



RESEARCH ARTICLE



WILEY

# Economic uncertainty, trading activity, and commodity futures volatility

Sumudu W. Watugala<sup>1,2</sup>

<sup>1</sup>Cornell University, Ithaca, New York

<sup>2</sup>Oxford-Man Institute of Quantitative Finance, University of Oxford, Oxford, UK

**Correspondence**

Sumudu W. Watugala, Cornell University, Ithaca, NY 14853.

Email: sumudu@cornell.edu

**Funding information**

Oxford-Man Institute

## Abstract

This paper investigates the dynamics of commodity futures volatility. I derive the variance decomposition for the futures basis and show unexpected excess returns result from new information about expected future interest rates, convenience yields, and risk premia. Measures of uncertainty in economic conditions have significant predictive power for realized volatility of commodity futures returns, after controlling for lagged volatility, returns, commodity index trading, hedging pressure, and other trading activity, even during the so-called “index financialization” period. During this period, hedge fund performance predicts volatility in grain commodities, which are affected by the US ethanol mandate.

## KEY WORDS

basis, commodity markets, financialization, futures, uncertainty, volatility

## JEL CLASSIFICATION

F36, G12, G13, G15, Q02

## 1 | INTRODUCTION

This paper investigates the time variation in commodity futures volatility and the factors explaining its dynamics. I analyze the impact of concentration and increased emerging market demand on commodity markets. This study builds on Bloom (2014), who presents evidence that emerging markets and recessionary periods are strongly associated with economic uncertainty, and Gabaix (2011), who shows the impact on aggregate volatility from power laws in size distributions. This paper adds to the literature on what explains fluctuations in volatility (see, e.g., Bloom, 2014; Engle & Rangel, 2008; Gabaix, 2011; Roll, 1984; W.G. Schwert, 1989), while also contributing to the current debate on commodity price dynamics and potential distortions arising from market frictions.<sup>1</sup> In particular, I examine how macroeconomic uncertainty and financial frictions are related to realized volatility in commodity futures markets.

Volatility dynamics are a key consideration in strategy formation for hedging, derivatives trading, and portfolio optimization. Moreover, producers and consumers benefit from understanding the factors explaining price fluctuations when evaluating real options embedded in investment choices (E.S. Schwartz, 1997). Distortions can lead to underinvestment or overinvestment, and even transitory deviations from fundamentals can lead to the long-term misallocation of resources (see, e.g., Bernanke, 1983; Bloom, Bond, & VanReenen, 2007). This is especially important

<sup>1</sup>For recent studies on factors affecting commodity markets, see, inter alia, US Senate Permanent Subcommittee on Investigations (2009, 2014), Tang and Xiong (2012), Irwin and Sanders (2012), Singleton (2014), Basak and Pavlova (2013, 2015), and Henderson, Pearson, and Wang (2015) on commodity derivative and exchange-traded fund trading and the involvement of banks in commodity markets; Etula (2013) and Acharya, Lochstoer, and Ramadorai (2013) on broker-dealer capacity for risk taking; Kilian (2009), Chen, Rogoff, and Rossi (2010), and Atugala (2017) on producer and consumer shocks; and Roberts and Schlenker (2013) on changes to regulation.

when there are nonconvex production functions and large fixed costs to entry and expansion (e.g., a copper producer considering the development of a new mine or a manufacturer considering the opening of a new factory that uses raw commodities as inputs). Uncertainty also increases the difficulty for both producers and consumers when formulating optimal hedging strategies, potentially leading to higher volatility in their cash flows. This can cause higher borrowing costs and lower debt in the presence of nonzero costs to bankruptcy and default, which can in turn lead to lower firm values. Consequently, understanding the relationship between volatility and economic factors is a first-order consideration. For commodities with derivative markets that are illiquid, opaque, or have little market depth or limited expirations, the findings in this paper can provide a useful aid to price discovery, real option evaluation, and risk management for end-users as well as financial investors. A better understanding of these futures return dynamics also enables policy-makers to consider the impact of possible market intervention and evaluate regulatory options aimed at achieving a desired welfare objective.<sup>2</sup>

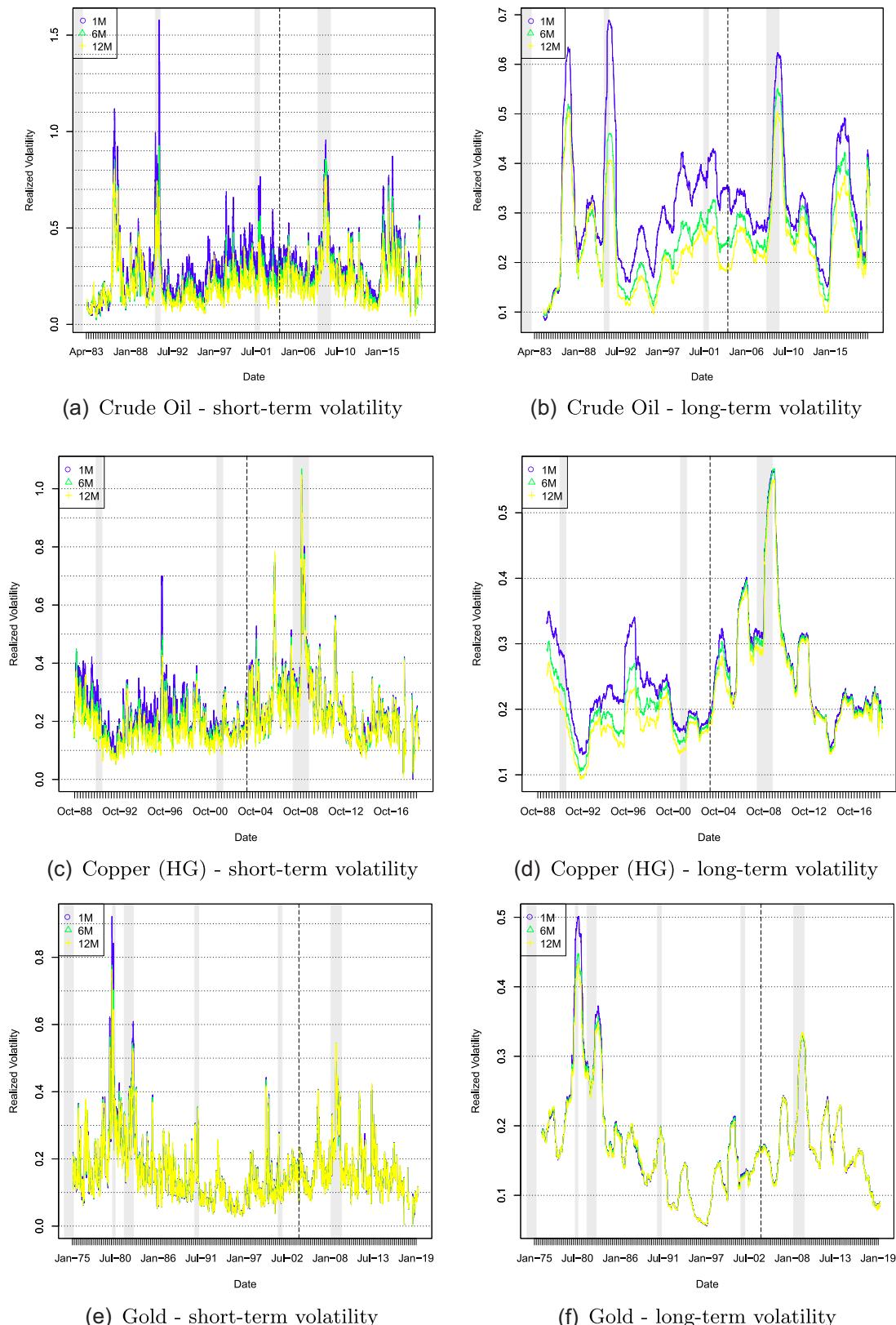
I collate data on 22 major commodity futures markets and analyze the extent to which commodity volatility is related to fundamentals such as inflation uncertainty, while controlling for financial frictions introduced by changing market structure and commodity index trading. I find predictability in commodity futures volatility using variables capturing macroeconomic uncertainty, with adjusted  $R^2$  gains of over 10% over the baseline specification. Moreover, controlling for recession periods further increases the explanatory power of the main predictive regressions by over 13%. These reflect economically significant gains for an investor, particularly those engaged in hedging, in evaluating real options embedded in investment choices, or in trading portfolios of derivatives.

I derive the variance decomposition for futures, building on Working (1949), Campbell and Shiller (1988), and Campbell (1991), to show how unexpected changes to the excess basis return of a commodity future are driven by changes to the expectation of future interest rates, convenience yield (the net benefit of holding the underlying physical commodity), and risk premia. These expectations are updated in response to new information about the future state of the economy (e.g., news on inflation and other variables related to the business cycle) and future commodity supply and demand (e.g., news about the economic health of commodity consumers and frictions to producer hedging). Similar to the analysis of stock market volatility by Engle and Rangel (2008), using this decomposition as the theoretical motivation, I examine the time variation in the relationship between commodity volatility and shocks to relevant factors.

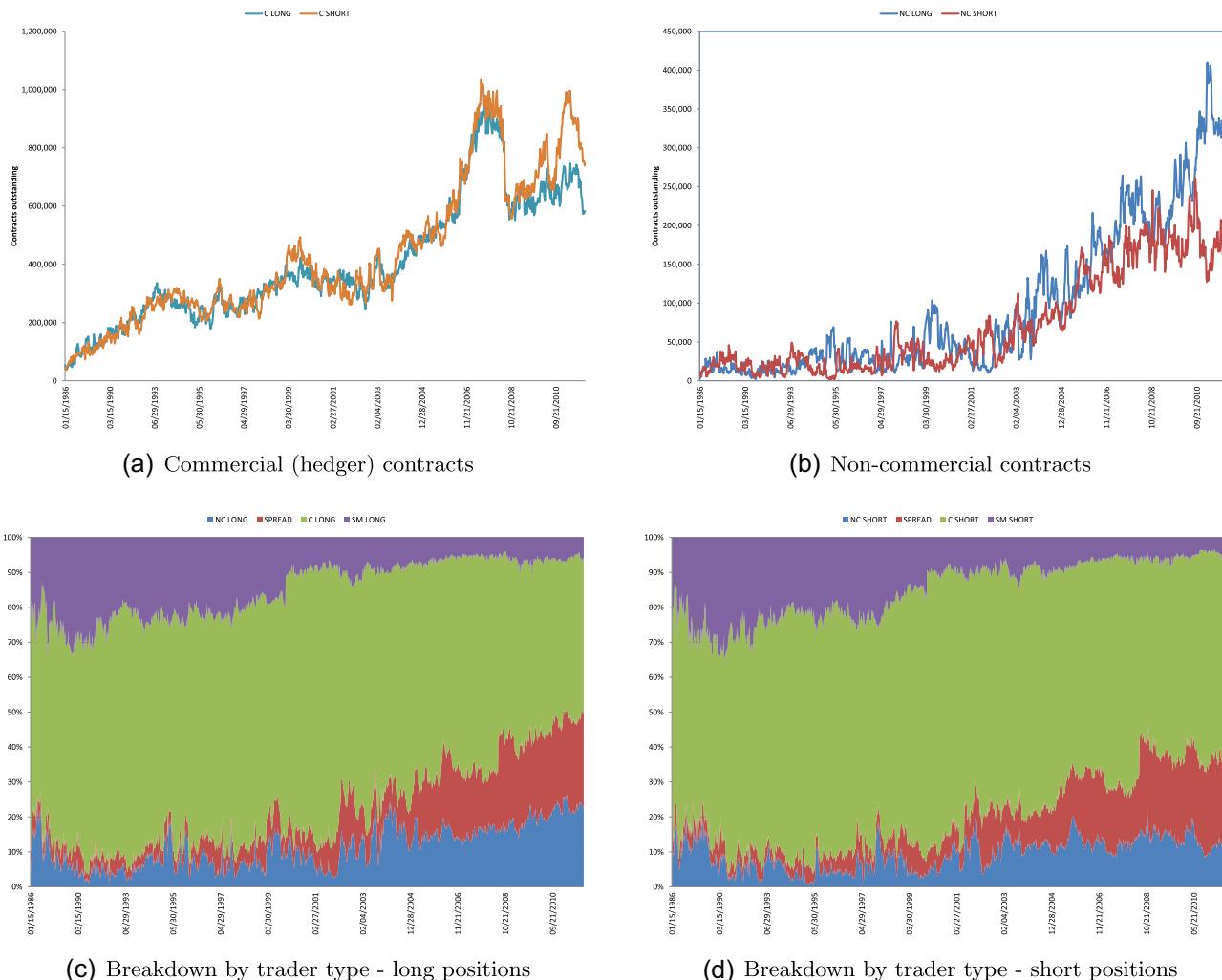
I find that there are significant fluctuations in both the realized volatility and the realized correlations of futures returns for the commodities analyzed in this study (e.g., Figures 1–6). This is true at different horizons corresponding to different holding periods, and throughout the entire trading history of a contract (e.g., beginning in the 1960s for most grain commodities, April 1983 for crude oil, etc.). Large fluctuations in price and volatility occurred for the commodities in the sample even before the popularization of commodity index and exchange-traded fund trading during the “index financialization” period, commonly identified in the literature as beginning by January 2004 (Basak & Pavlova, 2015; Hamilton & Wu, 2015; Singleton, 2014; Tang & Xiong, 2012). I analyze the determinants of this variation in volatility, selecting variables that capture the variation in global macroeconomic conditions, commodity supply-demand, and market frictions based on theory and past empirical studies on commodity risk premia (see, e.g., Acharya et al., 2013; Hong & Yogo, 2012). I add to this from the literature on analyzing the determinants of the realized volatility of financial assets; see, for example, Roll (1984) for an early study on the volatility dynamics of a commodity derivative; W.G. Schwert (1989) on understanding the time variation in equity volatility; Engle and Rangel (2008) on relating low-frequency macroeconomic factors to realized volatility in global equity market indices; Gabaix (2011) and Kelly, Lustig, and Van Nieuwerburgh (2013) on the granular origins of volatility; and Bloom et al. (2007), Bloom (2009, 2014), and Jurado, Ludvigson, and Ng (2014) on uncertainty and its relationship to volatility.

Part of the recent debate on commodity price fluctuations attempts to distinguish between the impact on commodity futures markets from changing market structure and investor composition as opposed to changing macroeconomic fundamentals and supply-demand dynamics (see Footnote 1). Several recent studies find in favor of the “financialization” or trader activity argument, citing, among other evidence, high commodity volatility and correlation (between crude oil prices and other financial markets) in the past decade (especially, after January 2004), when commodity index trading volumes increased substantially. However, I find that the commodity futures volatility observed during the past decade may in fact be largely in line with the high levels of futures volatility observed during

<sup>2</sup>See, for example, Sen (1983) on the direct and potentially catastrophic consequences of commodity price dynamics.



**FIGURE 1** Annualized realized volatility of daily returns of crude oil, copper, and gold futures maturing at 1, 6, and 12 months. (a) Crude oil—short-term volatility; (b) crude oil—long-term volatility; (c) copper (HG)—short-term volatility; (d) copper (HG)—long-term volatility; (e) gold—short-term volatility; and (f) gold—long-term volatility. It is noted that short-term volatility plots depict the annualized standard deviation for the previous month, while long-term volatility depicts the annualized standard deviation for the previous 12 months. The shaded areas highlight the NBER recession periods. The dotted line marks January 2004, commonly used as the start of increased trading in commodity indexes [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 2** The evolution of contract positions for crude oil futures broken down by trader type. (a) Commercial (hedger) contracts; (b) noncommercial contracts; (c) breakdown by trader type—long positions; and (d) breakdown by trader type—short positions. Source: CFTC Commitment of Traders reports [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

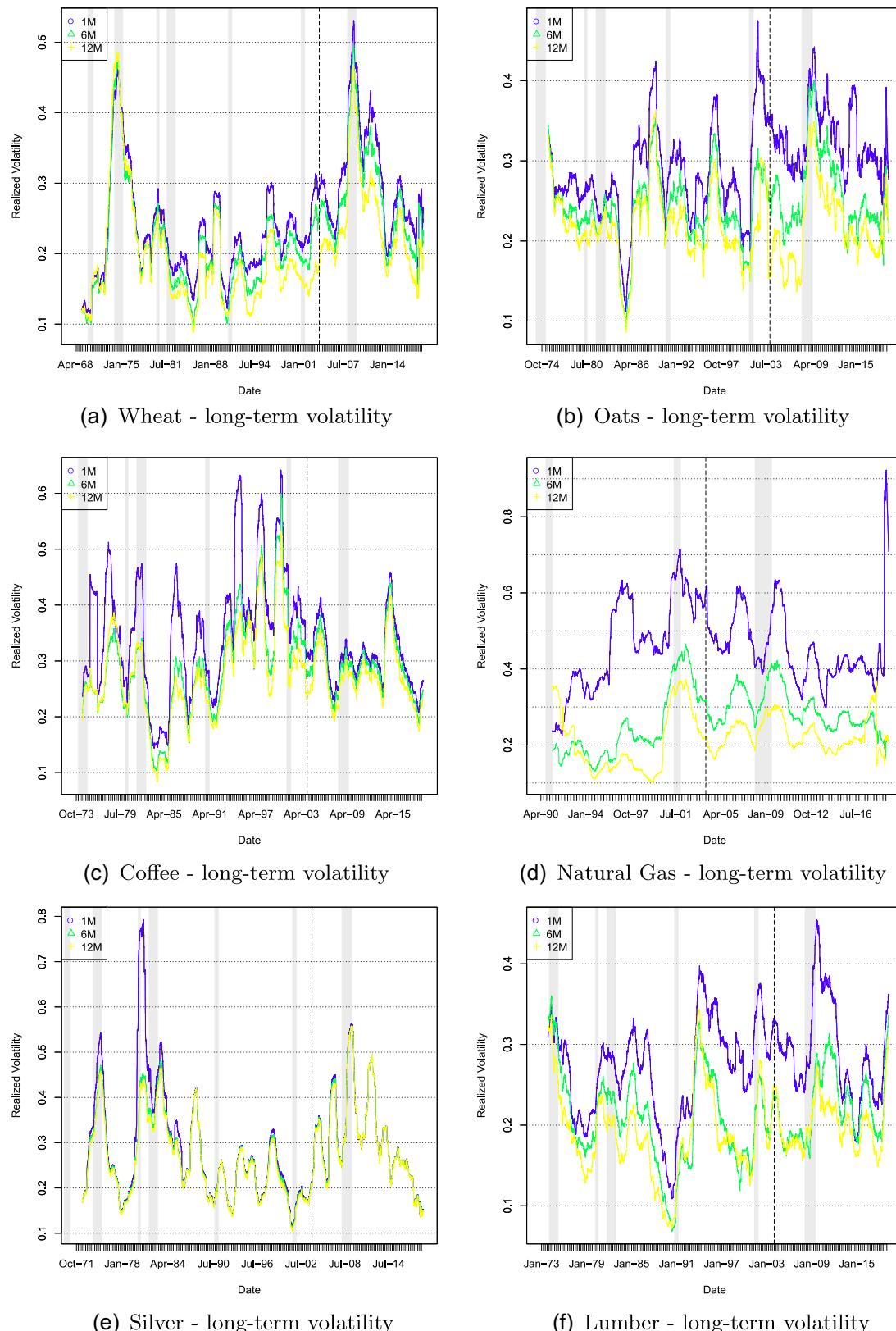
past periods of financial crisis and geopolitical uncertainty. Similarly, correlation levels show significant time variation over the full trading history of commodity futures (e.g., Figures 5 and 6).

The remainder of this paper is organized as follows. The next section presents the research framework, including the theoretical motivation and empirical methodology underpinning this study. Section 3 describes the data and variables used in the analyses. Section 4 presents the results from the main empirical analysis. The final section concludes.

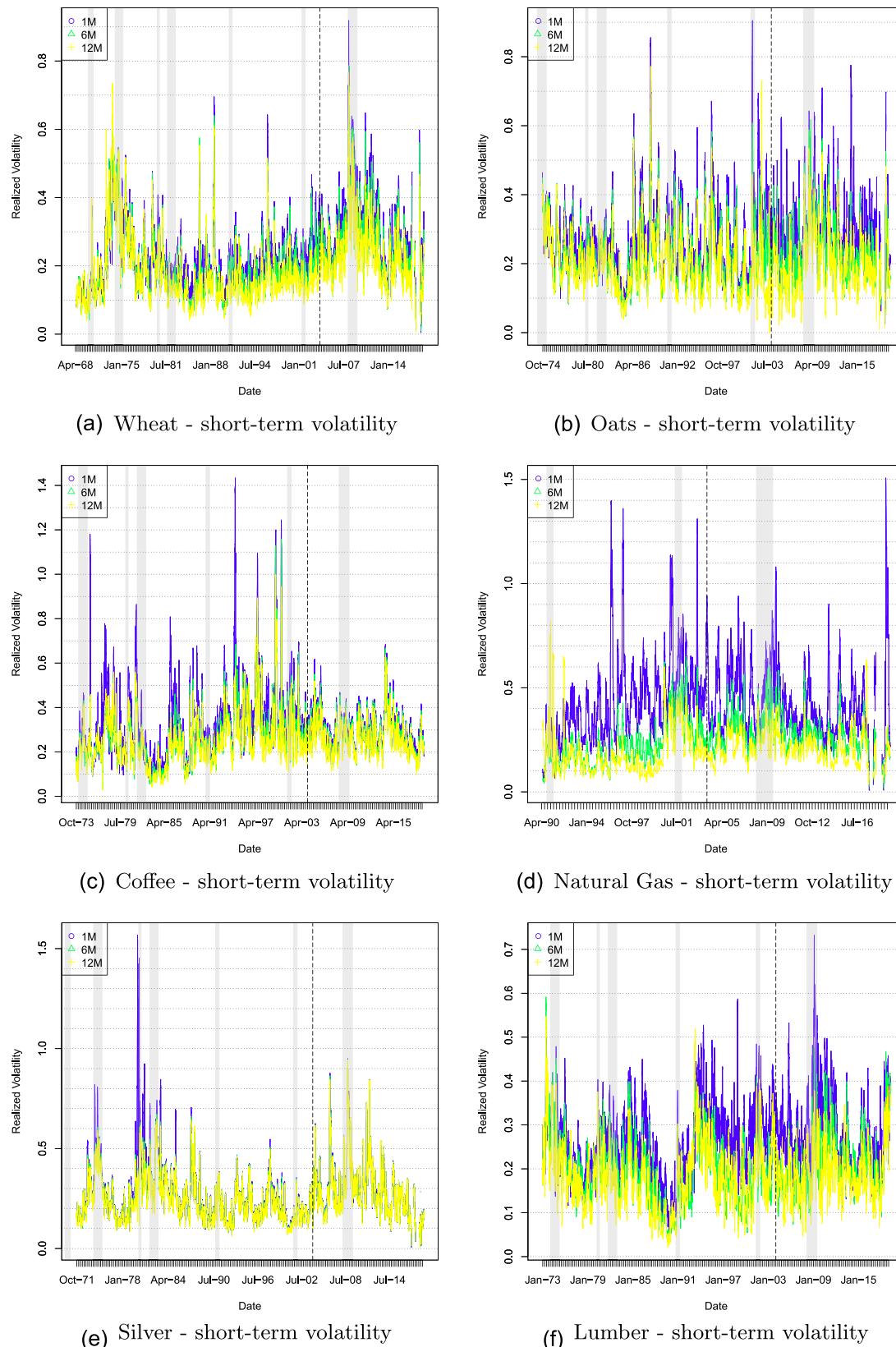
## 2 | RESEARCH FRAMEWORK

### 2.1 | Commodity futures volatility

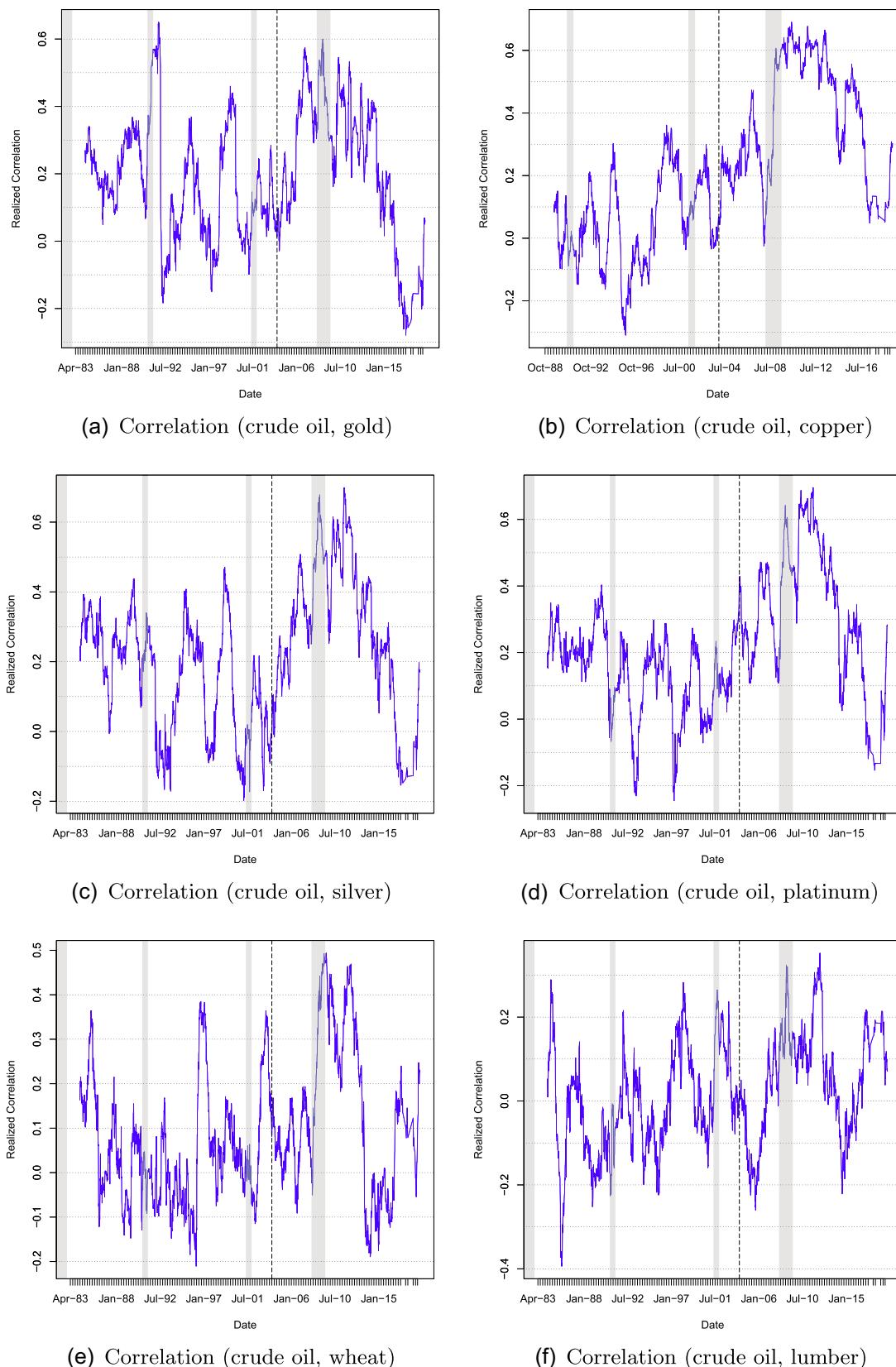
To understand the sources of variation in commodity futures returns, I build on present value models that show how changes in the current price of financial assets react to future changes to the underlying fundamentals. The stock variance decomposition presented in Campbell and Shiller (1988) and Campbell (1991) is widely used to identify the sources of financial asset volatility. This decomposition relates unexpected equity returns to news events that change expectations of future cash flows (stock dividends) and discount rates. Campbell and Ammer (1993) present the equivalent result for bond yields. A similar decomposition can be derived for commodity futures in terms of its basis. To understand this correspondence for a future on a storable commodity, begin with the no-arbitrage pricing formula for its futures price (Brennan, 1958; Kaldor, 1939; E.S. Schwartz, 1997; Working,



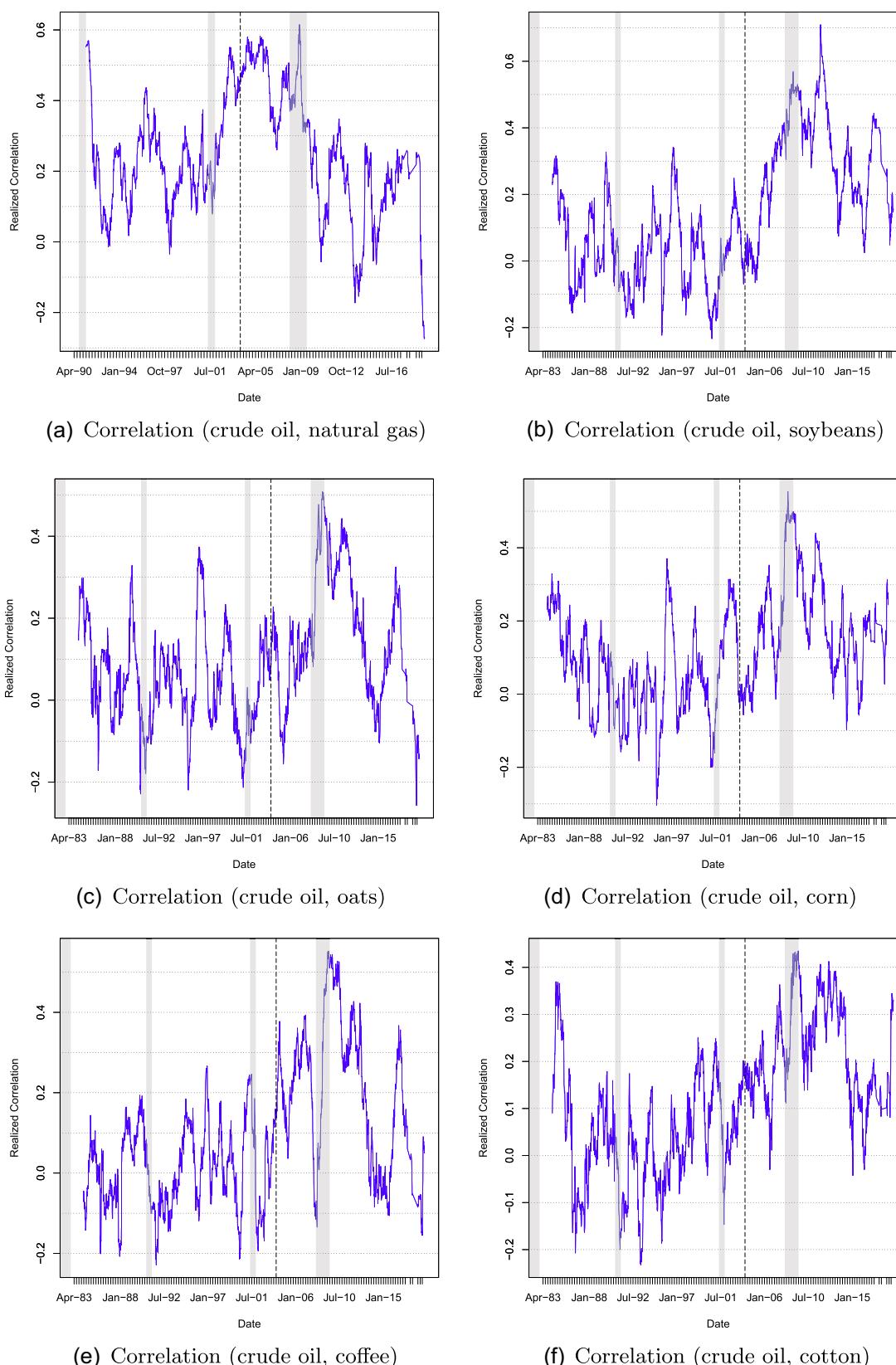
**FIGURE 3** Annualized (long-term) realized volatility of the daily returns for wheat, oats, coffee, natural gas, silver, and lumber futures maturing at 1, 6, and 12 months. (a) Wheat—long-term volatility; (b) oats—long-term volatility; (c) coffee—long-term volatility; (d) natural gas—long-term volatility; (e) silver—long-term volatility; and (f) lumber—long-term volatility. It depicts the time series of the standard deviation for the previous 12 months. The shaded areas highlight the NBER recession periods. The dotted line marks January 2004, commonly used as the start of increased trading in commodity indexes [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 4** Annualized (short-term) realized volatility of the daily returns for wheat, oats, coffee, natural gas, silver, and lumber futures maturing at 1, 6, and 12 months. (a) Wheat—short-term volatility; (b) oats—short-term volatility; (c) coffee—short-term volatility; (d) natural gas—short-term volatility; (e) silver—short-term volatility; and (f) lumber—short-term volatility. It depicts the time series of the standard deviation for the previous month. The shaded areas highlight the NBER recession periods. The dotted line marks January 2004, commonly used as the start of increased trading in commodity indexes [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 5** Pairwise correlation between returns of the crude oil 3-month future and gold, copper, silver, platinum, wheat, and lumber futures. (a) Correlation (crude oil, gold); (b) correlation (crude oil, copper); (c) correlation (crude oil, silver); (d) correlation (crude oil, platinum); (e) correlation (crude oil, wheat); and (f) correlation (crude oil, lumber). It is noted that each date shows the corresponding correlation for the previous 12-month period calculated on 3-day returns. The shaded areas highlight the NBER recession periods. The dotted line marks January 2004, commonly used as the start of increased trading in commodity indexes [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 6** Pairwise correlation between returns of the crude oil 3-month future and natural gas, soybeans, oats, corn, coffee, and cotton futures. (a) Correlation (crude oil, natural gas); (b) correlation (crude oil, soybeans); (c) correlation (crude oil, oats); (d) correlation (crude oil, corn); (e) correlation (crude oil, coffee); and (f) correlation (crude oil, cotton). It is noted that each date shows the corresponding correlation for the previous 12-month period calculated on 3-day returns. The shaded areas highlight the NBER recession periods. The dotted line marks January 2004, commonly used as the start of increased trading in commodity indexes [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

1933, 1949),  $F_{t,T} = S_t e^{(r-y)(T-t)}$ , where  $F_{t,T}$  is the futures price at time  $t$  of a unit of the commodity delivered at time  $T$ ,  $S_t$  is the spot price,  $r$  is the risk-free rate, and  $y$  is the convenience yield. Further,  $y$  can be decomposed into the “benefit” from holding the physical commodity,  $b$ , net of the storage (or carry) cost rate  $m$ ,  $y = b - m$ ;  $r = \pi + \psi$ , where  $\pi$  is the inflation rate and  $\psi$  the real interest rate. This decomposition and analysis that follow are applicable to any type of future, with the interpretation of  $y$  differing depending on the net benefit to holding the underlying asset, for example, replace  $y$  with dividend yield  $d$  for stock futures or with the foreign currency interest rate  $r^f$  for currency futures.<sup>34</sup>

Consider the discrete-time version of this formula, now with time-dependent  $r$  and  $y$ : the price at time  $t$  of a future expiring in  $n$  periods,

$$F_{n,t} = S_t \frac{(1 + R_{n,t})^n}{(1 + Y_{n,t})^n}, \quad (1)$$

$$(1 + Y_{n,t}) = \left( \frac{1 + B_{n,t}}{1 + M_{n,t}} \right). \quad (2)$$

Denote the log price at time  $t$  of a future expiring in  $n$  periods as  $f_{n,t}$  and the corresponding log spot price as  $s_t$ . Accordingly, the log price of the same future at time  $t+1$  is  $f_{n-1,t+1}$ , now with  $n-1$  periods to expiry, with an associated log spot price  $s_{t+1}$ . Define,  $r_{n,t} \equiv \ln(1 + R_{n,t}) = \pi_{n,t} + \psi_{n,t}$  and  $y_{n,t} \equiv \ln(1 + Y_{n,t}) = b_{n,t} - m_{n,t}$ . Note that  $r_{n,t}$  and  $y_{n,t}$  are *per period* rates at time  $t$ , corresponding to the interest and convenience yield for the next  $n$  periods. Using this notation, I can define the basis,  $p_{n,t}$ ,

$$f_{n,t} = s_t + n(r_{n,t} - y_{n,t}), \quad (3)$$

$$\begin{aligned} p_{n,t} &\equiv f_{n,t} - s_t \\ &= n(r_{n,t} - y_{n,t}). \end{aligned} \quad (4)$$

We can define  $\delta_{n,t+1}$ , the change in basis from  $t$  to  $t+1$ , and  $x_{n,t+1}$ , the return in excess of the cost-of-carry,

$$\begin{aligned} \delta_{n,t+1} &\equiv p_{n-1,t+1} - p_{n,t} \\ &= (n-1)(r_{n-1,t+1} - y_{n-1,t+1}) - n(r_{n,t} - y_{n,t}), \end{aligned} \quad (5)$$

$$x_{n,t+1} \equiv \delta_{n,t+1} + (r_{1,t} - y_{1,t}). \quad (6)$$

As discussed further in Sections 2.2, there can be deviations from the no-arbitrage condition due to nondiversifiable risks or market frictions such as producer hedging pressure and borrowing constraints (see, e.g., Acharya et al., 2013; Cootner, 1960; Hirshleifer, 1988, 1990; Keynes, 1930; deRoon, Nijman, & Veld, 2000). Further, the convenience yield associated with a commodity has features distinct from the yields on other financial contracts (see, e.g., Brennan, Williams, & Wright, 1997). In addition to such basis risks,  $x_{n,t+1}$  also captures the part of the futures risk premia due to deviations from the expectations hypothesis in the interest rate term structure, as shown in Appendix A.1, Equation (A1).

Given that  $p_{0,t} = 0$  for all  $t$ , solving (5) forward (for  $p_{n,t}$ ,  $p_{n-1,t+1}$ ,  $p_{n-2,t+2}, \dots, p_{1,t+n-1}$ ) until the maturity date  $t+n$ , and taking expectations at time  $t$  yields

$$p_{n,t} = -[\delta_{n,t+1} + \delta_{n-1,t+2} + \dots + \delta_{1,t+n}] \quad (7)$$

<sup>34</sup>This decomposition is exact for the forward price. Due to the mark-to-market gains and losses of the corresponding futures contract, differences can occur between the forward and future prices unless interest rates are deterministic.

<sup>4</sup>Several studies investigate the commodity convenience yield. Casassus and Collin-Dufresne (2005) nest several other models (including Gibson & Schwartz, 1990; E.S. Schwartz, 1997), concluding that convenience yield is increasing in the spot price, interest rates, and the extent to which the underlying commodity is used for production purposes.

$$= -E_t \sum_{i=0}^{n-1} \delta_{n-i,t+i+1}. \quad (8)$$

Equation (7) must hold ex post and ex ante, so taking its expectation yields Equation (8). Substituting (8) back into (5) gives the decomposition:

$$\delta_{n,t+1} - E_t \delta_{n,t+1} = -(E_{t+1} - E_t) \sum_{i=1}^{n-1} \delta_{n-i,t+i+1}. \quad (9)$$

Equation (6) can be substituted into (9) to obtain its unexpected change:

$$x_{n,t+1} - E_t x_{n,t+1} = (E_{t+1} - E_t) \left\{ \sum_{i=1}^{n-1} r_{1,t+i} - \sum_{i=1}^{n-1} y_{1,t+i} - \sum_{i=1}^{n-1} x_{n-i,t+i+1} \right\}. \quad (10)$$

Equation (10) means that, if there is an unexpected increase in the excess basis return, either expected future interest rates are higher, expected future convenience yields are lower, or future risk premia are lower. When the assumption that both the expectations hypothesis for the term structure of interest rates and the theory of storage hold exactly,  $E[\delta_{n,t+1}] = y_{1,t} - r_{1,t}$  for all  $n > 0$ , the third summation (of expected future excess basis returns) in (10) is zero. When this assumption is relaxed, the decomposition captures the risk premia reflecting the maturity and spot risk in interest rates and convenience yields. If we further decompose the excess basis return,  $x_{n,t+1}$ , to separate out the excess return due to the interest rate term structure (i.e., due to deviations from the expectations hypothesis), we can characterize the excess return purely due to the convenience yield and commodity risk premia (see Equations (A2) and (A3)) in Appendix A.1.

The decomposition can be rewritten explicitly in terms of news events relating to convenience yield, the risk-free rate, and risk premia:

$$x_{n,t+1} - E_t x_{n,t+1} = \eta_{t+1}^r - \eta_{t+1}^y - \eta_{t+1}^x. \quad (11)$$

Equation (11) shows that unexpected changes to the futures risk premium are due to innovations in the future expected convenience yields, interest rates, and excess basis returns. These expectations are updated in response to new information about the future state of the economy (e.g., the level and volatility of inflation and real interest rates) and commodity supply-demand (e.g., inventory levels and the economic health of consumers). A positive shock to future convenience yields (the net benefit from holding the underlying spot commodity) or risk premia has a negative effect on the futures risk premium. The volatility of the excess basis return is driven by unexpected news affecting interest rates, convenience yield, and risk premia. More explicitly, with correlated components,

$$\text{Var}(x_{n,t+1}) = \text{Var}(\eta_t^r) + \text{Var}(\eta_t^y) + \text{Var}(\eta_t^x) - 2\text{Cov}(\eta_t^r, \eta_t^y) - 2\text{Cov}(\eta_t^r, \eta_t^x) + 2\text{Cov}(\eta_t^y, \eta_t^x). \quad (12)$$

Engle and Rangel (2008) show that it is straightforward to model the unexpected return of a financial asset decomposed in this manner in terms of its stochastic volatility as

$$x_{n,t+1} - E_t x_{n,t+1} = \sigma_t \epsilon_t, \text{ where } \epsilon_t | \Omega_{t-1} \sim N(0, 1). \quad (13)$$

Given (11) and (13), we see that the stochastic volatility,  $\sigma_t$ , is driven by news on the future state of the economy and commodity supply-demand that directly impact convenience yield and interest rates. Models commonly used to estimate  $\sigma_t$  for financial assets and their implementation for commodity futures in this study are discussed in Section 3.2. Many studies attempting to understand equity risk premium dynamics decompose unexpected returns into  $K$  observable news sources or risk factors which affect expectations of future discount rates and cash flows to equity, that is, for the unexpected excess equity return,  $e_t - E_{t-1} e_t = \eta_t^d - \eta_t^r - \eta_t^e = \sum_{k=1}^K \beta_k \lambda_{k,t}$ . The equivalent for commodities should use the appropriate information set given the decomposition in (11).

## 2.2 | Impact of market frictions and limits to arbitrage

Deviations from the decomposition derived from no-arbitrage pricing conditions can occur for a variety of reasons in imperfect markets with frictions (e.g., information asymmetry or disagreement, limits to arbitrage via capital constraints) or due to the natural scarcity of the underlying asset, which is especially important for commodities, an asset class that has historically shown many episodes of market cornering and manipulation.<sup>5</sup> Such conditions can cause Equation (1) to no longer hold exactly for all investors in the market. In Equation (10), these deviations are captured in the third term.

The limits to arbitrage and its related literature look at standard theoretical asset pricing models with strong assumptions on the existence of perfect frictionless markets relaxed. Shleifer and Vishny (1997) show that since arbitrage in practice requires capital and is inherently risky, asset prices will diverge from fundamental values under a variety of possible conditions when informed arbitrageurs in the market are constrained from eliminating them. Gromb and Vayanos (2002) find that capital-constrained arbitrageurs may take more or less risk than in a situation where they face perfect capital markets, leading to equilibrium outcomes that are not Pareto optimal. Yuan (2005) uses a modified Grossman and Stiglitz (1980) framework where a fraction of informed investors face a borrowing constraint, which is a function of the risky asset price (the lower the price, the more constrained the investor), and shows that this can result in asymmetric price movements.

Garleanu, Pedersen, and Potoshman (2009) apply this reasoning to options markets, and consider the case where it is not possible to hedge equity option positions perfectly, leading to demand pressure having an impact on option prices. They show empirically with equity index and single stock data that this helps to explain asset pricing puzzles such as option volatility skewness and relative expensiveness, which are anomalies under the assumptions of the Black–Scholes–Merton model (Black & Scholes, 1973; Merton, 1973).

Basak and Pavlova (2013) model the impact on a stock market from institutional investors whose performance is measured against a benchmark equity index. As this results in institutional investors holding more index stocks than is otherwise optimal, there is demand pressure that boosts index stock prices (and not off-index stock prices). This amplifies the volatility of on-index stock prices and the correlations between them, as well as increasing overall market volatility.

The term financialization, in the context of commodities trading, is generally used to describe the increased noise and uninformed speculative trading (usually with no direct exposure to the underlying commodity) through a range of trading activities including index investment and financial portfolio hedging and rebalancing. Given market frictions, such trading can result in price volatility and correlation between markets to an extent that does not reflect underlying fundamentals (Basak & Pavlova, 2013; Pavlova & Rigobon, 2008). Implicit in the financialization argument is the assumption that there are binding constraints on investors or other significant frictions such as information asymmetries that lead to the persistence of market inefficiencies despite the existence of some informed players in that market. Such frictions render markets incomplete. Under such conditions, financial innovation or the introduction of even redundant assets can change equilibrium allocations and market volatility and efficiency could increase or decrease. Equilibrium outcomes in markets where arbitrageurs are constrained can be inefficient or indeterminate under a range of common market conditions.<sup>6</sup>

Several recent studies examine the predictive relationships between commodities and other markets, and investigate the possible impact of financialization and investor characteristics on commodity markets. Tang and Xiong (2012) find that nonenergy commodities have become increasingly correlated with oil prices, and that this relationship is stronger for constituent commodities of the SP-GSCI and DJ-UBS indices. They link this trend to increased financialization, (mainly via the increased investment in popular commodity indices since the early 2000s), and conclude that the underlying mechanism driving this phenomenon differs from other episodes of commodity price shocks and increased correlation, such as the crisis periods during the 1970s. Singleton (2014) surveys the recent literature in an attempt to explain the impact of trader activity on the behavior of energy markets, particularly crude oil futures prices, and finds futures open interest has important predictive power for crude oil prices, confirming the finding in Hong and Yogo (2012).

Acharya et al. (2013) consider the effect of capital-constrained speculators in a commodity futures market, where producers trade due to hedging needs and link producer default risk to inventories and prices in energy markets. They

<sup>5</sup>For example, Haase and Zimmermann (2013) on the scarcity premium in commodity futures prices and Jovanovic (2013) on the possibility of bubbles in the prices of exhaustible commodities.

<sup>6</sup>See, for example, Cass (1992), Bhamra and Uppal (2006), Basak, Cass, Licari, and Pavlova (2008), and Gromb and Vayanos (2010).

find that when speculator activity is constrained or reduced, the impact of hedging demand increases, that is, unconstrained speculator activity will assist the absorption of producer demand shocks. Etula (2013) also finds that the risk-bearing capacity of broker-dealers is predictive of commodity risk premia.

Based on the literature, I empirically test several hypotheses related to limits to arbitrage and the impact of trading activity on commodity markets. These test if the effects of “financialization” during the period from January 2004 had a discernible impact on market volatility. This requires that other (possibly more informed) market participants were constrained in their capacity to step in and engage in arbitrage trading to correct any mispricing. Any alternate explanation is that increased participation makes commodity markets more efficient and liquid, correcting any mispricings that may have existed previously due to limited participation and illiquidity.

### 3 | DATA AND VARIABLE DEFINITIONS

In this section, I describe the data used in the empirical analysis. I include a variety of factors that are potentially relevant for commodity prices based on theory and past empirical studies (see, among others, Bali, Brown, & Caglayan, 2014; Engle & Rangel, 2008; Hong & Yogo, 2012). Along the lines of the empirical analysis in Roll (1984) and Engle and Rangel (2008), I model the unexpected shocks to economic and financial variables that are potentially related to commodity prices and test the relationship between these variables and commodity futures volatility.

#### 3.1 | Price, returns, and volatility

I use daily closing prices for commodity options and futures obtained from Barchart.com Inc. These commodities are categorized into four groupings (energy, grain, metal, and softs), traded on NYMEX (energy), COMEX (metal), CBOT (grain), CME, CSCE, and NYCE (softs) as shown in Table 1. Options price data, where available, begin on January 2, 2006. I extend futures data history before January 3, 2005 with data from Pinnacle Data Corp. Futures data go back further for most commodities, with the earliest being July 1959 for cotton, cocoa, and all commodities except rough rice in the grain grouping. To obtain the longest time period within a balanced panel without stale prices, the main regressions exclude natural gas, propane, rough rice, soybean oil, and orange juice futures.

I calculate commodity futures returns (from holding and rolling futures) at a fixed maturity point in the term structure (1, 3, 6, 9, 12, 15, 18, 24, and 36 months) using the methodology described in Singleton (2014), and generate realized volatility time series for 1-, 3-, 6-, and 12-month horizons using these fixed-term daily returns.

#### 3.2 | Volatility estimation

Several recent papers have studied the observed behavior of market implied and realized volatilities, and the variation in the volatility risk premium in equity and currency markets. Of these, Engle and Rangel (2008) and Engle, Ghysels, and Sohn (2013) analyze directly the impact of macroeconomic shocks on equity volatility within GARCH-type models that decompose volatility into short-term and long-term components, and identify several macroeconomic variables with significant impact on low- and high-frequency equity volatility. Ang, Hodrick, Xing, and Zhang (2006, 2009) study the cross-sectional variation in risk premia and idiosyncratic volatility and find a significant positive relationship between the two. Campbell, Giglio, Polk, and Turley (2018) include stochastic volatility in an intertemporal CAPM framework and conclude that volatility risk is priced in US stocks and may explain stock return anomalies such as the value premium. Previous empirical studies on market volatility have mainly concentrated on the S&P 500 index, individual US stocks, and currency markets for a variety of reasons including easy access to the relevant data, long time periods, liquidity, coverage in time-strike space (for implied volatility), etc. A similar systematic analysis of commodity volatility remains a potentially rich area for furthering our understanding of these markets.

A flexible, first-pass estimate for the volatility of an asset over a certain period is its realized volatility over that horizon. Similar to the convention for returns in Singleton (2014), I denote the  $d$ -day rolling return of the (fixed-term)  $f$ -month future of commodity  $i$  as  $R_{i,t}^{fFdD}$ . For example, the 5-day rolling return of the (fixed-term) 3-month commodity future at time  $t$  is denoted  $R_{i,t}^{3FSD}$ . Consequently, the realized volatility of  $d$ -day returns of the  $f$ -month commodity future at time  $t$ , over a horizon of  $m$  months, is defined as the annualized standard deviation over that period,  $\sigma_{i,t} \approx Vol_{i,t}^{fFdDmM}$ , where  $d \in \{1, 3, 5, 21\}$  is the frequency in days of the return series used to construct the volatility series and the volatility horizon in months,  $m \in \{1, 3, 6, 9, 12, 15, 18, 24, 36\}$ , with a week, month, and year, defined

**TABLE 1** Commodity derivative contract and trade classification information

Contract code	Exchange code	Traded contract months	Commodity name	Futures data start
<i>Energy</i>				
CL	NYMEX	All months	Crude oil	03/30/1983
HO	NYMEX	All months	Heating oil	11/14/1978
NG	NYMEX	All months	Natural gas	04/04/1990
PN	NYMEX	All months	Propane	01/03/2005
<i>Metal</i>				
GC	COMEX	G   J   M   Q   V   Z	Gold	12/31/1974
SI	COMEX	H   K   N   U   Z	Silver	06/12/1963
HG	COMEX	H   K   N   U   Z	Copper	01/03/1989
PA	NYMEX	H   M   U   Z	Palladium	01/03/1977
PL	NYMEX	F   J   N   V	Platinum	03/04/1968
<i>Grain</i>				
W	CBOT	H   K   N   U   Z	Wheat	07/01/1959
C	CBOT	H   K   N   U   Z	Corn	07/01/1959
O	CBOT	H   K   N   U   Z	Oats	07/01/1959
S	CBOT	F   H   K   N   Q   U   X	Soybeans	07/01/1959
SM	CBOT	F   H   K   N   Q   U   V   Z	Soybean meal	07/01/1959
BO	CBOT	F   H   K   N   Q   U   V   Z	Soybean oil	07/01/1959
RR	CBOT	F   H   K   N   U   X	Rough rice	08/20/1986
<i>Soft</i>				
CT	NYCE	H   K   N   V   Z	Cotton	07/01/1959
OJ	NYCE	F   H   K   N   U   X	Orange juice	02/01/1967
KC	CSCE	H   K   N   U   Z	Coffee	08/16/1972
SB	CSCE	H   K   N   V	Sugar	01/04/1961
CC	CSCE	H   K   N   U   Z	Cocoa	07/01/1959
LB	CME	F   H   K   N   U   X	Lumber	10/01/1969

Note: It shows the 22 underlying commodities in the data set, categorized into four market groupings (energy, metal, grain, and soft). The naming convention for a futures contract is [Contract code][Expiry month code][Last digit of expiry year], for example, on January 5, 2005, the WTI Crude Oil futures contract expiring in December 2008 is “CLZ8” (month codes: F, Jan | G, Feb | H, Mar | J, Apr | K, May | M, Jun | N, Jul | Q, Aug | U, Sep | V, Oct | X, Nov | Z, Dec |).

as 5, 21, and 252 trading days, respectively.<sup>7</sup> The baseline panel regressions use the (nonoverlapping) end-of-month (EOM) volatility of daily returns of the 1-month future,  $Vol_{i,t} = Vol_{i,t}^{1F1DEOM}$  as the dependent variable, except where explicitly stated otherwise. The augmented Dickey–Fuller test (ADF) rejects the existence of a unit root in  $Vol_{i,t}$  for all commodities in the sample (Table A2 in the Supporting Information Appendix). The baseline predictive regressions take the form

$$Vol_{i,t} = \mu + \alpha|r_{i,t-1}| + \beta Vol_{i,t-1} + \mathbf{z}'_{i,t-1}\boldsymbol{\theta} + \eta_{i,t}, \quad (16)$$

where  $r_{i,t} = R_{i,t}^{1F21D}$ ,  $\mathbf{z}_{i,t}$  is a vector of  $K$  (nonnegative) explanatory variables,  $\alpha$ ,  $\beta$ , and the vector  $\boldsymbol{\theta}$  denotes regression coefficients. In Appendix A.2, I discuss the related volatility models and empirical work that attempt to explain realized volatility with economic variables, which inform the framework of my analysis and its future extensions.

Table 2 shows summary statistics for the realized volatility of the commodity futures in this study. Panel A shows the mean and standard deviation of 1-month (“short-term”) and 12-month (“long-term”) realized volatility for the three

<sup>7</sup>By this definition,

$$Vol_{i,t}^{fFdDmM} = \sqrt{\frac{252}{d}} \left[ \frac{1}{N} \sum_{p=1}^N (R_{i,t-N+p}^{fFdD})^2 - \left( \frac{1}{N} \sum_{p=1}^N R_{i,t-N+p}^{fFdD} \right)^2 \right]^{1/2}, \quad (14)$$

$$Vol_{i,t}^{1F1D1M} = \sqrt{\frac{252}{1}} \left[ \frac{1}{21} \sum_{p=1}^{21} (R_{i,t-21+p}^{1F1D})^2 - \left( \frac{1}{21} \sum_{p=1}^{21} R_{i,t-21+p}^{1F1D} \right)^2 \right]^{1/2}, \quad (15)$$

where  $N = m \times (21/d)$  is the number of  $d$ -day return periods in  $m$  months. The factor  $252/d$  annualizes the volatility.

**TABLE 2** Commodity futures volatility summary statistics

	Mean			SD		
	1M	3M	12M	1M	3M	12M
Crude	30.054	26.531	22.033	16.249	13.670	11.775
Heating oil	29.102	25.800	21.814	14.103	11.846	10.270
Natural gas	44.508	34.208	20.631	19.003	13.614	10.150
Gold	16.966	16.846	16.617	9.991	9.460	9.095
Silver	28.114	26.830	26.122	15.884	12.785	12.449
Palladium	28.896	28.258	28.305	14.481	13.604	13.305
Platinum	24.005	23.520	23.049	13.261	11.432	10.760
Copper	25.295	24.794	21.629	11.214	11.228	11.06
Wheat	23.868	22.889	19.641	10.571	10.453	10.335
Corn	20.269	19.881	17.761	10.411	10.069	9.321
Oat	27.939	26.060	20.689	10.978	9.776	9.517
Soybeans	21.065	20.519	18.227	11.248	10.022	9.920
Soybean meal	23.704	22.763	20.010	13.211	12.253	12.205
Soybean oil	24.322	23.790	20.921	10.269	9.919	10.170
Cotton	23.584	21.767	16.582	10.118	8.684	8.764
Coffee	32.138	30.115	25.745	15.885	14.351	11.438
Sugar	38.880	35.493	27.793	18.477	13.884	13.416
Cocoa	29.285	27.959	24.283	9.565	9.126	8.186
Lumber	27.263	24.579	18.635	8.586	7.703	7.426

Note: It shows the summary statistics of volatility of daily returns for each commodity future at 1-month (1M), 3-month (3M), and 12-month (12M) maturities. This includes the mean and standard deviation of short-term (1-month) and long-term (12-month) realized volatility for the full trading history of each commodity until December 31, 2011 (see Table 1 for futures data start dates by commodity). The standard deviation is a measure of the volatility of volatility. In appendix, I show the same summary statistics by decade.

maturity points on the futures curve (1M, 3M, and 12M). Plots of short-term and long-term realized volatility for the entire term structure are shown in Figures 1 and 3 for crude oil, copper, gold, natural gas, wheat, and lumber.

Relative to commodities in energy, grain, and softs, precious metals broadly show little variation on average volatility by contract month. This is also evident in the figures plotting realized volatility for the futures terms structures over time (figures in 1, 3, and 4). This is indicative of parallel shifts to the forward curve being more common for metals than for commodities in other groups. For crude oil, natural gas, wheat, orange juice, and lumber, etc., the contracts in the nearer term are more volatile than longer-dated contracts. This difference is potentially a risk characteristic driven by underlying fundamentals—inventory, storability, and the nature of the demand for a particular commodity. Relative to other commodity groups, metals are highly storable (dense and durable), easy to transport, and less exposed to supply-demand uncertainty due to weather or geopolitics. Casassus and Collin-Dufresne (2005), in addressing the disparities between the dynamics of convenience yields and futures term structure of crude oil and copper versus gold and silver, hypothesize that oil and copper have a primary function as inputs to production, while the latter two commodities are primarily stores of value. In this case, demand shocks driven by the prevailing economic conditions would drive price fluctuations in production commodities to a greater extent, and create greater variation along the term structure.

Table 3 shows that commodities generally exhibit volatility asymmetry in the opposite direction to equities, with significant gamma coefficients all negative. As documented by Bekaert and Wu (2000), Bollerslev and Todorov (2011), and others, equity indices tend to become more volatile as the price drops, to a greater extent than with index price increases, giving rise to positive gamma coefficients in GJR-GARCH(1,1) specifications. The causes commonly cited for this phenomenon in equities include financial and operating leverage effects, time-varying risk premia, and volatility feedback mechanisms. For commodities, volatility increases are generally larger with large price increases, and this effect merits further study. It appears likely that this effect is greater for commodities with increased inventory risk. In that case, such commodities would also show greater variation in the term structure of volatility.

**TABLE 3** GARCH(1,1) and GJR-GARCH(1,1) fits

	GARCH(1,1)				GJR-GARCH(1,1)					
	$\omega$	$\alpha$	$\beta$	LL	$\omega$	$\alpha$	$\beta$	$\gamma$	LL	LRT
S&P 500	0.010	0.068	0.925	-7,692.3	0.017	0.010	0.918	0.117	-7,618.7	147.2
FTSE 100	0.018	0.086	0.904	-8,474.5	0.022	0.028	0.913	0.090	-8,434.9	79.2
USD-GBP	0.003	0.038	0.954	-4,890.1	0.003	0.032	0.953	0.011	-4,888.5	3.2
All commodities	0.003	0.039	0.957	-6,339.4	0.002	0.045	0.959	-0.012	-6,337.4	4.0
Grain	0.017	0.069	0.923	-9,056.4	0.014	0.081	0.929	-0.032	-9,048.8	15.2
Metal	0.011	0.055	0.937	-8,107.8	0.010	0.062	0.938	-0.014	-8,106.1	3.4
Energy	0.052	0.073	0.917	-11,678.0	0.051	0.076	0.919	-0.009	-11,677.0	2.0
Softs	0.035	0.057	0.913	-8,177.7	0.035	0.059	0.912	-0.003	-8,177.7	0.0

Note: It shows the parameter and fit estimates of GARCH(1,1) and GJR-GARCH(1,1) models of commodity futures volatility, and for comparison, other financial data series. The first three rows show results from fitting daily log returns of the S&P 500 index, the FTSE 100 index, and the US-GBP exchange rate. Energy, Grain, Metal, Soft, and “All” correspond to equal-weighted indices of the constituent commodity futures as shown in the grouping in Table 1. LL is the log likelihood. The LRT column shows the likelihood ratio test statistic ( $K - K' = 1$ , critical value (at  $\alpha = 0.05$ ) is 3.841). The series cover the period from September 1988 to December 2011, yielding 5,636 observations of daily log returns.

### 3.3 | Market activity

I obtain information on the evolution of different types of traders (classified as commercial [hedger], noncommercial, spread, or nonreporting [small] traders) and their activity in commodity markets from the Commitment of Traders (COT) reports made available by the US Commodity Futures Trading Commission (CFTC). Figure 2 shows the variation in the type of traders holding outstanding long and short positions in commodities, from January 1986 or December 2011. While the fraction of commercial traders’ (hedgers’) positions has not changed markedly, the fraction of outstanding spread positions (which trade the basis) has increased substantially. Moreover, the imbalance in commercial positions generally appears to be the opposite of the imbalance in noncommercial positions.

The set of variables identified from previous work that examines the impact of speculator activity on commodity futures returns (Acharya et al., 2013; Hong & Yogo, 2012) used as explanatory variables in  $\mathbf{z}_{i,t}$  include changes to open interest and demand imbalance of hedgers, that is, using commercial (“hedger”) position values collated by the CFTC,<sup>8</sup>

$$HEDGER\_IMB_{i,t} = \frac{ShortOI_{i,t} - LongOI_{i,t}}{ShortOI_{i,t} + LongOI_{i,t}}.$$

I use an indicator for the period beginning January 2004, commonly cited in previous work as the period showing index “financialization” (see, e.g., Singleton, 2014; Tang & Xiong, 2012), as  $IndexPeriod_t$ , when testing for changes in the dynamics of volatility due to commodity index trading.

Finally, the state of the hedge funds industry is captured using the absolute value of the mean of monthly hedge funds returns ( $HF\_RET_t$ ) using hedge fund data collated from the Lipper-TASS, BarclayHedge, Morningstar, HFR, and CISDM databases.

### 3.4 | Macroeconomic uncertainty indicators

I use the IMF World Economic Outlook Database for aggregate economic variables, and the IMF Direction of Trade Statistics for country-to-country aggregate import-export data. Both of these sources provide data at annual frequency. All interest rates and exchange rates are from the Global Financial Database and Datastream. Wherever necessary, World Bank classifications are used to group world economies.<sup>9</sup> The US GDP and CPI (quarterly) forecast statistics are from the Philadelphia Federal Reserve Bank’s Survey of Professional Forecasters. Economic forecasts for other

<sup>8</sup>Hong and Yogo (2012) investigate the power of futures open interest to predict commodity, currency, stock, and bond prices, and find open interest growth is more informative than other common alternatives as it is reflective of future economic activity.

<sup>9</sup>WB Country and Lending Groups Page (<http://data.worldbank.org/about/country-classifications/country-and-lending-groups>).

countries are from analyst forecasts collated in Bloomberg. US recession period data are from the National Bureau of Economic Research (NBER).

The choice of variables used in constructing the macroeconomic uncertainty series is motivated by previous studies (Bali et al., 2014; Bloom, 2014; Campbell & Shiller, 1988; Campbell et al., 2018).  $INF\_U$ —US inflation from change in consumer price index.  $INFFC\_A$ —Survey of Professional Forecasters, dispersion in next quarter CPI forecasts.  $TERM\_U$ —Spread between 10-year and 3-month Treasury yields.  $RREL\_U$ —Difference between 3-month Treasury yield and its 12-month geometric mean.  $DEF\_U$ —Baa-Aaa (Moody's) rated corporate bond yield spread.  $TED\_U$ —1M LIBOR—1M-T-Bill rates.  $UNEMP\_U$ —US unemployment rate.  $GDP\_U$ —US real GDP growth rate per capita.  $CFNAI\_U$ —Chicago Fed Economic Activity Index.  $RDIV\_U$ —Aggregate real dividend yield on S&P 500.  $MKT\_U$ —S&P 500 index excess return.  $VXO\_A$ —S&P 100 implied volatility index level.

These variables are available from January 1960 to the end of the sample period, except for CFNAI (from May 1967), TED (from January 1971), and VXO (from January 1986).  $X\_U_t$  denotes the one-period-ahead GARCH(1,1) volatility prediction of variable  $X$  made using all available observations up to time  $t - 1$  and  $X\_A_t$  denotes the AR(1) forecast made using all available observations up to time  $t - 1$ .

## 4 | EMPIRICAL RESULTS

### 4.1 | Macroeconomic uncertainty

Tables 4–7 show the results from balanced panel regressions of the (time,  $t$ ) 1-month realized volatility of the front-month futures return over lagged (time,  $t - 1$ ) explanatory variables, as specified in Equation (16). The volatility series are at a nonoverlapping monthly frequency. The results shown are for the commodity groups: energy, metal, grain, soft, and all (of the 17 commodities in the sample, see Section 3.1). The panel regressions all include commodity and seasonal (month-of-year) fixed effects, and  $t$ -statistics clustered by month are shown in parenthesis below each coefficient estimate. All regressions for a particular dependent variable include observations on the same dates, allowing for the comparison of information criterion.<sup>10</sup>

In Table 4, Panel A shows the baseline regressions with only lagged (time,  $t - 1$ ) volatility and lagged (absolute) return as explanatory variables (*Vol* and *Return*). This is similar in concept to a GARCH(1,1) formulation, broadly capturing the same information set at time  $t - 1$ . The coefficients are positive and highly significant. This is similar to empirical observations of equity, bond and other financial markets. In Panel B, I include the variable *PositiveReturn*, which is the return series with negative values replaced with zero. This formulation is similar to a GJR-GARCH(1,1) specification (see Equation (A13)) and allows for the capture of any asymmetric effect on volatility from the direction of the lagged return. Similar to the model fits in Table 3, these results also show that, unlike in the case of equities (Bekaert & Wu, 2000), there is no unconditional directional bias in the relationship between lagged return and volatility for commodity futures. Given the information contained in this asymmetric effect on the concentration and direction of risk and investor demand (Bekaert & Wu, 2000; Bollerslev & Todorov, 2011; Garleanu et al., 2009), the conditional variation in this relationship bears further study in the commodities space.

In Table 5, I add the variables capturing macroeconomic uncertainty. This results in an adjusted  $R^2$  gain of over 10% (for the energy group) from the baseline specification in Table 4, Panel A. Other comparisons of model fit such as BIC and LRT also show a clear improvement for the commodities in the energy, metal, grain, and all groups. The softs group has the smallest gain in proportion of explained variation. The inclusion of economic controls consistently improves the adjusted  $R^2$  and information criterion measures of model fit.  $INFFC\_A$ ,  $CFNAI\_U$  and  $RDIV\_U$  have positive and significant coefficients in the regression including all commodities.<sup>11</sup> These results are in agreement with the implications of the derivation in Section 2.1, which shows that variation in commodity futures volatility arises due to changes to the expectations of future interest rates, convenience yield, and risk premia. The inflation variables capture information about future interest rates and is informative of future economic and inflation regimes (David & Veronesi, 2013). Economic activity is related to the convenience yield (Casassus & Collin-Dufresne, 2005). The results including

<sup>10</sup>In the discussion of regression results that follow, model fit is considered using likelihood ratio tests, denoted LRT (Engle, 1984; Neyman & Pearson, 1933; Wilks, 1938), and Schwarz Bayesian information criteria (also known as the Bayesian information criterion), denoted BIC (G. E. Schwarz, 1978). LRT can be used to compare two nested models. More generally, a comparison using BIC is possible when the LHS-dependent variable is exactly the same, even when two models do not nest. The standard errors shown in the panel regression results are clustered by month (Petersen, 2009).

<sup>11</sup>See Table A6 of the Supporting Information Appendix for the results of Granger causality tests, which show that the direction of predictive causality is from the economic uncertainty variables included here to commodity futures volatility, rather than vice versa.

**TABLE 4** Commodity futures volatility—Panel regression results

	<b>Energy</b>	<b>Metal</b>	<b>Grain</b>	<b>Soft</b>	<b>All</b>
<i>Panel A: Baseline panel regression</i>					
(Intercept)	20.671 (4.571)	6.765 (5.084)	10.213 (7.897)	16.249 (8.419)	11.335 (10.582)
Return <sub>(i,t-1)</sub>	0.617 (4.074)	0.416 (3.321)	0.227 (3.137)	0.371 (4.119)	0.391 (6.781)
Vol <sub>(i,t-1)</sub>	0.412 (4.872)	0.478 (8.795)	0.520 (12.203)	0.336 (5.458)	0.428 (10.647)
Adjusted R <sup>2</sup>	0.343	0.437	0.419	0.252	0.395
Number of commodity-months	524	1,310	1,310	1,310	4,454
<i>Panel B: Baseline panel regression allowing for asymmetric return effect</i>					
(Intercept)	20.739 (4.606)	6.768 (5.087)	10.227 (7.892)	16.260 (8.567)	11.349 (10.546)
Return <sub>(i,t-1)</sub>	0.721 (4.125)	0.399 (3.040)	0.208 (2.409)	0.297 (2.329)	0.362 (4.490)
I <sub>[Return<sub>(i,t-1)&gt;0</sub> ×  Return<sub>(i,t-1)</sub> ]</sub>	-0.141 (-0.733)	0.025 (0.145)	0.031 (0.313)	0.118 (1.094)	0.044 (0.525)
Vol <sub>(i,t-1)</sub>	0.400 (4.851)	0.479 (8.628)	0.520 (12.159)	0.338 (5.629)	0.429 (10.833)
Adjusted R <sup>2</sup>	0.343	0.436	0.419	0.253	0.395
Number of commodity-months	524	1,310	1,310	1,310	4,454
LRT statistic	1.60	0.20	0.20	3.20	1.40
c [(K - K' = 1), ( $\alpha = 0.05$ )]	= 3.841				

Note: It shows the results for balanced panel regressions of (time,  $t$ ) 1-month volatility of the front-month futures return,  $Vol_{(t)}$ , over lagged (time,  $t - 1$ ) volatility and (absolute) return.  $PositiveReturn$  is the absolute return series of the front-month future with the negative return months set to 0. The results reported here are for the groups energy, metal, grain, soft, and all commodities. All regressions include commodity and season (month) fixed effects. Return variables are in percentage.  $t$ -statistics clustered by month are shown in parenthesis below each coefficient estimate.

controls for recession periods (Section 4.4) also capture the variation in risk premia associated with the business cycle. Moreover, these findings broadly confirm observations on the effects of uncertainty in other markets (Bloom, 2014).

It is difficult to contemporaneously explain, let alone predict, financial asset volatility using economic factors (see, e.g., Engle & Rangel, 2008; Engle et al., 2013; Roll, 1984; W.G. Schwert, 1989), even when model results and economic intuition posit a relationship between economic conditions and volatility. Consequently, the results in Table 5 constitute a step forward in our understanding of the factors that drive volatility.

Moreover, such predictive power is economically significant for a mean-variance investor (see, e.g., Campbell & Thompson, 2008; Inoue & Kilian, 2004; Moskowitz, Ooi, & Pedersen, 2012 for further discussion on the value of time series predictability). An adjusted  $R^2$  gain over the baseline model is useful for investors who have a nonzero “vega” exposure in their portfolio ( $\delta V / \delta \sigma \neq 0$  in a portfolio with value  $V$  and volatility  $\sigma$ ) as, in that case, predicting volatility allows for the prediction of portfolio and position values. This is important for any hedger or derivatives trader as the value of their portfolios and trading strategies is directly tied to the volatility of the traded assets.

For tractability, in the regressions that follow, I use the first four principal components to capture the variation of the 11 economic uncertainty series in the main regressions. Table A3 of the Supporting Information Appendix presents the details of the principal component analysis. The panels in Table A4 show the regressions results with varying numbers of principal components included.<sup>12</sup> In future work, I include uncertainty proxies directly based on work by Jurado et al. (2014).

## 4.2 | Hedging and trading activity

Table 6 shows the results once the variables capturing momentum and hedging activity (Acharya et al., 2013; Hong & Yogo, 2012) are added to the specification in Table 5, which includes the macroeconomic controls. While there is some improvement, there is no consistent gain in predictive power. Table A5 of the Supporting Information Appendix shows the results without the inclusion of the macroeconomic controls. The regressions adding only economic uncertainty variables to the baseline specification as in Table 5 perform better on the dimensions of adjusted  $R^2$  and information criterion measures of model fit.

<sup>12</sup>See Bai and Ng (2002) and Stock and Watson (2002) on selecting the appropriate number of factors.

**TABLE 5** Commodity futures volatility and macroeconomic uncertainty

	<b>Energy</b>	<b>Metal</b>	<b>Grain</b>	<b>Softs</b>	<b>All</b>
<i>EconU_1<sub>(t-1)</sub></i>	-2.166*** (-4.975)	-1.030*** (-4.723)	-0.890*** (-4.991)	-0.630*** (-4.552)	-0.952*** (-6.934)
<i>EconU_2<sub>(t-1)</sub></i>	-1.356** (-2.139)	-0.417 (-1.473)	0.037 (0.123)	0.380 (1.468)	-0.155 (-0.759)
<i>EconU_3<sub>(t-1)</sub></i>	2.261*** (4.616)	0.538* (1.903)	0.291 (1.114)	0.426* (1.749)	0.586*** (3.968)
<i>EconU_4<sub>(t-1)</sub></i>	1.188** (2.005)	1.690*** (5.011)	1.513*** (3.981)	-0.320 (-0.997)	0.977*** (4.489)
All predictors in Table 4	Yes	Yes	Yes	Yes	Yes
Adjusted <i>R</i> <sup>2</sup>	0.442	0.474	0.459	0.263	0.423
Number of commodity-months	524	1,310	1,310	1,310	4,454
LRT statistic	90.0	93.2	97.4	24.2	217.2

$$c[(K - K' = 4), (\alpha = 0.05)] = 9.488$$

Note: It shows the results for balanced panel regressions of (time, *t*) 1-month volatility of the front-month futures return, *Vol<sub>(t)</sub>*, over lagged (time, *t* - 1) explanatory variables that capture macroeconomic uncertainty. These variables are the first four principal components of 11 lagged macroeconomic uncertainty series (Tables A3 and A4 in the Supporting Information Appendix contain details of this principal component analysis). The results reported here are for the groups energy, metal, grain, soft, and all commodities. All regressions include commodity and season (month) fixed effects. Return variables are in percentage. *t*-Statistics clustered by month are shown in parenthesis below each coefficient estimate. The LRT row shows the likelihood ratio test statistic comparing the fit shown in Table 4A to the current results.

The significance of the coefficient estimate is indicated by \**p* < .10, \*\**p* < .05, and \*\*\**p* < .01.

### 4.3 | Hedge fund performance

If financialization increased access to commodity futures markets by participants such as hedge funds, the comovement between futures returns and large-scale trading activity and portfolio shocks of hedge funds may have increased to such an extent that cannot be absorbed by other market participants due to borrowing constraints, illiquidity, or other market friction that introduces limits to arbitrage. I test the hypothesis that shocks to hedge funds during the financialization period are associated with higher commodity futures volatility using the regression specification:

$$\begin{aligned} Vol_{i,t} = \mu_i + \beta_1 HF\_RET_{t-1} + \mathbf{z}'_{t-1} \theta + \beta_2 I_{[t-1 \in IndexPeriod]} \\ + I_{[t-1 \in IndexPeriod]} \times \mathbf{z}'_{t-1} \theta^{INDEX} + \beta_3 I_{[t-1 \in IndexPeriod]} HF\_RET_{t-1} + \eta_{i,t}, \end{aligned} \quad (17)$$

**TABLE 6** Commodity futures volatility and commodity market risk factors

	<b>Energy</b>	<b>Metal</b>	<b>Grain</b>	<b>Softs</b>	<b>All</b>
<i>CMOM<sub>(i,t-1)</sub></i>	0.804 (0.681)	4.009*** (3.538)	0.503 (0.561)	-0.615 (-0.665)	1.767*** (2.882)
<i>HEDGER_OIG<sub>(i,t-1)</sub></i>	1.235 (1.245)	-0.177 (-0.388)	0.282 (0.863)	0.774 (1.201)	0.320 (1.149)
<i>HEDGER_IMB<sub>(i,t-1)</sub></i>	-0.153 (-0.806)	-0.021 (-0.472)	-0.069 (-0.786)	-0.114 (-1.338)	-0.089** (-2.328)
All predictors in Table 5	Yes	Yes	Yes	Yes	Yes
Adjusted <i>R</i> <sup>2</sup>	0.449	0.483	0.46	0.264	0.428
Number of commodity-months	524	1,310	1,310	1,310	4,454
LRT statistic	9.6	27.2	5.0	4.4	41.2

$$c[(K - K' = 3), (\alpha = 0.05)] = 7.815$$

Note: It shows the results for balanced panel regressions of (time, *t*) 1-month volatility of the front-month futures return, *Vol<sub>(t)</sub>*, over lagged (time, *t* - 1) commodity market variables in addition to the macroeconomic uncertainty factors included in Table 5. The results reported here are for the groups energy, metal, grain, soft, and all commodities. All regressions include commodity and season (month) fixed effects. Return variables are in percentage. *t*-Statistics clustered by month are shown in parenthesis below each coefficient estimate. The LRT row shows the likelihood ratio test statistic comparing the fit shown in Table 5 to the current results.

The significance of the coefficient estimate is indicated by \**p* < .10, \*\**p* < .05, and \*\*\**p* < .01.

**TABLE 7** Commodity futures volatility and hedge fund activity

	<b>Energy</b>	<b>Metal</b>	<b>Grain</b>	<b>Softs</b>	<b>All</b>	<b>1</b>	<b>2</b>	<b>3</b>
<i>HF_RET<sub>(t-1)</sub></i>	0.684 (1.435)	0.220 (0.611)	0.864*** (2.585)	0.106 (0.428)	0.421* (1.798)	-0.223 (-1.125)	0.120 (0.702)	
<i>IndexPeriod<sub>(t)</sub> × HF_RET<sub>(t-1)</sub></i>						1.085*** (2.627)		
<i>Recession<sub>(t)</sub> × HF_RET<sub>(t-1)</sub></i>							1.615** (1.983)	
All predictors in Table 6	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted <i>R</i> <sup>2</sup>	0.451	0.473	0.440	0.250	0.416	0.422	0.427	
Number of commodity-months	376	940	940	940	3,196	3,196	3,196	

Note: It shows the results for balanced panel regressions of (time, *t*) 1-month volatility of the front-month futures return, *Vol*<sub>(*t*)</sub>, over lagged (time, *t* - 1) absolute value of the hedge fund industry mean return, in addition to all explanatory variables included in Table 6. The results reported here are for the groups energy, metal, grain, soft, and all commodities, from January 1995 to December 2010. All regressions include commodity and season (month) fixed effects. Return variables are in percentage. *t*-Statistics clustered by month are shown in parenthesis below each coefficient estimate.

The significance of the coefficient estimate is indicated by \**p* < .10, \*\**p* < .05, and \*\*\**p* < .01.

where *HF\_RET<sub>t</sub>* denotes a proxy capturing shocks to hedge funds.  $\beta_3 > 0$  in the specification in (17).

Table 7 controls for hedge fund activity by including lagged hedge fund (absolute) return as an explanatory variable. *HF\_RET* has a positive and significant coefficient of 0.421 with a *t*-statistic of 1.798, even after the inclusion of all proxies for economic uncertainty and hedging activity included in Table 6. The coefficient for grain commodities is the most significant, and this potentially links to the consequences of market changes related to the US ethanol mandate (Roberts & Schlenker, 2013). After interacting for the indicator for the “index period,” I find that this positive relationship is limited to this period, when the coefficient is 1.085 with a *t*-statistic of 2.627. However, given that this period overlaps significantly with a major recession, this identification in a smaller sample (starting in January 1995) is weak.

#### 4.4 | Introduction of commodity indices

Table 8 shows the comparison of model fits, in addition to results with interactions for different time periods added to the specification in Table 6. In column 5, I interact for the NBER recession periods as in regression specification (18), and in column 6, I show the results from interacting with *IndexPeriod* as in regression specification (19):

$$Vol_t = \mu_i + NBER\_Recession + \mathbf{z}'_{t-1}\theta + NBER\_Recession \times \mathbf{z}'_{t-1}\theta^{REC} + \eta_t, \quad (18)$$

$$Vol_t = \mu_i + IndexPeriod + \mathbf{z}'_{t-1}\theta + IndexPeriod \times \mathbf{z}'_{t-1}\theta^{INDEX} + \eta_t. \quad (19)$$

Interacting for *NBER\_Recession* increases the model fit for all groups relative to the specification without the interaction with up to a 13.6% adjusted *R*<sup>2</sup> gain for energy commodities. Commodities in the grain and softs groups show a better fit under *IndexPeriod* interactions. Metal commodities show no significant difference between the two specifications, while energy commodities have less explanatory power under the interaction with *IndexPeriod*.

#### 5 | CONCLUSIONS

This paper conducts a systematic analysis to understand the dynamics of commodity futures volatility. I derive the variance decomposition for commodity futures to show how unexpected changes to the excess basis return are driven by changes to the expectation of future interest rates, convenience yield, and risk premia. These expectations are updated in response to news about the future state of the economy and future commodity supply and demand. I model time-varying commodity futures volatility and study the impact of variables that proxy for such economic uncertainty,

**TABLE 8** Commodity futures volatility during different time periods

	<b>Baseline 1</b>	<b>Macro uncertainty 2</b>	<b>EU_PC 1-4 3</b>	<b>Commodity market 4</b>	<b>Interaction for recession periods 5</b>	<b>Interaction for IndexPeriod 6</b>
<i>Panel A: Adjusted R<sup>2</sup></i>						
Energy	34.3	48.9	44.2	44.9	58.5	51.1
Metal	43.7	47.3	47.4	48.3	49.2	49.3
Grain	41.9	47.8	45.9	46.0	48.6	49.1
Softs	25.2	26.3	26.3	26.4	26.2	26.7
All	39.5	42.4	42.3	42.8	43.9	43.1
<i>Panel B: Bayesian information criterion (BIC)</i>						
Energy	4,136.6	4,057.1	4,071.6	4,080.8	3,983.6	4,069.7
Metal	9,760.8	9,733.7	9,696.2	9,690.7	9,728.9	9,727.6
Grain	9,306.6	9,256.4	9,237.8	9,254.3	9,252.7	9,239.6
Softs	10,210.0	10,251.1	10,214.4	10,231.6	10,295.9	10,287.9
All	33,500.9	33,351.5	33,317.3	33,301.3	33,285.9	33,353.4
<i>Panel C: Likelihood ratio comparison (LRT)</i>						
$K - K'$		$K_2 - K_1 = 10$	$K_3 - K_1 = 4$	$K_4 - K_2 = 3$	$K_5 - K_4 = 10$	$K_6 - K_4 = 10$
Critical value						
( $\alpha = 0.05$ )	18.307	9.488	7.815	18.307	18.307	18.307
Energy	142.2	90.0	9.6	159.8	73.6	
Metal	98.8	93.2	27.2	33.6	34.8	
Grain	122.0	97.4	5.0	73.6	86.6	
Softs	30.6	24.2	4.4	7.4	15.4	
All	233.4	217.2	41.2	99.4	31.8	

Note: It shows model fit measures for the balanced panel regressions of (time,  $t$ ) 1-month volatility of the front-month futures return over lagged (time,  $t-1$ ) explanatory variables under different specifications. Panel A shows the adjusted  $R^2$ , Panel B shows the Bayesian information criterion values, and Panel C shows the likelihood ratio statistic of the different models. Column 1 shows the fit measures from the regression specification in Table 4A (“Baseline”), column 2 shows those from including all 11 macroeconomic uncertainty series on the RHS, column 3 shows those from Table 5 (“EconU\_1 – 4”), and column 4 shows those from Table 6 (“Commodity market”). Column 5 shows the results with all explanatory variables in Table 6 included, together with interactions for NBER recession periods; the last column shows the results with all explanatory variables in Table 6 included, together with interactions with *IndexPeriod* (the indicator for the period after January 2004). The results reported here are for the groups energy, metal, grain, soft, and all commodities. All regressions include commodity and season (month) fixed effects.

while controlling for the impact of any frictions due to trading activity. Using data for major commodity futures markets, economic indicators, and commodity market trading activity, I analyze the extent to which commodity volatility is related to fundamentals that impact convenience yield and interest rates such as inflation uncertainty, as well as financial frictions introduced by changing market structure and commodity index trading. I find significant predictability in commodity futures volatility using variables capturing macroeconomic uncertainty.

Such explanatory power can be economically significant for market participants (Campbell & Thompson, 2008; Inoue & Kilian, 2004). Investors who have volatility-sensitivity (a nonzero “vega” exposure) in their portfolios would especially benefit as in that case, predicting volatility allows for the prediction of portfolio and position values. This is important for any hedger or derivatives trader, as the value of their portfolios and trading strategies is directly tied to the volatility of the traded assets. Investors and end-users (commodity producers and consumers) in commodity markets benefit from understanding how the observed price behavior relates to the prevailing economic conditions. Uncertainty can lead to the long-term misallocation of resources as end-users evaluate real options in their investment decisions. Moreover, for many commodities with illiquid or short-dated derivatives markets of little depth, these findings can be a useful aid to price discovery and risk management.

It is difficult to contemporaneously explain, let alone predict, financial asset volatility using factors reflecting economic conditions (Engle & Rangel, 2008; Roll, 1984; W.G. Schwert, 1989), even when model results and economic intuition posit such a relationship. Consequently, the results in this paper constitute a step forward in our understanding of the factors that drive volatility.

## ACKNOWLEDGMENTS

I thank Rui Albuquerque, Turan Bali, Greg Duffee, Michael Gallmeyer, Lena Koerber, Mathias Krutli, José Martinez, Thomas Noe, Diaa Noureldin, Tarun Ramadorai, Neil Shephard, Kevin Sheppard, Galen Sher, Peter Tufano, Raman Uppal, Mungo Wilson, Brian Wright, and various seminar participants for helpful comments, and the Oxford-Man Institute for financial support. Any errors remain the responsibility of the author.

## REFERENCES

- Acharya, V. V., Lochstoer, L., & Ramadorai, T. (2013). Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics*, 109, 441–465.
- Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 51, 259–299.
- Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91, 1–23.
- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70, 191–221.
- Bali, T. G., Brown, S. J., & Caglayan, M. O. (2014). Macroeconomic risk and hedge fund returns. *Journal of Financial Economics*, 114(1), 1–19.
- Basak, S., Cass, D., Licari, J. M., & Pavlova, A. (2008). Multiplicity in general financial equilibrium with portfolio constraints. *Journal of Economic Theory*, 142, 100–127.
- Basak, S., & Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review*, 103, 1728–1758.
- Basak, S., & Pavlova, A. (2015). A model of financialization of commodities. *Journal of Finance*, 71(4), 1511–1556.
- Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13, 1–42.
- Bernanke, B. (1983). Irreversibility, uncertainty and cyclical investment. *Quarterly Journal of Economics*, 98, 85–106.
- Bhamra, H. S., & Uppal, R. (2006). The effect of introducing a non-redundant derivative on the volatility of stock-market returns. *Review of Financial Studies*, 22, 2303–2330.
- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.
- Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77, 623–685.
- Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28, 153–176.
- Bloom, N., Bond, S., & VanReenen, J. (2007). Uncertainty and investment dynamics. *Review of Economic Studies*, 74, 391–415.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T., & Todorov, V. (2011). Tails, fears and risk premia. *The Journal of Finance*, 66, 2165–2211.
- Brennan, D., Williams, J., & Wright, B. D. (1997). Convenience yield without the convenience: A spatial-temporal interpretation of storage under backwardation. *The Economic Journal*, 107, 1009–1022.
- Brennan, M. (1958). The supply of storage. *American Economic Review*, 48, 50–72.
- Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal*, 101, 157–179.
- Campbell, J. Y., & Ammer, J. (1993). What moves the stock and bond markets? A variance decomposition for long-term asset returns. *The Journal of Finance*, 48, 3–37.
- Campbell, J. Y., Giglio, S., Polk, C., & Turley, R. (2018). An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics*, 128(2), 207–233.
- Campbell, J. Y., & Shiller, R. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1, 195–228.
- Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21, 1509–1531.
- Casassus, J., & Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance*, 60, 2283–2331.
- Cass, D. (1992). Sunspots and incomplete financial markets: The general case. *Economic Theory*, 2, 341–358.
- Chen, Y., Rogoff, K. S., & Rossi, B. (2010). Can exchange rates forecast commodity prices? *Quarterly Journal of Economics*, 125, 1145–1194.
- Cootner, P. H. (1960). Returns to speculators: Telser versus keynes. *Journal of Political Economy*, 68, 396–404.
- David, A., & Veronesi, P. (2013). What ties return volatilities to price valuations and fundamentals? *Journal of Political Economy*, 121, 682–746.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50, 987–1008.
- Engle, R. F. (1984). Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. In Z. Griliches, & M. D. Intriligator (Eds.), *Handbook of Econometrics* (1st ed., Vol. 2, pp. 775–826). Elsevier B. V. <https://www.sciencedirect.com/science/article/pii/S1573441284020055>
- Engle, R. F., & Gallo, G. M. (2008). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131, 3–27.
- Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95, 776–797.
- Engle, R. F., & Lee, G. (1999). A long run and short run component model of stock return volatility. In R. F. Engle, & H. White (Eds.), *Cointegration, causality, and forecasting: A festschrift in honour of Clive W. J* (pp. 475–497). Granger: Oxford University Press.
- Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 23, 1187–1222.
- Etula, E. (2013). Broker-dealer risk appetite and commodity returns. *Journal of Financial Econometrics*, 11, 486–521.
- Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79, 733–772.
- Garleanu, N., Pedersen, L. H., & Potoshman, A. (2009). Demand-based option pricing. *Review of Financial Studies*, 22, 4259–4299.
- Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45, 959–976.
- Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relationship between the expected value and the volatility of the nominal excess returns on stocks. *The Journal of Finance*, 48, 1779–1801.
- Gromb, D., & Vayanos, D. (2002). Equilibrium and welfare in markets with constrained arbitrageurs. *Journal of Financial Economics*, 66, 361–407.
- Gromb, D., & Vayanos, D. (2010). Limits of arbitrage: The state of the theory. *Annual Review of Financial Economics*, 2, 251–275.

- Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70, 393–408.
- Haase, M., & Zimmermann, H. (2013). Scarcity, risk premiums, and the pricing of commodity futures: the case of crude oil contracts. *The Journal of Alternative Investments*, 16(1), 43–71.
- Hamilton, J. D., & Wu, J. C. (2015). Effects of index-fund investing on commodity futures prices. *International Economic Review*, 56, 187–205.
- Henderson, B. J., Pearson, N. D., & Wang, L. (2015). New evidence on the financialization of commodity markets. *Review of Financial Studies*, 28, 1285–1311.
- Hirshleifer, D. (1988). Residual risk, trading costs, and commodity futures risk premia. *Review of Financial Studies*, 1, 173–193.
- Hirshleifer, D. (1990). Hedging pressure and futures price movements in a general equilibrium model. *Econometrica*, 58, 411–428.
- Hong, H. G., & Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics*, 105, 473–490.
- Inoue, A., & Kilian, L. (2004). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23, 371–402.
- Irwin, S. H., & Sanders, D. R. (2012). Testing the masters hypothesis in commodity futures markets. *Energy economics*, 34, 256–269.
- Jovanovic, B. (2013). Bubbles in prices of exhaustible resources. *International Economic Review*, 54, 1–34.
- Jurado, K., Ludvigson, S., & Ng, S. (2014). Measuring uncertainty. *American Economic Review*, 105(3), 1177–1216.
- Kaldor, N. (1939). Speculation and economic stability. *Review of Economic Studies*, 7, 1–27.
- Kelly, B. T., Lustig, H. N., & Van Nieuwerburgh, S. (2013). *Firm volatility in granular networks* (Working Paper). University of Chicago.
- Keynes, J. M. (1930). *A treatise on money*. London: Macmillan.
- Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99, 1053–1069.
- Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141–183.
- Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, 104, 228–250.
- Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London*, 231, 289–337.
- Pavlova, A., & Rigobon, R. (2008). The role of portfolio constraints in the international propagation of shocks. *Review of Economic Studies*, 75, 1215–1256.
- Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22, 435–480.
- Roberts, M. J., & Schlenker, W. (2013). Identifying supply and demand elasticities of agriculture commodities: Implications for the US ethanol mandate. *American Economic Review*, 103, 2265–2295.
- Roll, R. (1984). Orange juice and weather. *American Economic Review*, 74, 861–880.
- deRoon, F. A., Nijman, T. E., & Veld, C. (2000). Hedging pressure effects in futures markets. *The Journal of Finance*, 55, 1437–1456.
- Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52, 923–973.
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.
- Schwert, W. G. (1989). Why does stock market volatility change over time? *The Journal of Finance*, 44, 1115–1153.
- Sen, A. (1983). *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford, UK: Oxford University Press.
- Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52, 35–55.
- Singleton, K. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60, 300–318.
- Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97, 1167–1179.
- Tang, K., & Xiong, W. (2012). Index investment and financialization of commodities. *Financial Analysts Journal*, 68, 54–74.
- US Senate Permanent Subcommittee on Investigations. (2009). *Excessive speculation in the wheat market* (Report). Washington, D.C.: Committee on Homeland Security and Government Affairs. <https://www.hsgac.senate.gov/imo/media/doc/REPORTExcessiveSpeculationintheWheatMarketwoexhibitschartsJune2409.pdf>
- US Senate Permanent Subcommittee on Investigations. (2014). *Wall street bank involvement with physical commodities* (Report). Washington, D.C.: Committee on Homeland Security and Governmental Affairs. [https://www.hsgac.senate.gov/imo/media/doc/REPORT-Wall%20Street%20Bank%20Involvement%20With%20Physical%20Commodities%20\(12-5-14\).pdf](https://www.hsgac.senate.gov/imo/media/doc/REPORT-Wall%20Street%20Bank%20Involvement%20With%20Physical%20Commodities%20(12-5-14).pdf)
- Watugala, S. W. (2017). *Global trade and the term structure dynamics of commodity returns* (Working Paper). Cornell University.
- Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9, 60–62.
- Working, H. (1933). Price relations between July and September wheat futures at Chicago since 1885. *Wheat Studies of the Food Research Institute*, 9, 187–238.
- Working, H. (1949). The theory of the price of storage. *American Economic Review*, 39, 1254–1262.
- Yuan, K. (2005). Asymmetric price movements and borrowing constraints: A rational expectations equilibrium model of crises, contagion, and confusion. *The Journal of Finance*, 60, 379–411.

**How to cite this article:** Watugala SW. Economic uncertainty, trading activity, and commodity futures volatility. *J Futures Markets*. 2019;39:921–945. <https://doi.org/10.1002/fut.22018>

## APPENDIX A

## A.1 | Components of the excess basis return and term spread

We can further decompose the excess basis return,  $x_{n,t+1}$ , in Equation (10) to separate out the excess return due to the interest rate term structure and characterize the excess return purely due to convenience yield and commodity risk premia:

$$x_{n,t+1} \equiv x_{n,t+1}^y - x_{n,t+1}^r, \quad (\text{A1})$$

$$x_{n,t+1}^y - E_t x_{n,t+1}^y = (E_{t+1} - E_t) \left\{ - \sum_{i=1}^{n-1} y_{1,t+i} - \sum_{i=1}^{n-1} x_{n-i,t+i+1}^y \right\}, \quad (\text{A2})$$

$$x_{n,t+1}^r - E_t x_{n,t+1}^r = (E_{t+1} - E_t) \left\{ - \sum_{i=1}^{n-1} \pi_{1,t+i} - \sum_{i=1}^{n-1} \psi_{1,t+i} - \sum_{i=1}^{n-1} x_{n-i,t+i+1}^r \right\}, \quad (\text{A3})$$

where  $\pi_{1,t}$  is the one-period inflation rate and  $\psi_{1,t}$  is the one-period real interest rate at time  $t$ . The derivation of (A3) is discussed in Campbell and Ammer (1993).

In this appendix, I also show the extension of the decomposition presented in Section 2.1 for the commodity futures term spread.

Define the term spread as  $s_{n,t} \equiv y_{n,t}^* - y_{1,t}^*$ , where  $y_{n,t}^* = y_{n,t} - r_{n,t}$ . Given  $p_{n,t} = n(r_{n,t} - y_{n,t})$  and  $p_{n,t} = -E_t \sum_{i=0}^{n-1} \delta_{n-i,t+i+1}$ ,

$$\begin{aligned} y_{n,t}^* &= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} \delta_{n-i,t+i+1} \\ &= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + y_{1,t+i}^*). \end{aligned} \quad (\text{A4})$$

Given Equation (A4) and the definition of futures term spread  $s_{n,t}$ , the unexpected excess return can be decomposed into the unanticipated change in  $y_{1,t}^*$  and the unexpected change in futures term spread:

$$\begin{aligned} x_{n,t+1} - E_t x_{n,t+1} &= -(n-1)(y_{n-1,t+1}^* - E_t y_{n-1,t+1}^*) \\ &= -(n-1)(E_{t+1} - E_t)[y_{1,t+1}^* + s_{n-1,t+1}]. \end{aligned} \quad (\text{A5})$$

Given the definition of  $y_{1,t}^*$ , it is straightforward to show that  $y_{1,t+1}^* - E_t y_{1,t+1}^* = (E_{t+1} - E_t)[y_{1,t+1} - r_{1,t+1}]$ . To derive the unexpected change in futures term spread, I start with Equation (A4) to relate futures term spread to expectations of future excess return and changes to the basis yield of the front-month future. Then I have,

$$\begin{aligned} s_{n,t} &\equiv y_{n,t}^* - y_{1,t}^* \\ &= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + y_{1,t+i}^* - ny_{1,t}^*) \\ &= \left( \frac{1}{n} \right) \left( E_t \sum_{i=0}^{n-1} x_{n-i,t+i+1} \right) + E_t[-ny_{1,t}^* + y_{1,t}^* + (n-1)y_{1,t+1}^* - (n-2)y_{1,t+1}^* \\ &\quad + \dots + 3y_{1,t+n-3}^* - 2y_{1,t+n-3}^* + 2y_{1,t+n-2}^* - y_{1,t+n-2}^* + y_{1,t+n-1}^*] \\ &= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + (n-i-1)\Delta y_{1,t+i+1}^*). \end{aligned} \quad (\text{A6})$$

It follows that the unanticipated change in futures term spread can be derived as

$$\begin{aligned}s_{n-1,t+1} - E_t s_{n-1,t+1} &= \left( \frac{E_{t+1} - E_t}{n-1} \right) \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + (n-i-1)\Delta y_{1,t+i+1}^*) \\ &= \nu_{t+1}^x + \nu_{t+1}^y - \nu_{t+1}^r.\end{aligned}\quad (\text{A7})$$

When both the expectations hypothesis for the term structure of interest rates and the theory of storage hold exactly, the first term on the right-hand side is zero, and innovations in futures term spread are driven by changes to expectations of future changes to front-month convenience yield and short rates.

## A.2 | Volatility models and extensions

In this appendix, I describe the realized volatility models that form the basis of the empirical analysis.

### A.2.1 | GARCH-type models

Drawing on previous work on equity market volatility (Engle & Lee, 1999; Engle & Gallo, 2008; Engle & Rangel, 2008), I use a GARCH-type model of volatility to check the robustness of the baseline regression analysis. A standard GARCH (1,1) process (Bollerslev, 1986; Engle, 1982) for a particular asset is defined as

$$r_t = \mu_t + \sigma_t u_t, \quad (\text{A8})$$

$$\text{where } u_t | \Omega_{t-1} \sim N(0, 1) \quad \text{and} \quad \varepsilon_t = \sigma_t u_t, \quad (\text{A9})$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}, \quad (\text{A10})$$

$$\sigma_t = \sqrt{h_t}. \quad (\text{A11})$$

It follows that the unconditional variance in the model will be  $E[(r_t - E_{t-1} r_t)^2] = E[(r_t - \mu_t)^2] = \omega/(1 - \alpha - \beta)$ . In its simplest form, extensions to the standard GARCH(1,1) process that include  $K$  (weakly) exogenous lagged explanatory variables in  $\mathbf{z}_t$ , with  $\xi_t = \mathbf{z}_t / (E[\mathbf{z}_t])$ , take the form of GARCH-X(1,1):

$$\begin{aligned}g_t &= \omega + \alpha \varepsilon_{t-1}^2 + \beta g_{t-1} + \xi_t' \theta, \\ \sigma_t &= \sqrt{g_t}.\end{aligned}\quad (\text{A12})$$

Then the unconditional variance  $E[(r_t - E_{t-1} r_t)^2] = (\omega + \gamma_1 + \dots + \gamma_K)/(1 - \alpha - \beta)$ .

Note that, unlike with equity (Bekaert & Wu, 2000; Bollerslev & Todorov, 2011), there is no direct equivalent to the firm leverage effect for commodities and risk can be concentrated in either direction depending on the shock to supply or demand.

A model capturing asymmetry in a manner such as the GJR-GARCH model<sup>13</sup> may be useful for learning about the conditional demand- or supply-side pressures in a commodity market. As seen in Tables 3 and 4B, for commodity futures, there is no unconditional asymmetric volatility effect when controlling solely for the sign of lagged returns.

<sup>13</sup>From Glosten, Jagannathan, and Runkle (1993), a GJR-GARCH(1,1) process allows for an asymmetric return effect, and differs from the specification of GARCH(1,1) in (A10) by the specification of  $h_t$ :

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \gamma I_{t-1}^{(-)} \varepsilon_{t-1}^2 + \beta h_{t-1}, \quad (\text{A13})$$

where  $I_{t-1}^{(-)}$  is 1 when  $\varepsilon_{t-1} < 0$ , and 0 otherwise.

### A.2.2 | Long-run and short-run volatility components

Consider the short-term and long-term components of the data-generating process within a framework similar to the models of equity volatility presented in Engle and Rangel (2008) and Engle et al. (2013), differing only in terms of the definition of the slow-moving component of volatility:

$$\begin{aligned} r_t &= \mu_t + \sigma_t u_t, \\ \text{where } u_t | \Omega_{t-1} &\sim N(0, 1) \quad \text{and} \quad \varepsilon_t = \sigma_t u_t, \\ h_t &= (1 - \alpha - \beta) + \alpha \frac{\varepsilon_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1}, \\ \sigma_t &= \sqrt{\tau_t h_t}, \end{aligned} \tag{A14}$$

where  $\tau_t$  represents the long-term volatility component and, for a set of  $K$  lagged explanatory variables in  $\mathbf{z}_t$ , is defined as

$$\log \tau_t = m + \mathbf{z}'_t \boldsymbol{\theta}. \tag{A15}$$

The size of the set of estimated parameters in the model,  $\Theta = \{\mu, \alpha, \beta, m, \gamma_1, \dots, \gamma_K\}$ , is on the same order as the GARCH-X model presented in the previous section. In this model, the unconditional variance corresponds exactly to the low-frequency component as  $E[(r_t - E_{t-1} r_t)^2] = \tau_t E[h_t] = \tau_t$ .

Engle et al. (2013), in their analysis of the macroeconomic determinants of equity market volatility, separately consider the impact of the level and volatility of two variables: inflation and industrial production growth. They find a significant impact from these macroeconomic variables even on daily volatility. Their model differs in the definition of  $\tau$  in (A15) by including multiple lags of each explanatory variable with an imposed weighting function. This limits the number of factors that can be included together as each adds three parameters to  $\Theta$ .

In contrast, Engle and Rangel (2008), in their spline-GARCH specification (also differing solely in their definition of (A15), estimate  $\tau$  nonparametrically using an exponential quadratic spline:

$$\tau_t = c \exp \left( w_0 t + \sum_{i=1}^k w_i ((t - t_{i-1})_+)^2 \right), \tag{A16}$$

where  $(t - t_i)_+ = \{t - t_i \text{ if } t > t_i, \text{ otherwise } 0\}$  and  $k$  is the optimal number of equally-spaced knots, selected using information criteria (AIC and BIC). This partitions the time series into  $k$  equally-spaced intervals, demarcated by  $\{t_0 = 0, t_1, \dots, t_k = T\}$ . The estimated time series of the slow-moving component ( $\tau$ ) is subsequently used as the dependent variable in an independent regression, with up to 11 explanatory variables in their model: economic development level, market capitalization, inflation level, GDP level and growth, market size (number of listed companies), and volatilities of the short-term interest rate, exchange rate, GDP, and inflation.

Correspondingly, I include a number of variables in my analysis that are potentially relevant for commodity markets in  $\mathbf{z}$  that can capture the impact of macroeconomic uncertainty, supply-demand shocks, and trading activity.