidate variable. In a third step, we regresses VOL_{t-1} on $Candi_{t-1}$:

$$VOL_{i,t-1} = c_{t-1} + \delta_{t-1} Candi_{i,t-1} + \varepsilon_{i,t-1} \tag{13}$$

The purpose of this step is to decompose $VOL_{i,t-1}$ into two orthogonal components: $\delta_{t-1} Candi_{i,t-1}$ and $(c_{t-1} + \varepsilon_{i,t-1})$. Finally in Step 4, β_t from Step 1 is decomposed as:

$$\beta_t = \frac{cov(r_{i,t}, VOL_{i,t-1})}{Var(VOL_{i,t-1})} = \frac{cov(r_{i,t}, \delta_{t-1} Candi_{i,t-1})}{Var(VOL_{i,t-1})} + \frac{cov(r_{i,t}, (c_{t-1} + \varepsilon_{i,t-1}))}{Var(VOL_{i,t-1})} = \beta_t^C + \beta_t^R. \tag{14}$$

The fraction of the negative VOL return relation explained by the candidate variable is measured by $E(\beta_t^C/\beta_t)$, and the fraction unexplained by the candidate is measured by $E(\beta_t^R/\beta_t)$. The statistical approximations of the mean and variance of the ratios are given by:

$$E(\beta_t^C/\beta_t) \approx E(\beta_t^C)/E(\beta_t) \tag{15}$$

$$Var(\beta_t^C/\beta_t) \approx \left(\frac{E(\beta_t^C)}{E(\beta_t)}\right)^2 \times \left(\frac{Var(\beta_t^C)}{(E(\beta_t^C))^2} + \frac{Var(\beta_t)}{(E(\beta_t))^2} - 2\frac{Cov(\beta_t^C,\beta_t)}{E(\beta_t^C)E(\beta_t)}\right) \tag{16}$$

In addition to adding one candidate variable each time, this approach can also be applied to the multiple-candidate case. We examine both situations. Furthermore, it is worth mentioning that a candidate variable that is highly correlated with VOL does not necessarily explain a high fraction of the predictive ability of VOL, because the highly correlated part with VOL may not be the component that drives the predictive ability of VOL.

Table 12 presents the results. In all the steps of each candidate case we only include those commodity-month observations where the respective candidate observations are available. Therefore, the average number of commodities per month in the bottom line varies. Panel A reports the results including one candidate each time. In each of the six cases in Step 1, the VOL coefficient is -0.06 on average and statistically significant at the 5% level. The results are robust after including the candidate variables (Step 2).

In the third step we decompose VOL into a part explained by the candidate variable and an orthogonal part by regressing VOL on the candidate variable each month. Among the lottery preference variables, VOL is positively related to $Skew$. The coefficient is 0.7% with t-statistics of 3.42. This evidence confirms the findings in Boyer et al. (2010) that high volatility is positively correlated with high skew, which leads to overpricing and lower future returns. However, the adjusted R^2 is relatively low (1.7%). VOL is positively related to $Maxret$ with adjusted R^2 of 13.5%. As $Maxret$ is essentially a range-based measure of volatility, the high correlation with VOL is not surprising. VOl is negatively correlated with *Basis*, suggesting that the detrended volatility tends to be low when the futures curve is in contango (upward-sloping).

In the fourth step, we decompose the VOL coefficient β_t into a component that is related to the candidate variable β_t^C and a residual component β_t^R following Equation (14). Subsequently, we can calculate the fraction of β_t attributable to the candidate variable by $E(\beta_t^C/\beta_t)$. The time series average of β_t^C of *Skew* is -0.6% with t-stat of 1.29, and the explained fraction of the coefficient is 11%. As a range-based measure of volatility, $Maxret$ explains 35% of the VOL coefficient. Though *Basis* is weakly related to VOL, the explained proportion is closed to zero.

Among the market friction variables, VOL is neither significantly related to *Lagret* or *Open*. The coefficient of both variables in Step 3 is 0.4% (t-stat=0.18) and 1.6% (t-stat=0.81) respectively. The explained fraction of β_t in Step 4 is 14% and 4%. The coefficient of the Amihud-illiquidity measure is 0.019 and is statistically significant at the 1% level, showing that high VOL is associated with high illiquidity. However, in Step 2 when controlling for *Amihud*, the VOL coefficient still remains significant at the 5% level, meaning that a substantial fraction of it is unexplained by this measure, whereas the coefficient of *Amihud* is statistically insignificant from zero. The portion of the Stage 1 coefficient of VOL explained by *Amihud* is 26%, which is statistically significant at the 5% level. A large fraction of the VOL coefficient (74%) is unexplained by this illiquidity measure.

After examining each of the candidate variables, we next turn to multivariate analysis by combining lottery preferences, market frictions, or all candidates together. The purpose of this exercise is to find out the total fraction these two categories can explain and the marginal contribution of each candidate variable. In Panel B of Table 12, the first three columns examine the lottery preference candidates. In total, the lottery preference candidates explain

33% of the VOL coefficient. The residual component accounts for 67% (t-stat=4.39) of the VOL coefficient. *Skew* and $Maxret$ explain around 7% (t-stat=0.8) and 36% (t-stat=2.49) of the return-VOL relation. Column 4 – 6 report the results with market frictions candidates. *Lagret* and *Amihud* explain 10% and 26% of the VOL coefficient. Market friction candidates make up together 38% of the relation, whereas the rest as high as 72% (t-stat=5.64) is unexplained.

The last three columns display results by including all the candidates discussed above. In Step 2, after controlling for all the candidates, the coefficient of VOL remains robust. The coefficient of *Lagret* is -0.05 (t-stat=2.41) and is statistically significant at the 5% level. However, the forecasting ability of VOL does not mainly stem from the short-term mean reversal channel. When regressing VOL on all candidates in Step 3, the correlation with *Lagret* is small and statistically insignificant. In Step 4 of decomposition, the marginal contribution of it is merely 3%. Similarly, after controlling for the other candidates, the marginal contribution of *Amihud* is 9% and is statistically insignificant. $Maxret$ is the strongest candidate and accounts for 34% of the return–VOL relation. Again, it is essentially a ranged-based measure of volatility. In total, 52% of the relation is explained by the candidate variables, whereas the remaining 48% (t-stat=3.82) is unexplained.

In addition our approach can also evaluate the forecasting ability of the candidate variable for commodity returns after controlling for VOL. For example, *Amihud* explains 26% of the VOL coefficient, but it has no independent predictive power for commodity returns after controlling for VOL. The same applies to $Maxret$.

To sum up, lottery preference and market friction variables show some promise in explaining the predictive ability of VOL for the cross-section of commodity returns. This evidence suggests that some of the predictive ability come from these sources. However, these candidates still leave a sizable portion of the return–VOL relation unexplained.

7 Predictive Ability for Exchange Rate Returns

Is the aggregate commodity VOL factor able to price other assets? To answer this question, we examine the FX returns. We follow Menkhoff et al. (2012) to construct the currency sample, which contains 48 currencies from January 1990 to November 2014. We use s to denote the log of the spot exchange rate in units of foreign currency per US dollar, and f for the log of the forward exchange rate, also in units of foreign currency per US dollar. Monthly excess returns for holding foreign currency k are computed as

$$rx_{t+1}^{k} \equiv i_t^k - i_t - \Delta s_{t+1}^{k} \approx f_t^k - \Delta s_{t+1}^{k} \tag{17}$$

In Equation (17), the excess returns can be decomposed into the spot return Δs_{t+1}^{k} and the forward discount $f_t^k = i_t^k - i_t$.

Our strategy is constructed as follows: At each month t, we regress the excess returns of currency k from $t-35$ to t on the average commodity VOL of the same period to obtain their beta loading $\beta_{k,t}$. We then sort currencies into five portfolios depending on $\beta_{k,t}$ and calculate their equal weighted portfolio returns at $t+1$. Thus, we use rolling estimates of beta with a rolling window of 36 months (as in Lustig et al. (2011)) and rebalance portfolios at the end of each month. We find commodities VOL, especially that of precious metals, metals and industrial materials are able to generate significant FX returns, therefore, we apply the average VOL of these commodities in our strategy. We first construct the mean within each of the three sub-categories before taking the average of the category means. As the time series contains some noises, we smooth it by taking three months' average similar to Bakshi and Panayotov (2013).

Descriptive statistics for portfolio excess returns are shown in Table 12. The commodity VOL is able to predict currency returns. The strategy that is short in the highest commodity VOL beta and long in the lowest commodity VOL beta yields a significant annualized return of 4.77% (t-stat=2.92). Our strategy return decomposition distinguishes from that of the carry trade strategy (Lustig et al., 2011), for which the return is mainly attributed to the forward discount component, and from the variance risk premia strategy (Corte et al., 2016) where the spot return drives the result: In our strategy both the forward discount and spot return component are comparatively large and statistically significant. Though the spot return component is large in magnitude (3.32, t-stat=2.00), the forward discount component is also very stable (1.45, t-stat=4.75) over time. This evidence suggests that the strategy return can benefit both from the predictive ability of commodity VOL and the interest rate difference. In addition, the commodity VOL strategy has low correlations with the carry trade strategy (0.21), providing incremental contribution to the current existing strategies.

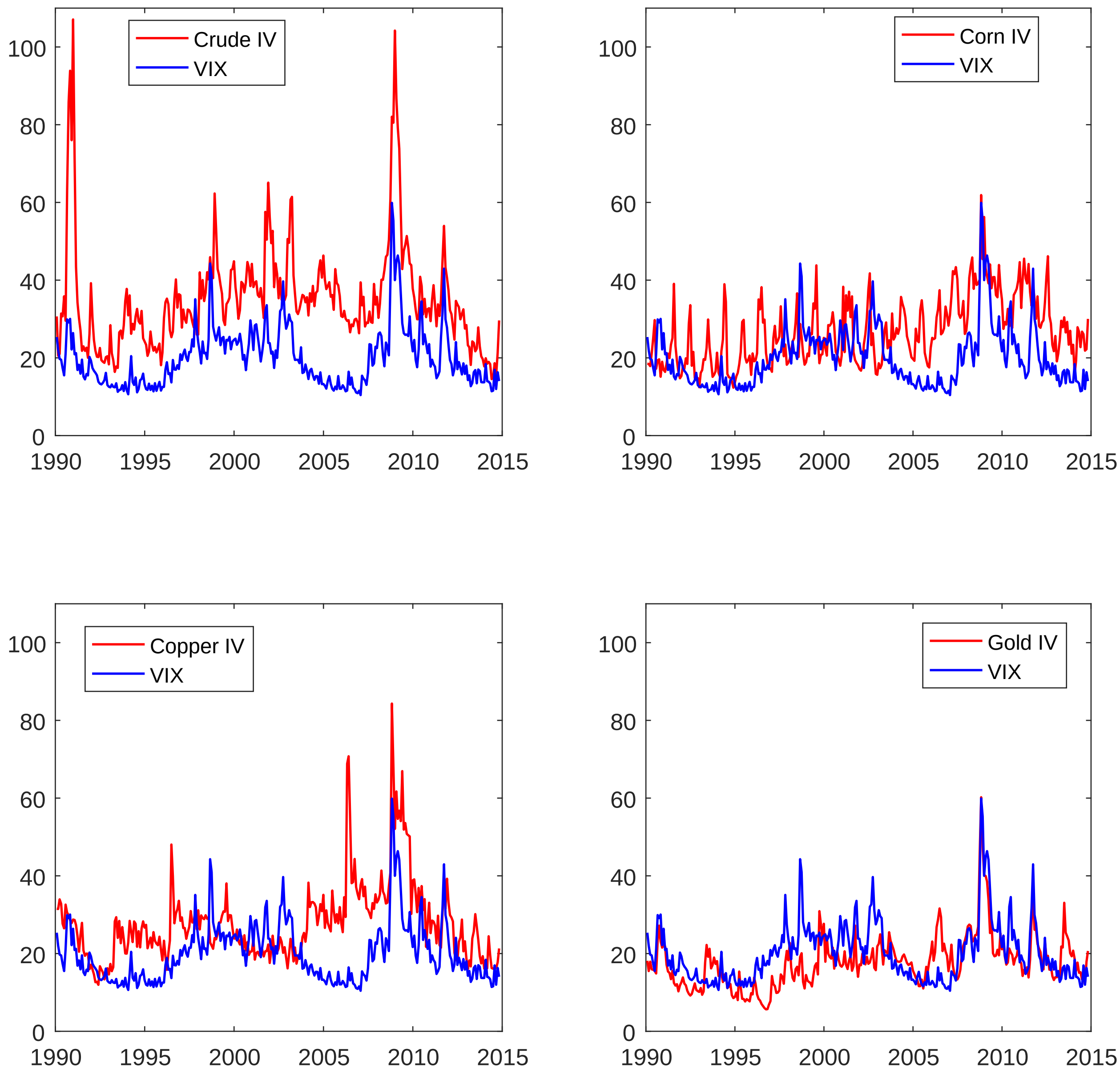
8 Conclusion

This paper uncovers the significant predictive ability of the commodity option implied volatility for the cross-sectional commodity returns. Commodities with high implied volatility tend to decrease in prices in the next period, whereas those with low volatility increase. A strategy that is long in commodities with low volatility and short in those with high volatility earns substantial returns in the next month. The predictive power of the option implied volatility is superior to alternative measures such as the realized volatility or option implied skewness. The asset pricing tests show that the volatility factor contributes to the model fit and represents a distinct source of premium.

What distinguishes our strategy from the existing commodity strategies is that our strategy returns emanate almost entirely from the predictive power for the nearby return component, whereas the other strategies mainly earn positive returns from the roll return component. Therefore, our strategy demonstrates low correlations with the other strategies and provide substantial diversification potential. This fact becomes highly desirable especially during the recessions.

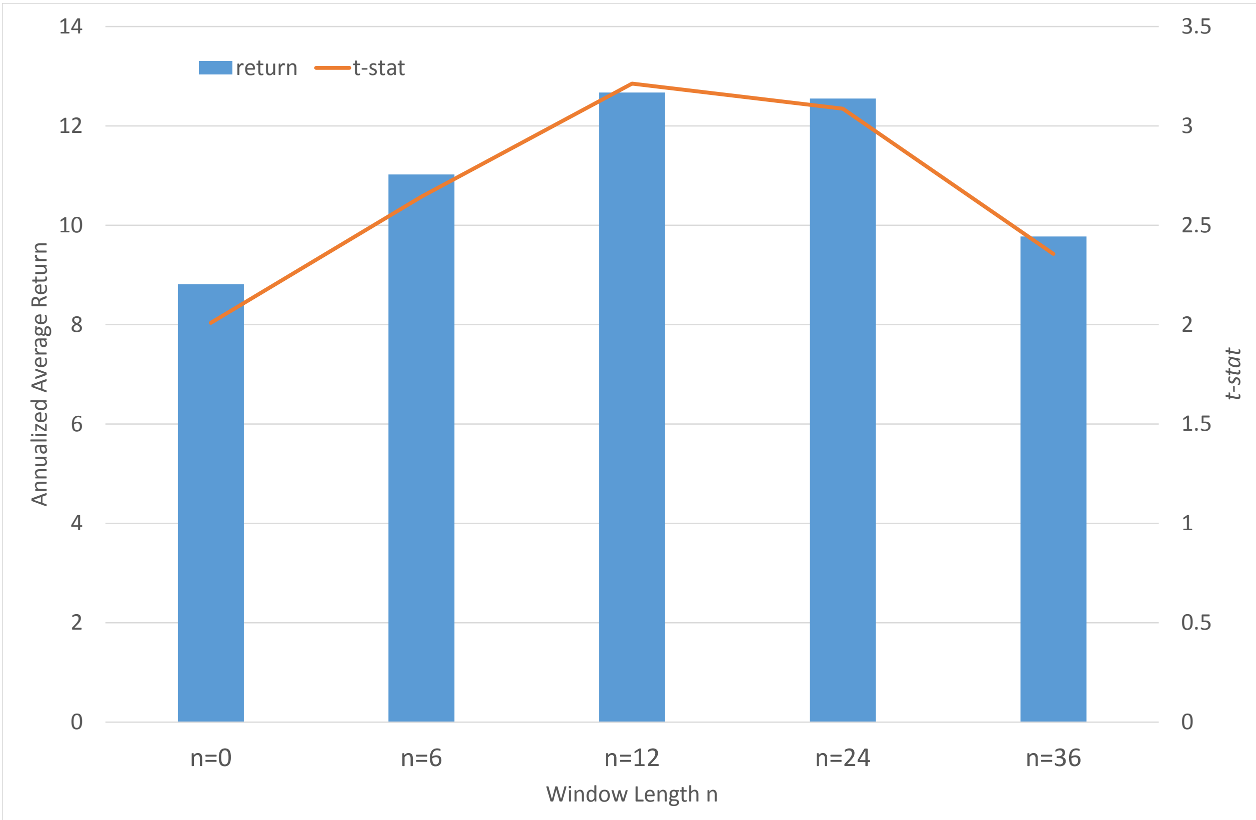
We show that our strategy is not related to the risk factors associated with other strategies or business cycle risks. The detrended implied volatility of commodities is related to the hedging pressure and the limits to arbitrage on the commodity markets. When few producers are willing to hedge their production on the futures or options market, commodity insurance becomes more expensive, which leads to a decline in the future spot prices due to the selling pressure from producers. We further show that the uncertainty proxied by media coverage can generate similar patterns to the implied volatility. However, they represent difference source of premium. Neither is the premium related to the aggregate commodity market risk. The predictive ability of VOL is associated with lottery preferences of investors and market frictions. However, these variables cannot fully explain the return–VOL relation either.

Figure 1: **Commodity Implied Volatilities vs VIX**



The figure plots the implied volatilites of crude oil, corn, copper, and gold (red line) versus the VIX index (blue line). The commodity implied volatilities are computed based on the method in Bakshi et al. (2003). The data are monthly. The sample period spans from January 1990 to October 2014.

Figure 2: **Sensitivity of Strategy Return to the Choice of n**



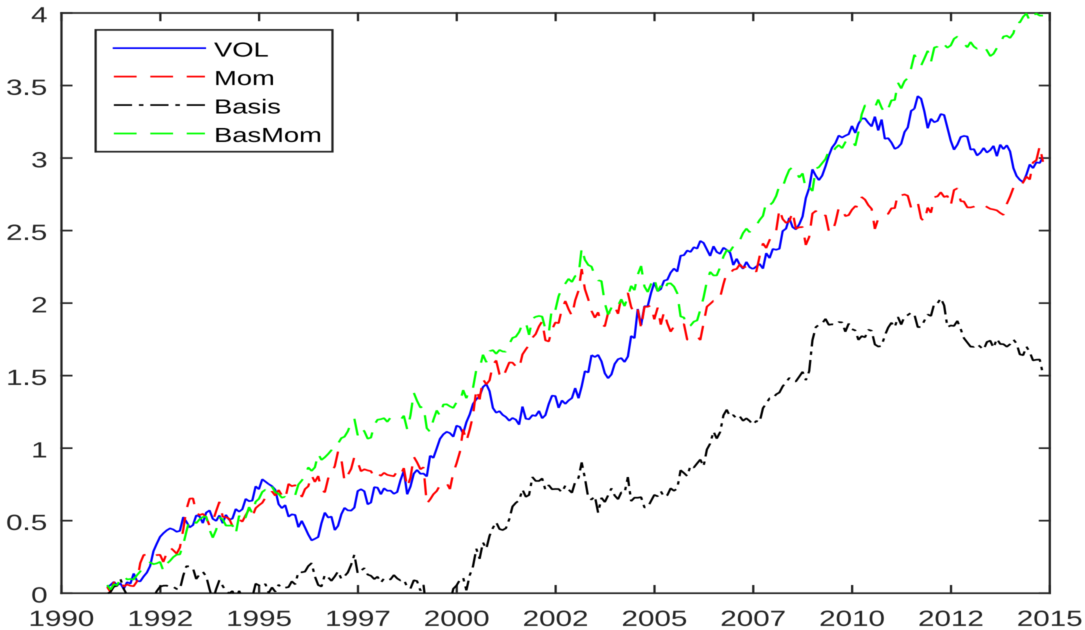
The figure plots the annualized strategy returns with regard to the choice of the window length n as in Equation 1. The blue bar denotes the annualized average strategy returns sorted on the detrended implied volatility. The detrended implied volatility is computed as the commodity implied volatility in month t, subtracted by the average volatility of window n up to month $t-1$. The orange line denotes the corresponding t-stat of the strategy returns.

Figure 3: **Portfolio Formation Period Returns**

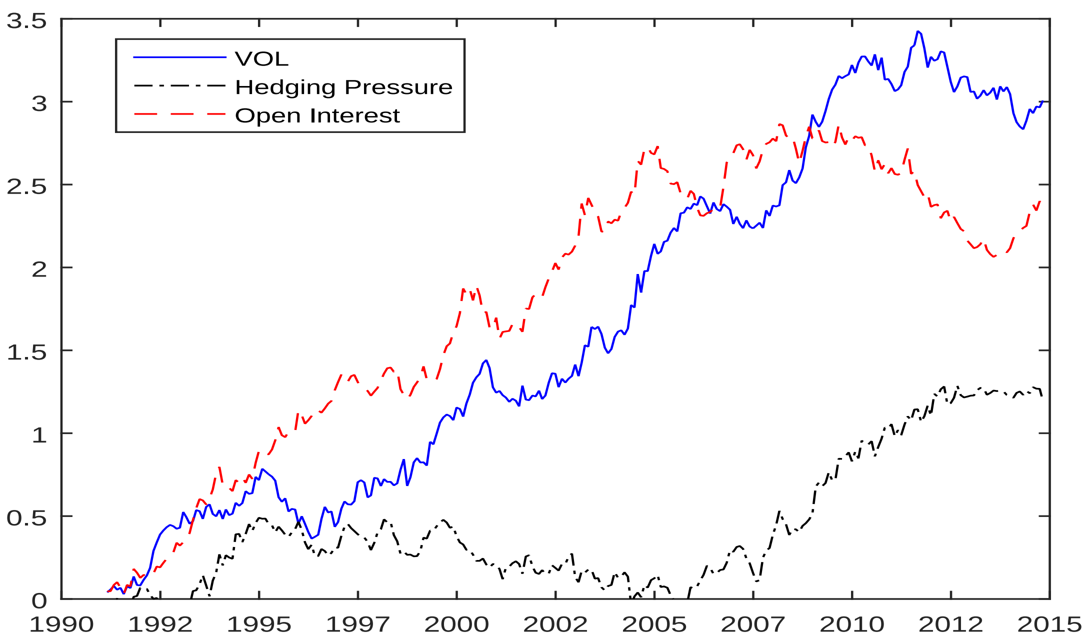
The figure presents the average annualized 1-month holding period returns. We sort the detrended implied volatility at t and presents the average 1-month holding period returns for month $t+1$, or $t+2$,...,$t+12$. Returns from one month up to one year after the portfolio formation are presented. Panel A displays the excess return, nearby returns, and roll returns of the VOL strategy which goes long in the lowest VOL portfolio and short in the highest VOL portfolio. Panel B shows the long and short leg of the strategy over the different holding periods. Panel C and D decompose the long and short leg returns respectively into nearby returns and roll returns. The returns are constructed as the long-short portfolios from a sort of 25 commodities on each of the respective signals. The sample period is from January 1990 to October 2014, so the first return is observed in February 1991.

Figure 4: **Cumulative Returns from Portfolio Sorting Strategies**

VOL vs. Mom vs. Basis vs. BasMom



VOL vs. Hedging Pressure vs. Open Interest



The figure shows the cumulative excess returns based on the strategies: volatility, momentum, basis, basis-momentum, hedging pressure, and open interest. The returns are constructed as the long-short portfolios from a sort of 25 commodities on each of the respective signals. The sample is monthly and spans from January 1990 to October 2014, so the first return is observed in February 1991.

Table 1: **Portfolios Sorted by Detrended Implied Volatility**

	High	2	3	Low	Low−High
Panel A: Excess Returns					
Mean	-10.08	0.66	-0.65	2.58	12.66
t-stat	(-2.50)	(0.21)	(-0.22)	(0.83)	(3.34)
Sdev	19.68	15.49	14.62	15.14	18.48
Skew	-0.57	-0.43	-0.31	-0.29	0.04
Kurtosis	5.58	4.62	5.66	4.32	3.28
Sharpe ratio	-0.51	0.04	-0.04	0.17	0.69
AC(1)	0.05	0.07	0.02	-0.05	-0.06
Panel B: Nearby Returns					
Mean	-5.66	5.50	4.12	8.77	14.42
t-stat	(-1.37)	(1.68)	(1.37)	(2.78)	(3.67)
Sdev	20.07	15.94	14.64	15.36	19.13
Skew	-0.58	-0.29	-0.24	-0.35	0.18
Kurtosis	5.43	4.37	4.99	4.27	3.85
Sharpe ratio	-0.28	0.34	0.28	0.57	0.75
AC(1)	-0.01	0.06	0.00	-0.09	-0.09
Panel C: Roll Returns					
Mean	-4.41	-4.80	-4.78	-6.20	-1.80
t-stat	(-2.92)	(-5.70)	(-6.89)	(-7.63)	(-1.17)
Sdev	7.35	4.10	3.38	3.96	7.52
Skew	0.31	-0.10	0.18	-0.09	-0.39
Kurtosis	5.65	4.26	3.95	5.23	5.39
Sharpe ratio	-0.60	-1.17	-1.41	-1.57	-0.24
AC(1)	0.33	0.18	0.17	0.03	0.19
Panel D: Transition Matrix					
P_{high}	0.57	0.26	0.13	0.05	
P_2	0.23	0.39	0.27	0.11	
P_3	0.12	0.26	0.39	0.23	
P_{low}	0.06	0.11	0.23	0.60	
$\overline{\pi}$	*0.25*	*0.25*	*0.25*	*0.25*	

Panel E: Characteristics before and after Portfolio Formation					
	High	2	3	Low	Low−High
pre-VOL	0.08	0.01	-0.02	-0.07	-0.15
	(18.67)	(4.26)	(-10.80)	(-22.66)	(-35.47)
post-VOL	0.05	0.01	-0.01	-0.05	-0.10
	(13.12)	(2.82)	(-5.23)	(-14.78)	(-25.55)
pre-MOM $\times 10^2$	-0.06	-0.03	-0.17	-0.37	-0.30
	(-0.39)	(-0.27)	(-1.80)	(-3.61)	(-2.34)
post-MOM $\times 10^2$	-0.12	-0.00	-0.15	-0.33	-0.21
	(-0.76)	(-0.02)	(-1.64)	(-3.26)	(-1.62)
pre-Basis $\times 10^2$	0.26	0.29	0.36	0.54	0.29
	(2.17)	(5.13)	(7.06)	(9.43)	(2.44)
post-Basis $\times 10^2$	0.33	0.29	0.34	0.50	0.18
	(2.91)	(5.08)	(6.29)	(8.50)	(1.58)
pre-BasMom	-0.02	-0.02	-0.02	-0.03	-0.00
	(-7.75)	(-7.28)	(-7.92)	(-8.88)	(-1.46)
post-BasMom	-0.02	-0.02	-0.02	-0.03	-0.00
	(-7.76)	(-7.21)	(-7.86)	(-8.88)	(-0.89)
pre-HP	-0.12	-0.10	-0.09	-0.09	0.03
	(-15.16)	(-15.78)	(-13.54)	(-14.34)	(3.38)
post-HP	-0.12	-0.10	-0.09	-0.10	0.02
	(-15.64)	(-15.46)	(-12.88)	(-13.82)	(2.25)
pre-OI$\times 10^{10}$	0.21	0.18	0.20	0.23	0.02
	(9.53)	(9.78)	(8.55)	(8.42)	(0.74)
post-OI$\times 10^{10}$	0.21	0.18	0.21	0.24	0.03
	(9.53)	(9.79)	(8.66)	(8.42)	(1.35)

This table presents descriptive statistics of commodity portfolios sorted on the implied volatility at $t-1$ detrended by the previous 12 months mean of implied volatility $\frac{1}{12}\sum_{i=t-13}^{t-2}\sigma_i$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) volatilities. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1) and the annualized Sharpe ratio. Panel A displays the overall excess return, whereas Panel B and Panel C report the nearby return and the roll return component only. Pane D presents the transition probability from portfolio i to portfolio j between time t and time $t+1$. $\overline{\pi}$ indicates the steady state probability. Panel E summarizes the characteristics of the portfolios before $(t-1)$ and after portfolio formation (t). *MOM* is the past 12-month futures returns. *Basis* is the difference between the second nearby and the nearby futures prices adjusted by the maturity difference. *BasMom* is the difference between momentum in the first- and second-nearby futures. HP is the hedging pressure computed as the net short positions of producers divided by the total positions of producers. OI is the dollar value of open interest which is the number of contracts outstanding times the contract size. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table 2: **Commodity Portfolio Strategies**

	VOL	MOM	Basis	BasMom	Hedging Pressure	Open Interest
Panel A: Excess Returns						
Mean	12.66	12.52	6.35	16.77	5.02	9.94
t-stat	(3.34)	(2.88)	(1.75)	(4.63)	(1.56)	(2.66)
Sdev	18.48	21.19	17.70	17.66	15.73	18.22
Skew	0.04	-0.22	-0.13	-0.09	0.20	0.18
Kurtosis	3.28	3.83	4.34	4.18	4.28	4.08
Sharpe ratio	0.69	0.59	0.36	0.95	0.32	0.55
AC(1)	-0.06	-0.05	-0.06	-0.05	-0.01	-0.00
Panel B: Nearby Returns						
Mean	14.42	-12.63	-45.05	-5.29	1.72	4.68
t-stat	(3.67)	(-2.78)	(-11.21)	(-1.33)	(0.53)	(1.23)
Sdev	19.13	22.13	19.59	19.39	15.93	18.52
Skew	0.18	-0.41	-0.28	-0.23	0.08	0.16
Kurtosis	3.85	4.01	4.67	4.15	3.98	4.21
Sharpe ratio	0.75	-0.57	-2.30	-0.27	0.11	0.25
AC(1)	-0.09	-0.03	0.00	-0.08	-0.06	-0.08
Panel C: Roll Returns						
Mean	-1.80	24.89	51.41	21.91	3.49	5.27
t-stat	(-1.17)	(17.30)	(34.87)	(14.74)	(3.09)	(4.09)
Sdev	7.52	7.01	7.18	7.24	5.49	6.28
Skew	-0.39	0.72	0.55	0.55	0.20	0.41
Kurtosis	5.39	3.98	2.93	3.39	4.08	4.55
Sharpe ratio	-0.24	3.55	7.16	3.02	0.63	0.84
AC(1)	0.19	0.06	-0.06	0.06	0.15	0.06
Panel D: Correlations						
VOL	1.00	-0.06	-0.02	0.02	-0.07	-0.02
MOM	-0.06	1.00	0.45	0.39	0.26	0.39
Basis	-0.02	0.45	1.00	0.38	0.15	0.10
BasMom	0.02	0.39	0.38	1.00	0.18	0.17
Hedging Pressure	-0.07	0.26	0.15	0.18	1.00	0.35
Open Interest	-0.02	0.39	0.10	0.17	0.35	1.00

This table presents descriptive statistics of commodity strategies formed using time $t-1$ information. *VOL* is the strategy that buys (sells) the top quartile of all commodities with the lowest (highest) detrended implied volatilities. Similarly, *MOM* is the momentum strategy that buys (sells) commodities with the highest (lowest) past 12-month futures returns. *Basis* is the carry strategy that buys (sells) commodities with the lowest (highest) basis. *BasMom* is the strategy that buys (sells) commodities with the highest (lowest) basis-momentum. *Hedging Pressure* is the strategy that buys (sells) commodities with the highest (lowest) hedging pressure. *Open Interest* is the strategy that buys (sells) commodities with the highest (lowest) open interest growth. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The table also reports first order autocorrelation coefficient (AC). Panel A displays the overall commodity excess return, whereas Panel B and Panel C report the nearby return and the roll return component only. Panel D presents the sample correlations of the commodity excess returns among the strategies. The strategies are rebalanced monthly from February 1991 to October 2014.

Table 3: **Static Commodity Portfolio Strategies**

Panel A: Excess Returns						
	$\overline{VOL}$	$\overline{MOM}$	$\overline{Basis}$	$\overline{BasMom}$	$\overline{\text{Hedging Pressure}}$	$\overline{\text{Open Interest}}$
Mean	6.10	17.20	15.85	16.08	3.47	4.58
t-stat	(1.26)	(4.11)	(4.67)	(4.65)	(1.03)	(1.20)
Sdev	23.57	20.39	16.56	16.86	16.41	18.66
Skew	-0.06	0.02	0.33	0.35	-0.14	0.05
Kurtosis	3.58	2.92	4.02	3.98	3.24	3.22
Sharpe ratio	0.26	0.84	0.96	0.95	0.21	0.25
AC(1)	0.03	0.07	0.13	0.06	0.09	0.06
Panel B: Nearby Returns						
Mean	0.53	2.64	1.35	2.39	1.15	1.54
t-stat	(0.11)	(0.58)	(0.36)	(0.63)	(0.34)	(0.39)
Sdev	24.45	22.11	18.08	18.36	16.61	18.98
Skew	-0.13	-0.11	-0.01	0.11	-0.22	-0.03
Kurtosis	3.40	3.15	3.54	3.54	3.13	3.22
Sharpe ratio	0.02	0.12	0.07	0.13	0.07	0.08
AC(1)	0.03	0.12	0.12	0.06	0.01	0.01
Panel C: Roll Returns						
Mean	5.50	15.15	14.32	13.74	2.32	3.04
t-stat	(4.55)	(10.25)	(9.87)	(9.10)	(1.79)	(2.55)
Sdev	5.89	7.20	7.07	7.36	6.30	5.81
Skew	-0.08	0.20	0.02	-0.15	0.16	0.02
Kurtosis	4.07	3.00	2.66	2.85	3.47	5.01
Sharpe ratio	0.93	2.10	2.03	1.87	0.37	0.52
AC(1)	0.26	0.35	0.16	0.20	0.05	0.17
Panel D: Correlations						
VOL	1.00	0.67	0.41	0.40	-0.48	0.49
MOM	0.67	1.00	0.67	0.78	-0.24	0.41
Basis	0.41	0.67	1.00	0.72	0.15	0.03
BasMom	0.40	0.78	0.72	1.00	0.26	0.03
Hedging Pressure	-0.48	-0.24	0.15	0.26	1.00	-0.66
Open Interest	0.49	0.41	0.03	0.03	-0.66	1.00

This table presents descriptive statistics of static commodity strategies formed using full-sample information. $\overline{VOL}$ is the strategy that buys (sells) the top quartile of all commodities with the lowest (highest) detrended implied volatilities. Similarly, $\overline{MOM}$ is the momentum strategy that buys (sells) commodities with the highest (lowest) past 12-month futures returns. $\overline{Basis}$ is the carry strategy that buys (sells) commodities with the lowest (highest) basis. $\overline{BasMom}$ is the strategy that buys (sells) commodities with the highest (lowest) basis-momentum. $\overline{HedgingPressure}$ is the strategy that buys (sells) commodities with the highest (lowest) hedging pressure. $\overline{OpenInterest}$ is the strategy that buys (sells) commodities with the highest (lowest) open interest growth. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The table also reports first order autocorrelation coefficient (AC). Panel A displays the overall commodity excess return, whereas Panel B and Panel C report the nearby return and the roll return component only. Panel D presents the sample correlations of the commodity excess returns among the strategies. The sample spans monthly from February 1991 to October 2014.

Table 4: **Portfolios Sorted by Good and Bad Volatility**

	High	2	3	Low	Low−High
Panel A: Good Volatility					
Mean	-6.16	-1.30	-0.11	0.57	6.72
t-stat	(-1.58)	(-0.41)	(-0.04)	(0.18)	(1.80)
Sdev	18.99	15.42	15.20	15.63	18.15
Skew	-0.77	-0.25	-0.74	-0.16	0.37
Kurtosis	6.07	5.09	5.16	4.17	3.58
Sharpe ratio	-0.32	-0.08	-0.01	0.04	0.37
AC(1)	0.05	0.08	0.01	0.03	0.01
Panel B: Bad Volatility					
Mean	-8.29	-1.26	2.35	0.28	8.57
t-stat	(-2.16)	(-0.40)	(0.74)	(0.09)	(2.41)
Sdev	18.67	15.42	15.60	15.03	17.36
Skew	-0.68	-0.36	-0.80	-0.18	0.41
Kurtosis	5.78	6.66	5.49	3.90	4.38
Sharpe ratio	-0.44	-0.08	0.15	0.02	0.49
AC(1)	0.02	0.04	0.09	-0.02	-0.09
Panel C: Tail Risk					
Mean	-9.12	-1.36	1.04	2.41	11.53
t-stat	(-2.32)	(-0.43)	(0.33)	(0.73)	(2.94)
Sdev	19.18	15.25	15.23	15.99	19.11
Skew	-0.62	-0.47	-0.75	0.06	0.37
Kurtosis	5.81	6.38	6.27	4.26	3.83
Sharp ratio	-0.48	-0.09	0.07	0.15	0.60
AC(1)	0.01	0.07	0.11	-0.01	-0.03

This table presents descriptive statistics of commodity portfolios sorted on the good and bad implied volatility, as well as the tail risk measure at $t-1$ detrended by the previous 12 months mean $\frac{1}{12}\sum_{i=t-13}^{t-2}\sigma_i$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) volatilities or tails. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1) and the annualized Sharpe ratio. Panel A displays the excess returns of portfolios sorted by good volatility. Panel B reports those of portfolios sorted by bad volatility. Panel C shows the results with the tail risk measure. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table 5: **Interacting with Other Variables / Strategies**

Panel A: NBER Recessions and Expansions

	VOL	MOM	Basis	BasMom	Hedging Pressure	Open Interest
Recession	32.94	10.45	37.39	25.33	14.19	8.64
	(2.60)	(0.72)	(2.86)	(2.33)	(1.25)	(0.70)
Expansion	10.45	12.75	2.97	15.83	4.02	10.09
	(2.66)	(2.80)	(0.81)	(4.12)	(1.22)	(2.57)
Nearby Return Component						
Recession	31.71	-20.98	-17.04	-6.95	6.53	4.67
	(2.17)	(-1.53)	(-1.20)	(-0.54)	(0.55)	(0.35)
Non-Recession	12.54	-11.72	-48.11	-5.11	1.20	4.68
	(3.11)	(-2.43)	(-11.73)	(-1.22)	(0.36)	(1.18)

Panel B: Interacting with Other Strategies

		High VOL	t-stat	Low VOL	t-stat	Low−High	t-stat
MOM	High	-4.06	(-0.70)	6.07	(1.44)	13.18	(2.03)
	Low	-15.98	(-3.45)	-2.69	(-0.72)	12.61	(2.40)
	High−Low	11.26	(1.72)	9.01	(1.92)		
Basis	High	-15.92	(-3.03)	2.15	(0.54)	17.73	(2.99)
	Low	-6.01	(-1.10)	7.01	(1.67)	13.50	(2.45)
	High−Low	-7.99	(-1.14)	-4.86	(-1.07)		
BasMom	High	0.59	(0.10)	9.75	(2.49)	8.02	(1.38)
	Low	-21.10	(-4.08)	-7.32	(-1.84)	14.31	(2.59)
	High−Low	22.79	(3.21)	17.14	(3.67)		
Hedg Pres	High	-11.86	(-2.39)	2.13	(0.54)	14.48	(2.70)
	Low	-10.67	(-1.81)	2.60	(0.71)	12.98	(2.20)
	High−Low	-1.16	(-0.18)	0.11	(0.02)		
Open Intst	High	-6.14	(-1.14)	7.04	(1.61)	11.93	(1.95)
	Low	-14.03	(-2.60)	-4.46	(-1.13)	9.63	(1.80)
	High−Low	7.70	(1.18)	11.22	(2.20)		

This table presents the mean returns of the VOL strategy when we interact it with the NBER recession/expansion periods or other commodity strategies. Panel A reports the VOL strategy returns during NBER recession and expansion periods separately. Panel B reports the double-sorts results on VOL and one of the other commodity strategies listed in the first column. Namely, we first split the commodities at median into two groups based on the column variables, before sorting on VOL. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table 6: **Predictive Evidence of VOL for Commodity Returns**

	VOL	Basis	RV	VIX	All
β_{VOL}	-7.39				-7.62
(t-stat)	(-3.16)				(-3.21)
β_{Basis}		5.94			2.26
(t-stat)		(0.66)			(1.05)
β_{RV}			-7.26		4.17
(t-stat)			(-1.55)		(0.46)
β_{VIX}				-0.03	1.58
(t-stat)				(-1.26)	(0.33)
Time FE	Yes	Yes	Yes	Yes	Yes
Comdt FE	Yes	Yes	Yes	Yes	Yes
adj.R^2	0.173	0.169	0.169	0.169	0.172

This table presents the predictive evidence of the detrended implied volatility VOL for commodity futures returns in the next month $t+1$. The results are estimated from panel regressions with both time and commodity fixed effects. For control we also include additional variables such as basis, the realized volatility RV which is the sum of squared return over the month t, and the VIX. The cluster robust errors are estimated both at the time and commodity dimension. The sample period is from February 1991 to October 2014.

Table 7: **Exploring Risk Factors in the VOL Strategy**

Panel A: Strategy Factors

	c	COMDT	MOM	Basis	BasMom	Hedging Pressure	Open Interest	adj.R^2
β	0.009	-0.331	-0.075	-0.006	0.110	-0.025	0.102	0.046
t-stat	(2.76)	(-3.33)	(-0.99)	(-0.06)	(1.09)	(-0.27)	(1.40)	

Panel B: Hedge Fund Factors

	c	PTFSBD	PTFSFX	PTFSCOM	PTFSIR	PTFSSTK		adj.R^2
β	0.94	-3.67	3.61	-1.09	1.03	-0.39		0.00
t-stat	(2.34)	(-1.45)	(1.83)	(-0.38)	(0.76)	(-0.12)		

Panel C: Equity and Bond Market Factors

	c	Equity Market	Size Spread	Bond Market	ΔVIX			adj.R^2
β	1.09	-4.52	-4.17	0.29	-4.54			0.00
t-stat	(3.15)	(-0.39)	(-0.36)	(0.17)	(-2.18)			

Panel D: Business Cycle Factors

	c	Emerging Market	BDI	Credit Spread	Bloom	$\overline{CIV}$		adj.R^2
β	-3.37	5.84	1.43	1.31	-1.14	19.34		0.02
t-stat	(-1.75)	(1.17)	(0.68)	(0.47)	(-1.16)	(3.19)		

Panel E: Liquidity Factors

	Broker Dealer	Pastor	TED	Funding Liquidity	BaB	Bank
c	2.92	1.10	-0.05	1.06	1.21	0.97
t-stat	(2.74)	(3.20)	(-0.08)	(3.13)	(3.70)	(2.60)
β	-32.11	-8.63	2.20	0.07	-17.64	9.54
t-stat	(-1.90)	(-0.82)	(2.61)	(0.20)	(-1.45)	(1.30)
adj.R^2	0.04	0.00	0.02	0.00	0.01	0.01

This table presents estimates from the time series regressions. The dependent variable is the monthly return from the VOL strategy. Panel A employs explanatory return variables from the commodity strategies described in Table 2 and a strategy called "COMDT" that is a simple average of all commodity futures returns at time t. Panel B uses the Hedge Fund factors from Fung and Hsieh (2001). Panel C adopts the equity market (S&P 500 index) return, the size spread (Russell 2000 index monthly total return − S&P 500 monthly total return, a bond market factor (change in the 10-year treasury constant maturity yield), and the change in VIX. Panel D documents the emerging market return (MSCI Emerging Market index monthly total return), the log change in the BDI, credit spread (Moody's Baa less 10-year treasury constant maturity yield), the Bloom Economic Policy Uncertainty index, and the average commodity option implied volatility $\overline{CIV}$. Panel E explores the liquidity factors: the shock to Broker-Dealer leverage from Adrian et al. (2014), the Pastor and Stambaugh (2003) liquidity factor, the TED spread (spread between 3-month LIBOR and 3-month treasury bond), the funding liquidity factor from Fontaine and Garcia (2012), the Betting-against-Beta factor from Frazzini and Pedersen (2014), and the FTSE World Bank Index return. The Broker-Dealer measure is at quarterly frequency, whereas all other measures are monthly. The returns from the VOL strategy and from the other strategies in Panel A are in percentage points.

Table 8: **Cross-Sectional Asset Pricing Tests**

	λ_{Comdt}	λ_{Basis}	λ_{Mom}	λ_{BasMom}	λ_{VOL}	b_{Comdt}	b_{Basis}	b_{Mom}	b_{BasMom}	b_{VOL}	R^2_{OLS} (R^2_{GLS})	χ^2_{NW} p-value	χ^2_{SH} p-value	KR p-value (HJ-Dist.)
Panel A: Pricing Model with VOL														
FMB	-0.001	0.006	0.014	0.013	0.011						0.86	25.69	22.91	
	(-0.42)	(2.01)	(4.18)	(3.42)	(3.09)						(0.84)	(0.37)	(0.53)	
	[-0.41]	[2.01]	[4.23]	[3.36]	[3.06]									
GMM	-0.001	0.006	0.014	0.013	0.011	-0.522	-0.753	2.789	3.585	3.671				0.01
	(-0.28)	(1.62)	(3.89)	(3.65)	(3.13)	(-4.14)	(-1.89)	(10.08)	(10.57)	(13.94)				(0.36)
Panel B: Pricing Model with Factors from Boons and Prado (2016)														
FMB	-0.001	0.006	0.014	0.013							0.70	35.20	32.48	
	(-0.43)	(2.00)	(4.14)	(3.52)							(0.49)	(0.08)	(0.14)	
	[-0.43]	[2.01]	[4.17]	[3.48]										
GMM	-0.001	0.006	0.013	0.013		-1.741	-0.341	2.177	4.083					
	(-0.32)	(2.18)	(4.46)	(4.48)		(-18.53)	(-0.86)	(7.90)	(12.11)					(0.37)
Panel C: Pricing Model with Factors from Bakshi et al. (2016)														
FMB	-0.001	0.008	0.015								0.58	49.46	46.46	
	(-0.43)	(2.64)	(4.47)								(0.52)	(0.00)	(0.01)	
	[-0.42]	[2.64]	[4.49]											
GMM	-0.001	0.008	0.015			-1.413	1.098	3.548						
	(-0.46)	(3.43)	(6.40)			(-15.88)	(2.99)	(13.46)						(0.47)
Panel D: Average Excess Return of the Factors														
$\overline{R}_{Factor}$	-0.001	0.005	0.010	0.014	0.011									

This table reports the factor risk premia (λ) and the SDF parameters (b) from the asset pricing tests. The SDF specification in Panel A is of the from: $m_{t+1} = 1 - b_{Comdt}Comdt_{t+1} - b_{Basis}Basis_{t+1} - b_{Mom}Mom_{t+1} - b_{BasMom}BasMom_{t+1} - b_{VOL}VOL_{t+1}$. *Comdt* is the average excess return obtained from a strategy that holds all available commodities. *Basis*, *Mom*, *BasMom*, and *VOL* correspond to the long-short portfolios returns from the strategies described in Section 3. The test assets are excess returns to four portfolios sorted on basis, momentum, basis-momentum, hedging pressure, and open interest respectively, and sectoral portfolios including energy, grains, livestock, softs, and metals. In the row marked "GMM", the parameters are estimated following a one-step procedure by putting on a large weight on the mean and covariance parameters in the weighting matrix. The row "FMB" represents the two-step Fama-MacBeth cross-sectional regressions. For the Fama-MacBeth procedure, the t-stat are compuated using both the Newey-West standard errors without (with) the Shanken correction in parentheses (blankets). We report the OLS cross-sectional uncentered R^2 and the GLS uncentered R^2 in parentheses (Lewellen et al. (2010), Prescription 3). The χ^2 test value corresponds to the null hypothesis that the pricing errors are zero. The Hansen-Jagannathan distance (HJ-Dist.) is reported in the last column in parentheses. The KR p-value shows the Kan and Robotti (2009) test p-value of the null hypothesis that the Hansen-Jagannathan distance of Model A with the VOL factor and Model B without VOL factor are equal.

Table 9: **Hedging Pressure and the Commodity VOL**

Panel A: Commercial Traders			
	Futures Only	Futures & Options	Options Only
Hedging Pressure	-0.35	-0.99	-1.13
t-stat	(-0.56)	(-1.25)	(-2.38)
Controls?	Yes	Yes	Yes
Cluster Time	Yes	Yes	Yes
Cluster Comdt	Yes	Yes	Yes
R^2	0.53	0.52	0.51
N	6545	5466	5324
Panel B: Disaggregated Futures Only			
	Producer	Swap Dealer	Money Manager
Hedging Pressure	-1.20	-0.08	1.05
t-stat	(-1.34)	(-0.21)	(2.01)
Controls?	Yes	Yes	Yes
Cluster Time	Yes	Yes	Yes
Cluster Comdt	Yes	Yes	Yes
R^2	0.63	0.62	0.62
N	6545	5466	5324
Panel C: Disaggregated Futures & Options Combined			
	Producer	Swap Dealer	Money Manager
Hedging Pressure	-1.58	-0.20	0.99
t-stat	(-1.65)	(-0.59)	(1.73)
Controls?	Yes	Yes	Yes
Cluster Time	Yes	Yes	Yes
Cluster Comdt	Yes	Yes	Yes
R^2	0.63	0.62	0.62
N	6545	5466	5324

This table presents panel estimates of $VOL_{i,t+1} = \beta \cdot HP_{i,t} + \lambda' Control_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}$. HP is the hedging pressure defined as the net short position of a certain type of traders divided by the sum of long and short position of that type of traders. The standard error estimation includes both clustering at the time and commodity level following Petersen (2009). The coefficient β for the hedging pressure, its t-statistics, R^2, and the length of the available time series N in each regression are reported. The control variables are the lagged dependent variable, the lagged basis, the lagged realized volatility, and the lagged VIX. The CFTC data are collected on the last Tuesday of each month. Panel A shows results of the commercial traders (hedgers typically) from the Commitment of Traders report of CFTC. They include futures only, futures & options, and options only. The latter is backed out from the first two. Panel B and C presents the disaggregate data which classify traders into producers, swap dealers, and money managers, etc. Panel B show the futures only data while Panel C show the combined data with both futures and options. Data are monthly. The futures only data in Panel A start from March 1995 and the futures & options from February 1990. The disaggregated COT data are from June 2006.

Table 10: **Portfolios Sorted by the Average Commodity VOL (AVOL) Beta**

	High	2	3	Low	Low−High
Panel A: Excess Returns					
Mean	-5.66	-2.72	-3.19	3.71	9.37
t-stat	(-1.32)	(-0.78)	(-0.91)	(1.00)	(2.12)
Sdev	19.53	15.88	16.07	16.94	20.20
Skew	-0.49	-1.11	-0.36	-0.00	-0.06
Kurtosis	5.46	8.65	4.04	4.29	3.03
Sharp ratio	-0.29	-0.17	-0.20	0.22	0.46
AC(1)	0.08	0.01	0.06	0.06	0.06
Panel B: Nearby Returns					
Mean	-0.71	3.36	3.72	7.05	7.76
t-stat	(-0.17)	(0.90)	(1.05)	(1.85)	(1.74)
Sdev	19.06	17.03	16.23	17.40	20.33
Skew	-0.49	-0.98	-0.29	-0.11	-0.10
Kurtosis	5.33	7.63	4.22	4.87	3.11
Sharp ratio	-0.04	0.20	0.23	0.41	0.38
AC(1)	0.05	0.01	0.04	0.01	0.02
Panel C: Roll Returns					
Mean	-4.95	-6.08	-6.91	-3.34	1.61
t-stat	(-5.00)	(-5.68)	(-6.81)	(-2.72)	(1.20)
Sdev	4.51	4.88	4.63	5.62	6.10
Skew	-0.26	1.11	0.47	0.59	0.28
Kurtosis	7.04	7.33	5.24	8.68	5.99
Sharp ratio	-1.10	-1.25	-1.49	-0.60	0.26
AC(1)	0.33	0.20	0.13	0.39	0.18
Panel D: Beta before and after Portfolio Formation					
pre-β	0.622	0.128	-0.144	-0.601	
post-β	0.600	0.119	-0.137	-0.576	

This table presents descriptive statistics of portfolios sorted on their beta exposure to the aggregate commodity detrended implied volatility at $t-1$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) betas. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1), the annualized Sharpe ratio. Panel A displays the overall excess returns, whereas Panel B and Panel C report the nearby return and the roll return component only. Pane D presents the pre- and post-formation βs. Returns are expressed in percentage per annum. t-stat is based on the Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table 11: **Media Predictive Power**

Panel A: Portfolio Sorted by RMC					
	High	2	3	Low	Low − High
Mean	-3.93	-2.89	-2.25	5.36	9.29
t-stat	(-0.69)	(-0.67)	(-0.45)	(1.08)	(2.03)
Sdev	21.16	15.94	18.39	18.48	16.96
Skew	-0.80	-0.90	-0.81	-0.09	-0.12
Kurtosis	5.86	4.82	4.99	4.65	3.65
Sharp ratio	-0.19	-0.18	-0.12	0.29	0.55
AC(1)	0.11	-0.05	0.05	0.13	-0.09

Panel B: VOL and RMC					
	Media	Basis	RV	VIX	All
β_{Media}	0.05				0.04
t-stat	(0.51)				(0.44)
β_{Basis}		5.49			5.87
t-stat		(0.52)			(0.54)
β_{RV}			-16.31		-16.41
t-stat			(-2.98)		(-2.93)
β_{VIX}				0.02	0.03
t-stat				(1.12)	(1.97)
Time FE	Yes	Yes	Yes	Yes	Yes
Comdt FE	Yes	Yes	Yes	Yes	Yes
R^2	0.539	0.539	0.541	0.539	0.541

Panel C: Double Sort by VOL on RMC					
High RMC & High VOL	t-stat	Low RMC & Low VOL	t-stat	Diff	t-stat
-13.26	(-1.98)	7.44	(1.35)	20.70	(2.74)

This table presents the predictive results of uncertainty proxied by the media coverage. In each month, residual media coverage is obtained from the cross-sectional regressions of the number of news articles on its own lag, the open interest, a dummy variable for the membership in the S&P GSCI index and a dummy variable for the membership in the Bloomberg Commodity Index. Panel A presents the returns and descriptive statistics of portfolios sorted by the media coverage residual (RMC) and the long-short portfolio that is short in the portfolio with the highest RMC and long in portfolio with the lowest RMC. The strategy is rebalanced monthly. Panel B presents the results from panel regressions of VOL of the individual commodities on media coverage and other variables with both time and commodity fixed effects. The left hand side variable is the VOL variable, whereas the right hand side variables are the lagged RMC, Basis, the realized volatility, the VIX, or all the aforementioned variables. For control we also include the lagged VOL. The cluster robust errors are estimated both at the time and commodity dimension. Panel C displays the independent double sort by both RMC and VOL and return from the long-short strategy that is short in high RMC and high VOl commodities and long in low VOL and low RMC commodities. The sample period is from February 2001 to October 2014.

Table 12: **Decomposing the Predictive Ability of VOL**

Panel A: Decomposing the VOL Coefficient: Univariate Analysis

Stage			Coeff.	Fract	*t*-stat	Coeff.	Fract	*t*-stat	Coeff.	Fract	*t*-stat
				Skew			Maxret			Basis	
1	Return on VOL	VOL	-0.058		(-2.31)	-0.058		(-2.31)	-0.058		(-2.31)
2	Return on VOL	VOL	-0.054		(-2.19)	-0.053		(-2.17)	-0.060		(-2.46)
	and candi	Candi	0.000		(0.02)	0.003		(0.03)	-0.099		(-1.21)
3	VOL on candi	Candi	0.007		(3.42)	1.199		(14.03)	-0.140		(-1.82)
		avg adj R^2	0.017			0.135			0.057		
4	Decomp Stage 1	Candi	-0.006	0.11	(1.29)	-0.021	0.35	(2.70)	-0.000	0.00	(0.01)
	VOL coef	Resid	-0.052	0.89	(10.80)	-0.037	0.65	(4.92)	-0.058	1.00	(8.14)
	Avg ♯ comdt/m		22.96			22.96			22.96		
				Lagret			Open			Amihud	
1	Return on VOL	VOL	-0.058		(-2.31)	-0.063		(-2.49)	-0.066		(-2.47)
2	Return on VOL	VOL	-0.058		(-2.38)	-0.066		(-2.63)	-0.054		(-2.09)
	and candi	Candi	-0.051		(-2.62)	0.057		(2.40)	-0.001		(-0.25)
3	VOL on candi	Candi	0.004		(0.18)	0.016		(0.81)	0.019		(11.06)
		avg adj R^2	0.048			0.027			0.062		
4	Decomp Stage 1	Candi	-0.008	0.14	(1.29)	-0.003	0.04	(0.51)	-0.017	0.26	(2.22)
	VOL coef	Resid	-0.050	0.86	(7.89)	-0.060	0.96	(11.41)	-0.049	0.74	(6.33)
	Avg ♯ comdt/m		22.96			22.92			22.33		

Panel B: Decomposing the VOL Coefficient: Multivariate Analysis

Stage			Coeff.	Fract	t-stat	Coeff.	Fract	t-stat	Coeff.	Fract	t-stat
1	Return on VOL	VOL	-0.058		(-2.31)	-0.071		(-2.64)	-0.071		(-2.64)
2	Return on VOL	VOL	-0.060		(-2.37)	-0.071		(-2.55)	-0.070		(-2.51)
	and candi	Skew	0.000		(0.18)				-0.000		(-0.17)
		Maxret	0.008		(0.09)				-0.076		(-0.76)
		Basis	-0.094		(-1.10)				-0.188		(-2.12)
		Lagret				-0.050		(-2.47)	-0.052		(-2.41)
		OI				0.040		(1.58)	0.037		(1.53)
		Amihud				0.001		(0.45)	0.000		(0.18)
3	VOL on candi	Skew	0.000		(0.20)				0.000		(0.20)
		Maxret	1.094		(12.62)				0.923		(10.86)
		Basis	-0.094		(-1.31)				-0.112		(-1.59)
		Lagret				0.008		(0.37)	-0.011		(-0.57)
		OI				0.059		(2.69)	0.028		(1.36)
		Amihud				0.020		(10.59)	0.012		(6.51)
		avg adj R^2	0.180			0.143			0.280		
4	Decomp Stage 1	Skew	-0.004	0.07	(0.80)				-0.004	0.05	(0.76)
	VOL coef	Maxret	-0.021	0.36	(2.49)				-0.024	0.34	(2.76)
		Basis	0.006	-0.10	(-0.79)				-0.008	0.11	(1.11)
		Lagret				-0.007	0.10	(1.08)	-0.002	0.03	(0.24)
		OI				0.005	-0.08	(-0.88)	0.007	-0.10	(-1.14)
		Amihud				-0.018	0.26	(2.20)	-0.006	0.09	(0.86)
		Resid	-0.039	0.67	(4.39)	-0.051	0.72	(5.64)	-0.034	0.48	(3.82)
	Avg ♯ comdt/m		22.96			22.29			22.29		

This table decomposes the predictive ability of VOL by candidate variables related to lottery preferences of investors and market frictions following Hou and Loh (2016) using four-step Fama-MacBeth cross-sectional regressions. Step 1 regresses month t commodity returns on VOL in month $t-1$ ($r_{i,t} = \alpha_t + \beta_t VOL_{i,t-1} + \epsilon_{i,t}$). Step 2 adds a candidate variable $Candi_{t-1}$ to the regression. Step 3 regresses VOL_{t-1} on the candidate variable $Candi_{t-1}$ ($VOL_{t-1} = c_{t-1} + \delta_{t-1} Candi_{i,t-1} + \varepsilon_{i,t-1}$) to decompose $VOL_{i,t-1}$ into two orthogonal components: $\delta_{t-1} Candi_{i,t-1}$ and $(c_{t-1} + \varepsilon_{i,t-1})$. In Step 4 the β_t from Stage 1 is decomposed as: $\beta_t = \frac{cov(r_{i,t}, VOL_{i,t-1}}{Var(VOL_{i,t-1}} = \frac{cov(r_{i,t}, \delta_{t-1} Candi_{i,t-1}}{Var(VOL_{i,t-1}} + \frac{cov(r_{i,t}, (c_{t-1} + \varepsilon_{i,t-1})}{Var(VOL_{i,t-1}} = \beta_t^C + \beta_t^R$. The fraction of the negative VOL return relation explained by the candidate variable is measured by $E(\beta_t^C)/E(\beta_t)$, and the fraction unexplained by the candidate is measured by $E(\beta_t^R)/E(\beta_t)$. Panel A examines the individual candidate variables, where *Skew* is the respective commodity option implied skew in month $t-1$, *Maxret* is the maximum daily return in month $t-1$, *Basis* is the difference between the second nearby futures contract and the first nearby futures contract, *Lagret* is the month $t-1$ return, *Open* is the open interest growth, *Amihud* is the Amihud (2002) illiquidity measure. Panel B examines all the candidate variables jointly. Time series averages of the coefficients, fraction explained by the variables, and t-statistics (in parentheses) are reported.

Table 12: **FX Portfolios Sorted by Commodity VOL**

	High	2	3	4	Low	Low−High
Panel A: Excess Returns						
Mean	-1.51	-0.53	1.74	1.89	3.26	4.77
t-stat	(-1.02)	(-0.37)	(1.14)	(0.98)	(1.56)	(2.92)
Sdev	6.76	6.44	6.96	8.73	9.48	7.43
Skew	-0.36	-0.04	-0.52	-0.82	-0.18	-0.04
Kurtosis	4.41	4.84	5.86	5.91	4.04	3.52
Sharpe ratio	-0.22	-0.08	0.25	0.22	0.34	0.64
AC(1)	0.06	0.08	0.03	0.12	0.06	-0.04
Panel B: Spot Returns						
Mean	2.58	0.96	-0.30	-0.21	-0.75	-3.32
t-stat	(1.71)	(0.69)	(-0.20)	(-0.11)	(-0.36)	(-2.00)
Sdev	6.85	6.40	6.94	8.69	9.39	7.57
Skew	0.31	0.09	0.63	0.87	0.25	0.04
Kurtosis	4.40	4.96	6.05	6.08	4.09	3.59
Sharpe ratio	0.38	0.15	-0.04	-0.02	-0.08	-0.44
AC(1)	0.07	0.08	0.02	0.12	0.05	-0.03
Panel C: Forward Discount						
Mean	1.06	0.44	1.44	1.68	2.51	1.45
t-stat	(5.99)	(2.65)	(8.68)	(9.41)	(9.57)	(4.75)
Sdev	0.81	0.75	0.75	0.81	1.19	1.38
Skew	0.71	-0.92	0.84	0.50	-0.00	-0.54
Kurtosis	5.85	10.52	4.76	3.47	4.37	3.71
Sharpe ratio	1.32	0.58	1.91	2.07	2.10	1.04
AC(1)	0.73	0.53	0.61	0.64	0.83	0.81
Panel D: Transition Matrix						
P_{high}	0.85	0.12	0.02	0.01	0.00	
P_2	0.11	0.71	0.15	0.03	0.00	
P_3	0.02	0.13	0.69	0.14	0.02	
P_4	0.01	0.02	0.15	0.71	0.11	
P_{low}	0.00	0.01	0.02	0.13	0.84	
$\overline{\pi}$	*0.20*	*0.20*	*0.20*	*0.20*	*0.20*	

This table presents descriptive statistics of currency return portfolios sorted on their beta exposure to the aggregate commodity detrended implied volatility at $t-1$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) betas. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1), the annualized Sharpe ratio. Panel A displays the overall excess return, whereas Panel B and Panel C report the nearby return and the roll return component only. Pane D presents the transition probability from portfolio i to portfolio j between time t and time $t+1$. $\overline{\pi}$ indicates the steady state probability. Returns are expressed in percentage per annum. t-stat is based on the Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

A Appendix Tables

Table A.1: **Overview of the Commodities Sample**

Number	Name	Sector	Mnemonic	Start	End	Maturity Months	Exchange	GSCI	Bloomberg /DJ	small sample?
1	Crude Oil	energy	CL	1990:01	2014:10	1-12	NYMEX	yes	yes	yes
2	Heating Oil	energy	HO	1990:01	2014:10	1-12	NYMEX	yes	yes	yes
3	Natural Gas	energy	NG	1992:10	2014:10	1-12	NYMEX	yes	yes	yes
4	Gasoline, Blendstock	energy	RB	2006:08	2014:10	1-12	NYMEX	yes	yes	yes
5	Gasoline, Unleaded	energy	HU	1990:11	2006:11	1-12	NYMEX	yes	yes	yes
6	Wheat	grains	W-	1990:01	2014:10	3,5,7,9,12 + serial	CBOT	yes	yes	yes
7	Kansas Wheat	grains	KW	1990:08	2014:10	3,5,7,9,12	KCBT	yes	no[a]	no
8	Corn	grains	C-	1990:01	2014:10	3,5,7,9,12 + serial	CBOT	yes	yes	yes
9	Soybean	grains	S-	1990:01	2014:10	1,3,5,7,8,9,11	CBOT	yes	yes	yes
10	Soybean Oil	grains	BO	1990:01	2014:10	1,3,5,7,8,9,10,12 + serial	CBOT	no	yes	no
11	Soybean Meal	grains	SM	1990:01	2014:10	1,3,5,7,8,9,10,12 + serial	CBOT	no	no[b]	no
12	Rough Rice	grains	RR	1992:04	2014:10	1,3,5,7,9,11 + serial	CBOT	no	no	no
13	Oats	grains	O-	1990:05	2014:10	3,5,7,9,12 + serial	CBOT	no	no	no
14	Feeder Cattle	livestock	FC	1990:01	2014:10	1,3,4,5,8,9,10,11	CME	yes	no	no
15	Lean Hogs	livestock	LH	1996:06	2014:10	2,4,5,6,7,8,10,12	CME	yes	yes	yes
16	Live Cattle	livestock	LC	1990:01	2014:10	2,4,6,8,10,12 + serial	CME	yes	yes	yes
17	Cocoa	softs	CC	1990:03	2014:10	3,5,7,9,12	NYBOT	yes	no	no
18	Coffee	softs	KC	1990:03	2014:10	3,5,7,9,12	NYBOT	yes	yes	yes
19	Sugar	softs	SB	1990:03	2014:10	3,5,7,10,12	NYBOT	yes	yes	yes
20	Orange Juice	softs	JO	1990:03	2014:10	1,3,5,7,9,11	NYBOT	no	no	no
21	Copper	metals	HG	1990:01	2014:10	1-12	NYMEX	yes	yes	yes
22	Gold	(precious) metals	GC	1990:01	2014:10	1-12	NYMEX	yes	yes	yes
23	Silver	(precious) metals	SI	1999:12	2014:10	3,5,7,9,12 + serial	NYMEX	yes	yes	yes
24	Cotton	industrial materials	CT	1990:01	2014:10	3,5,7,10,12 + serial	NYBOT	yes	yes	yes
25	Lumber	industrial materials	LB	1990:01	2014:10	1,3,5,7,9,11 + serial	CME	no	no	no

This table describes the 25 commodities examined in our sample. It provides the names, sectors, mnemonics, start and ending months in our sample, as well as the option maturity months and exchanges on which they are traded. The second and third last columns document whether the commodity is a component of the current S&P GSCI (formerly the Goldman Sachs Commodity Index) and Bloomberg (formerly DJ UBS or DJ AIG) Commodity indices. The last column states whether the commodity is included in the small sample.

[a] except since 2013

[b] except since 2013

Table A.2: **Open interest of the Commodity Options**

	Mean	Median	St.Dev	Min	Max	AR(1)	1. Quartile	4. Quartile
Crude Oil	606611644	440484000	493475894	68469000	2147255000	0.98	173751500	973557000
Heating Oil	1325737030	1056468000	788766922	228354000	4663176000	0.88	813477000	1550157000
Natural Gas	1843691459	1488840000	1468621965	169720000	5342510000	0.98	440520000	2880705000
Gasoline, Blendstock	874685511	759318000	502330804	172074000	3129756000	0.81	477414000	1096830000
Gasoline, Unleaded	874685511	759318000	502330804	172074000	3129756000	0.81	477414000	1096830000
Wheat	3135450	2314450	1837365	757900	8403700	0.82	1653225	4537250
Kansas Wheat	422682	359300	239823	71830	1328000	0.75	251075	554550
Corn	14485442	9486850	10258264	2725550	42514550	0.92	5841625	21940550
Soybean	5693240	4158100	4054365	1456600	23284300	0.88	2718265	7701825
Soybean Oil	18119673	16422000	10775389	3068400	59793600	0.85	9765300	23952600
Soybean Meal	2311909	1929500	1432404	443400	9356100	0.87	1398850	2671400
Rough Rice	1977339	1680000	1148472	344000	6458000	0.67	1160000	2539000
Oats	76574	66900	41754	8500	222450	0.70	46095	95275
Feeder Cattle	1873511	1650440	965655	611160	6740360	0.74	1221220	2278760
Lean Hogs	7347045	2767600	8120629	562000	44739200	0.95	1832400	11062800
Live Cattle	15888532	9289200	12608507	3112800	57198000	0.94	6492400	25058400
Cocoa	151830	122890	90520	38150	512580	0.79	91440	183020
Coffee	12275527	10753500	7412271	2503125	33033000	0.90	5572313	17389875
Sugar	155976007	134062880	118968108	25313120	530860960	0.94	56247520	218024240
Orange Juice	1621246	1484250	798376	361500	5249850	0.76	1014450	2033550
Copper	570563	345750	617888	6750	4045500	0.89	117875	900250
Gold	13281818	11001300	7617030	3796500	44569300	0.91	7415350	17716350
Silver	1431072	1312600	705236	376200	3410300	0.86	811975	1906925
Cotton	30930330	23941500	22249094	7465000	134364000	0.88	15458500	38163000
Lumber	55811	47850	33085	9130	243540	0.71	35310	67825

This table presents the statistical properties of the dollar value open interest of commodity options. These numbers are backed out from the Commitment of Traders report and are calculated as the total open interest of futures and options combined minus that of futures only. The statistical properties reported are mean, median, standard deviation, the first auto-correlation coefficient, and the open interest at the first and last quartile. The data is monthly and the sample runs from March 1995 to October 2014.

Table A.3: **Lowest 4 versus Highest 4 Volatility Commodities in Portfolio Formation**

		Lowest volatility comdt				Highest volatility comdt			
		Lowest	Two lowest	Three lowest	Four lowest	Four highest	Three highest	Two highest	Highest
1	CL	16	34	47	57	72	54	38	14
2	HO	8	27	42	50	64	51	36	15
3	NG	44	60	71	77	106	96	87	68
4	RB	4	9	11	12	25	19	10	5
5	HU	6	18	31	37	35	28	22	15
6	W-	7	15	24	41	18	14	6	3
7	KW	3	12	25	36	20	13	5	2
8	C-	14	23	38	62	55	40	23	13
9	S-	5	16	31	47	46	28	17	7
10	BO	1	7	15	23	21	9	5	0
11	SM	11	16	29	42	35	28	15	4
12	RR	6	19	29	38	44	31	21	14
13	O-	7	18	37	53	53	42	28	15
14	FC	5	9	15	23	8	4	1	1
15	LH	9	18	25	37	36	30	14	4
16	LC	0	7	15	23	15	10	3	1
17	CC	8	22	36	48	50	26	16	7
18	KC	32	49	61	72	70	54	37	18
19	SB	17	33	49	54	64	44	24	6
20	JO	25	44	58	71	87	77	59	33
21	HG	9	13	20	37	52	35	22	7
22	GC	5	16	31	44	26	17	7	2
23	SI	10	24	33	40	37	30	24	15
24	CT	12	25	32	45	38	27	18	5
25	LB	21	36	50	71	63	48	32	11

This table reports the number of months the commodity is among those with the lowest or highest volatility. At the end of each month over the sample period February 1991 to October 2014, we rank-order the commodity futures according to their detrended implied volatilites. The column labeled "Lowest" shows how many months the respective commodity have the lowest volatilities. The next three columns show likewise how many months the volaility of the respective commodities is among the two, three, or four lowest ones. Analogously, the last four columns show how often the volatility of a commodity has been the highest, or among the two, three or four highest ones. For example, on 14 (54) occasions crude oil volatility has been the lowest (among the three lowest ones) and is hence in the long portfolio. On 16 (47) occasions it has been the highest volatility (among the three highest ones).

Table A.4: **Portfolios Sorted by Alternative Volatility Measures**

	High	2	3	Low	Low−High
Implied Volatility					
Mean	-8.07	-0.73	0.84	0.53	8.60
t-stat	(-1.73)	(-0.21)	(0.26)	(0.24)	(2.02)
Sdev	22.76	16.84	15.73	10.70	20.75
Skew	-0.33	-0.44	-0.71	-0.02	0.04
Kurtosis	4.06	5.20	5.90	6.21	2.87
Sharpe ratio	-0.35	-0.04	0.05	0.05	0.41
AC(1)	0.06	0.00	0.05	0.13	0.03
Realized Volatility					
Mean	-5.06	-0.48	-0.04	0.01	5.06
t-stat	(-1.22)	(-0.13)	(-0.01)	(0.00)	(1.30)
Sdev	20.26	17.57	16.81	10.85	18.94
Skew	-0.21	-0.24	-0.97	-0.46	-0.21
Kurtosis	4.50	4.36	7.02	4.34	3.40
Sharpe ratio	-0.25	-0.03	-0.00	0.00	0.27
AC(1)	0.04	0.04	0.03	0.08	0.03
Variance Risk Premium					
Mean	-1.22	1.27	-0.93	-4.66	-3.44
t-stat	(-0.36)	(0.47)	(-0.32)	(-1.19)	(-0.90)
Sdev	16.44	13.31	14.29	19.13	18.68
Skew	-0.84	-0.27	-0.14	-0.34	0.14
Kurtosis	8.28	4.55	4.94	4.19	3.53
Sharpe ratio	-0.07	0.10	-0.07	-0.24	-0.18
AC(1)	0.06	0.04	-0.00	0.01	-0.00
Implied Skewness					
Mean	-5.76	-1.14	0.30	-0.61	5.15
t-stat	(-1.66)	(-0.29)	(0.09)	(-0.22)	(1.49)
Sdev	16.88	19.31	17.12	13.30	16.89
Skew	0.11	-0.49	-0.66	-0.94	-0.13
Kurtosis	3.80	4.35	6.48	6.62	2.96
Sharpe ratio	-0.34	-0.06	0.02	-0.05	0.30
AC(1)	-0.04	0.13	0.05	0.14	-0.05

This table presents descriptive statistics formed using alternative volatility measures, variance risk premia, and skewness at $t-1$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) volatilities. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1) and the annualized Sharpe ratio. Panel A displays the excess return of portfolios sorted by the raw implied volatility, whereas Panel B and Panel C report the excess return sorted by the realized volatility and the variance risk premia. Realized volatility is calculated as the square root of the sum of daily squared return in month t. Variance risk premium is the difference between the implied variance at the end of month t and the realized variance over the month t. Returns are expressed in percentage per annum. t-stat is based on the Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table A.5: **Portfolios Sorted by Detrended Implied Volatility – Small Sample**

	High	2	3	Low	Low−High
Panel A: Excess Returns					
Mean	-9.09	1.00	0.72	3.82	12.90
t-stat	(-1.83)	(0.27)	(0.22)	(0.95)	(2.49)
Sdev	24.19	17.82	15.80	19.66	25.31
Skew	-0.29	-0.32	-0.53	-0.05	0.05
Kurtosis	4.71	4.87	7.31	3.70	3.00
Sharp ratio	-0.38	0.06	0.05	0.19	0.51
AC(1)	0.07	0.10	0.03	0.05	0.02
Panel B: Nearby Returns					
Mean	-6.48	4.71	6.06	9.51	15.98
t-stat	(-1.29)	(1.25)	(1.85)	(2.26)	(2.85)
Sdev	24.55	18.44	15.95	20.52	27.32
Skew	-0.41	-0.25	-0.51	-0.15	0.22
Kurtosis	4.98	4.38	7.15	3.54	3.42
Sharp ratio	-0.26	0.26	0.38	0.46	0.59
AC(1)	-0.03	0.09	0.01	0.04	-0.01
Panel C: Roll Returns					
Mean	-2.61	-3.72	-5.34	-5.69	-3.08
t-stat	(-1.28)	(-3.55)	(-6.85)	(-4.84)	(-1.36)
Sdev	9.95	5.10	3.80	5.73	11.02
Skew	0.37	-0.36	-0.25	0.37	-0.62
Kurtosis	7.47	5.15	4.52	7.62	7.01
Sharp ratio	-0.26	-0.73	-1.40	-0.99	-0.28
AC(1)	0.37	0.25	0.08	0.10	0.29
Panel D: Transition Matrix					
P_{high}	0.61	0.26	0.09	0.04	
P_2	0.23	0.42	0.27	0.08	
P_3	0.10	0.27	0.42	0.21	
P_{low}	0.05	0.10	0.26	0.59	
$\overline{\pi}$	*0.25*	*0.25*	*0.25*	*0.25*	
Panel E: VOL before and after Portfolio Formation					
pre-VOL	0.08	0.01	-0.03	-0.08	
post-VOL	0.06	0.01	-0.02	-0.06	

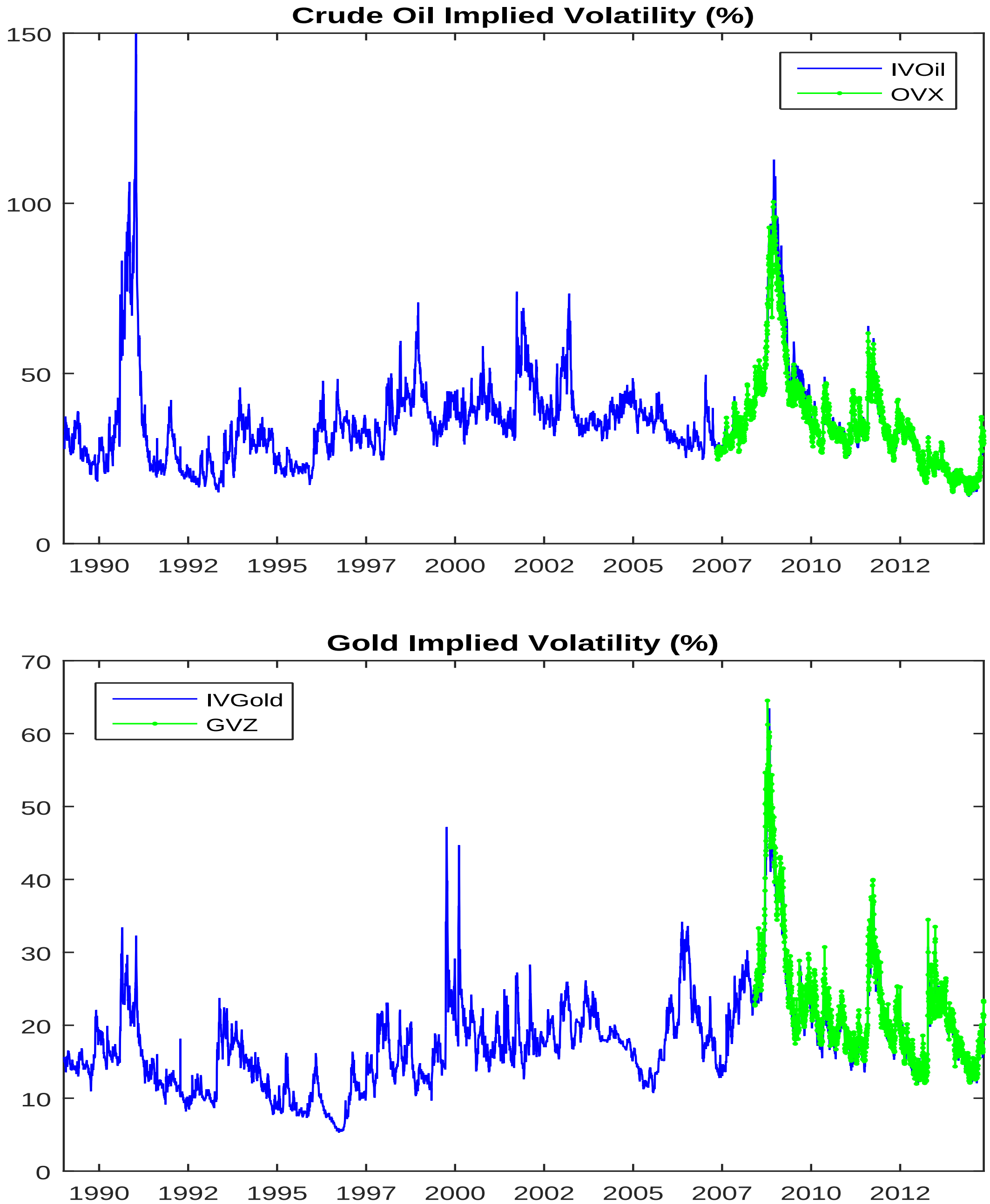
This table presents descriptive statistics of commodity portfolios sorted on the implied volatility at $t-1$ detrended by the previous 12 months mean of implied volatility $\frac{1}{12}\sum_{i=t-13}^{t-2}\sigma_i$. The "Low" ("High") portfolio contains the top 25% of all commodities with the lowest (highest) volatilities. "Low−High" denotes the long-short strategy that buys "Low" and sells "High". The table also reports the first-order autocorrelation coefficient AC(1), the annualized Sharpe ratio. Panel A displays the overall excess return, whereas Panel B and Panel C report the nearby return and the roll return component only. Pane D presents the transition probability from portfolio i to portfolio j between time t and time $t+1$. $\overline{\pi}$ indicates the steady state probability. Returns are expressed in percentage per annum. t-stat is based on Newey-West standard errors. The strategies are rebalanced monthly from February 1991 to October 2014.

Table A.6: **Risk Premia Given the Possibility of Omitted Factors**

	PC=6			PC=7			PC=8		
Factors	Risk Premium	t-stat	R_g^2	Risk Premium	t-stat	R_g^2	Risk Premium	t-stat	R_g^2
λ_{Comdt}	-0.007	[-2.01]	0.998	-0.003	[-1.05]	0.998	0.003	[0.82]	0.998
λ_{Basis}	0.008	[3.12]	0.460	0.010	[3.61]	0.585	0.009	[3.02]	0.651
λ_{Mom}	0.013	[3.95]	0.776	0.013	[3.85]	0.776	0.012	[3.66]	0.780
λ_{BasMom}	0.007	[2.68]	0.439	0.010	[3.37]	0.585	0.011	[3.87]	0.586
λ_{VOL}	0.006	[1.80]	0.369	0.008	[2.49]	0.696	0.007	[2.27]	0.720

This table reports the results of the three-pass cross-sectional regressions based on Giglio and Xiu (2017) with six, seven, and eight latent factors. The Giglio and Xiu (2017) method addresses the potential missing variable problem in the linear factor models and the resulted estimation bias. In the first pass, principle components are used as latent factors to recover the factor space. In a second step, using OLS the factor loadings are estimated to obtain the risk premia of the factors. In a third pass, the observable factor premia are estimated based on the projection of the observable factors on the latent factors. The number of the latent factors is determined by the information criteria following Bai and Ng (2002) which give an optimal number of six or seven. For each panel, the first and second column report the estimated risk premium and its t-statistics based on the Newey-West standard errors. The third column R_g^2 reports the R^2 of the third pass, namely the R^2 of the factors (Comdt, Basis, Mom, BasMom, VOL) explained by the latent factors.

Figure A.1: **The Implied Volatility of Oil and Gold**



The figure shows the 30-day implied volatility of crude oil (upper panel) and gold (lower panel). In the upper panel we also contrast our index with the CBOE crude oil volatility index (OVX). In the lower panel, we coplot the CBOE gold volatility index (GVZ). The sample period is from January 1990 to October 2014.

References

Acharya, V. V., L. A. Lochstoer, and T. Ramadorai (2013). Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics 109*(2), 441–465.

Adam, T. R. and C. S. Fernando (2006). Hedging, speculation, and shareholder value. *Journal of Financial Economics 81*(2), 283–309.

Adrian, T., E. Etula, and T. Muir (2014). Financial intermediaries and the cross-section of asset returns. *Journal of Finance 69*(6), 2557–2596.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets 5*, 31–56.

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

Bai, Y. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Bakshi, G., X. Gao, and A. Rossi (2016). Understanding the sources of risk underlying the cross-section of commodity returns. *Working Paper*.

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

Bakshi, G. and G. Panayotov (2013). Predictability of currency carry trades and asset pricing implications. *Journal of Financial Economics 110*(1), 139–163.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2001). Maxing out: Stocks as lotteries and the cross section of expected returns. *Journal of Financical Economics 99*(2), 427–446.

Barberis, N. and M. Huang (2001). Mental accounting, loss aversion, and individual stock returns. *Journal of Finance 56*(4), 1247–1292.

Barone-Adesi, G. and R. E. Whaley (1987). Efficient analytic approximation of american option values. *Journal of Finance 42*(2), 301–320.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics 3*(1-2), 167–179.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*(11), 4463–4492.

Boons, M. and M. P. Prado (2016). Basis-momentum in the futures curve and volatility risk. *Working Paper*.

Boyer, B., T. Mitton, and Vorkink (2010). Expected idiosyncratic skewness. *Review of Financial Studies 23*(1), 169–202.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies 22*(6), 2201–2238.

Burnside, C. (2011). The cross-section of foreign currency risk premia and consumption growth risk: Comment. *American Economic Review 101*(7), 3456–76.

Cheng, I.-h., A. Kirilenko, and W. Xiong (2015). Convective risk flows in commodity futures markets. *Review of Finance 19*(5), 1733–81.

Christoffersen, P. and X. Pan (2015). Oil volatility risk and expected stock returns. *Working Paper*.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Corte, P. D., T. Ramadorai, and L. Sarno (2016). Volatility risk premia and exchange rate preditability. *Journal of Financial Economics 120*(1), 21–40.

De Roon, F. A., T. E. Nijman, and C. Veld (2000). Hedging pressure effects in futures markets. *The Journal of Finance 55*(3), 1437–1456.

Dhume, D. (2011). Essays on measuring and explaining commodities returs. *Harvard University Dissertation*.

Dusak, K. (1973). Futures trading and investor returns: An investigation of commodity market risk premiums. *Journal of Political Economy 81*(6), 1387–1406.

Erb, C. and C. Harvey (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal 62*(2), 69–97.

Etula, E. (2013). Broker-dealer risk appetite and commodity returns. *Journal of Financial Econometrics 11*(3), 486–521.

Fama, E. and K. French (1987). Commodity futures prices: Some evidence on forecast power. *The Journal of Business 60*(1), 55–73.

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

Fontaine, J.-S. and R. Garcia (2012). Bond liquidity premia. *Review of Financial Studies 25*(4), 1207–1254.

Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics 111*(1), 1–125.

Fung, W. and D. A. Hsieh (2001). The risk in hedge fund strategies: Theory and evidence from trend followers. *Review of Financial Studies 14*(2), 313–341.

Gao, G., P. Gao, and Z. Song (2017). Do hedge funds exploit rare disaster concerns? *Review of Financial Studies forthcoming.*

Gao, G., X. Lu, and Z. Song (2017). Tail risk concerns everywhere. *Management Science forthcoming.*

Giglio, S. and D. Xiu (2017). Inference on risk premia in the presence of omitted factors. *Working Paper*.

Gorton, G., F. Hayashi, and G. Rouwenhorst (2013). The fundamentals of commodity futures returns. *Review of Finance 17*(1), 35–105.

Gorton, G. and G. Rouwenhorst (2006). Facts and fantasies about commodity futures. *Financial Analysts Journal 62*(2), 47–68.

Henderson, B. J., N. D. Pearson, and L. Wang (2014). New evidence on the financialization of commodity markets. *Review of Financial Studies 28*(5), 1285–1311.

Hicks, J. M. (1930). Value and capital. *Cambridge, U.K.: Oxford University Press*.

Hillert, A., H. Jacobs, and S. Müller (2014). Media makes momentum. *Review of Financial Studies 27*(12), 3467–3501.

Hirschleifer, D. (1990). Hedging pressure and futures price movements in a general equilibrium model. *Review of Financial Studies 58*(2), 411–428.

Hong, H., T. Lim, and J. C. Stein (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 M. Yogo (2012). What does futures market interest tell us about the macroeconomy and asset prices? *The Journal of Financial Economics 105*(3), 473–490.

Hou, K. and R. K. Loh (2016). Have we solved the idiosyncratic volatility puzzle? *The Journal of Financial Economics 121*(1), 167–194.

Huang, W., Q. Liu, S. G. Rhee, and L. Zhang (2010). Return reversals, idiosyncratic risk, and expected returns. *The Review of Financial Studies 23*(1), 147–168.

Jiang, G. J. and Y. S. Tian (2005). The model-free implied volatility and its information content. *Review of Financial Studies 18*(4), 1305–1342.

Kan, R. and C. Robotti (2009). Model comparison using the hansen-jagannathan distance. *Review of Financial Studies 22*(9), 3449–3490.

Keynes, J. M. (1930). A treatise on money. *Macmillan: London*.

Kilic, M. and I. Shaliastovich (2017). Good and bad variance premia and expected returns. *Management Science forthcoming.*

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics 96*(2), 175–194.

Lustig, H., N. Roussanov, and A. Verdelan (2011). Common risk factors in currency markets. *Review of Financial Studies 24*(11), 3731–3777.

Marshall, B. R., N. H. Nguyen, and N. Visaltanachoti (2012). Commodity liquidity measurement and transaction costs. *Review of Financial Studies 25*(2), 599–638.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012). Carry trades and global foreign exchange volatility. *The Journal of Finance 67*(2), 681–718.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica 41*.

Merton, R. C. (1987). Presidential address: A simple model of capital market equilibrium with incomplete information. *The Journal of Finance 42*(3), 485–510.

Moreira, A. and T. Muir (2017). Volatility-managed portfolios. *Journal of Finance forthcoming*.

Mou, Y. (2011). Limits to arbitrage and commodity index investment: Front-running the goldman roll. *Working Paper*.

Neuhierl, A. and A. J. Thompson (2016). Liquidity timing in commodity markets and the impact of financialization. *Working Paper*.

Ng, V. K. and S. C. Pirrong (1994). Fundamentals and volatility: Storage, spreads, and the dynamics of metals prices. *Journal of Business 67*.

Pan, X. and S. B. Kang (2015). Commodity variance risk premia and expected futures returns: Evidence from the crude oil market. *Working Paper*.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies 22*(1), 435–480.

Plante, M. (2015). Opec in the news: The effect of opec on oil price uncertainty. *Working Paper*.

Prokopczuk, M. and C. W. Simen (2014). Variance risk premium in commodity markets. *Working Paper*.

Szymanowska, M., F. de Roon, T. Nijman, and R. van den Goorbergh (2014). An anatomy of commodity futures risk premia. *The Journal of Finance 69*(1), 453–482.

Tang, K. and W. Xiong (2012). Index investment and the financialization of commodities. *Financial Analyst Journal 68*, 54–57.

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