Economic Uncertainty and Commodity Futures Volatility*

Sumudu W. Watugala†

December 7, 2014‡

Abstract

This paper investigates the dynamics of commodity futures volatility. I derive the variance decomposition for the commodity futures basis to show how unexpected excess returns result from new information about the expected future interest rates, convenience yields, and risk premia. This motivates my empirical analysis of the volatility impact of economic and inflation regimes and commodity supply-demand shocks. Using data on major commodity futures markets and global bilateral commodity trade, I analyze the extent to which commodity volatility is related to fundamental uncertainty from increased emerging market demand and macroeconomic forecast uncertainty, while controlling for the potential impact of financial frictions introduced by changing market structure and commodity index trading. Higher concentration in emerging market importers of a commodity is associated with higher futures volatility. I find commodity futures volatility is significantly predictable using variables capturing macroeconomic uncertainty. I examine the conditional variation in the asymmetric relationship between returns and volatility, and how this relates to the futures basis and sensitivity to consumer and producer shocks.

JEL classification: F36, G12, G13, G15, Q02.

*I thank Rui Albuquerque, Greg Duffee, Michael Gallmeyer, Lena Koerber, Mathias Kruttli, José Martinez, Thomas Noe, Diaa Noureldin, Tarun Ramadorai, Neil Shephard, Kevin Sheppard, Galen Sher, Emil Siriwardane, Peter Tufano, Raman Uppal, Mungo Wilson, and seminar participants at the Saïd Business School, Oxford-Man Institute, 2013 LBS-TADC Conference, Commodity Futures Trading Commission, and US Department of the Treasury for helpful comments, and the Oxford-Man Institute for financial support. Any errors remain the responsibility of the author. The views expressed in this paper are those of the author’s and do not necessarily represent the official positions or policy of the Office of Financial Research or US Department of the Treasury.

†Saïd Business School, University of Oxford, Oxford-Man Institute of Quantitative Finance, and Office of Financial Research, US Department of the Treasury. Email: sumudu.watugala@sbs.ox.ac.uk

‡The latest version of this paper can be found at: http://people.csail.mit.edu/sumudu/docs/Sumudu_Watugala_CommodityVolatility.pdf
1. Introduction

This paper investigates the time-variation in commodity futures volatility and the factors explaining its dynamics. I analyze the impact of increased emerging market demand on commodity markets and the impact of concentration in these markets. This research builds on Bloom (2014), which presents evidence that emerging markets and recessionary periods are strongly associated with economic uncertainty, and Gabaix (2011), which shows the impact on volatility from power laws in size distributions. This adds to the literature on what explains fluctuations in volatility (see, for example, Roll (1984); Schwert (1989); Engle and Rangel (2008); Gabaix (2011); Bloom (2014)), while contributing to the current debate on commodity price dynamics and potential distortions. I examine how supply-demand shocks, 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 true dynamics when evaluating real options embedded in investment choices (Schwartz 1997). Distortions can lead to underinvestment or overinvestment, and even transitory deviations from fundamentals can lead to long-term misallocation of resources (see, for example, Bernanke (1983); Bloom, Bond, and Reenen (2007)). This is especially important when there are non-convex production functions and large fixed costs to entry and expansion (e.g., a producer of copper considering the development of a new mine or a manufacturer considering the opening of a new factory). 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 non-zero costs to bankruptcy and default, which can in turn lead to lower firm values. As such, understanding the relationship between volatility and economic factors is a first-order consideration. For commodities with derivative markets that are illiquid, opaque, with little market depth, or with limited expirations, the findings in this paper can be 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 intervention and evaluate regulatory options that achieve a desired welfare objective.²

Using a reduced form model of a commodity market with power-law distributed consumers and producers, I obtain several hypotheses on how concentration and emerging market demand impacts commodity volatility, and I test these in the data. When commodity supply and demand are dominated by a handful of countries, their idiosyncratic (country-specific) shocks affect global commodity markets. Even in the case where trading partners face homogeneous shocks, market concentrations can have an impact on volatility. Heterogeneous consumers and producers may face supply-demand shocks with different variance. When the larger consumers are also riskier and more volatile (experience higher variance shocks), their impact on market volatility is amplified through concentration. This is important when considering the impact of growing emerging market demand on commodity prices. Many of these markets are highly segmented and pose non-diversifiable risks to hedgers and international investors.

I collate data on 22 major commodity futures markets and the global bilateral trade in the underlying commodities and analyze the extent to which commodity volatility is related to increased emerging market demand and other fundamentals such as inflation uncertainty, while controlling for financial frictions introduced by changing market structure and commodity index trading. Higher concentration in emerging market importers of a commodity is associated with higher futures volatility. The results imply that a 1.00% gain in market concentration by developing country consumers is associated with a 1.19% increase in commodity futures volatility. I find predictability in commodity futures volatility using variables capturing macroeconomic uncertainty, with adjusted R-squared 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

²Amartya Sen (Poverty and Famines (Oxford University Press, 1983)) and others, highlight the direct and potentially catastrophic consequences of commodity price dynamics.
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 in 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 there are significant fluctuations in the realized volatility and correlations of futures returns for the commodities in this study (e.g., Figures 1, 5, and 6). This is true at different horizons corresponding to different holding periods, and over the entire trading history (e.g., beginning in the 1960s for most grain commodities, April 1983 for crude oil, etc.). Large fluctuations in prices and volatility have occurred for the commodities in the sample even prior to the popularization of commodity index and ETF trading. I analyze the determinants of this variation in volatility, selecting potentially relevant proxy variables capturing the variation in global macroeconomic factors, supply-demand, and market frictions based on theory and past empirical studies on commodity risk premia (see, for example, Chen, Rogoff, and Rossi (2010); Hong and Yogo (2012) and Acharya, Lochstoer, and Ramadorai (2013)).

Global commodity markets have undergone major transformations in real economic demand and supply stemming from a sharp increase in demand from emerging market economies over the last two decades. The speed and extent of this increase is larger compared to similar episodes of major global market transformation in recent history. Emerging market economies have become increasingly significant players in many commodity markets. On the supply side, this

3The “index financialization” period is commonly identified in the literature as beginning in January 2004 (Tang and Xiong, 2012; Hamilton and Wu, 2012; Singleton, 2014; Basak and Pavlova, 2013b).

4For instance, Japan’s emergence as a global financial power during the post-World War II period (1960-1970) was accompanied by slower and smaller market share changes to commodity markets compared with the change in China’s share of the major commodity markets since 1990.
has been the case for several decades for some commodities. More recently, global demand has undergone significant changes. As can be seen in the data from UN Comtrade, countries that are not members of OECD and G7 have become the largest buyers in many commodity markets (see, for example, Figures 2 and 3). Developing and emerging market countries have more volatile economies and pose higher levels of legal, political, and economic policy uncertainty (Bekaert and Harvey 1995 1997 2000 Bloom 2014). Bernanke (1983); Bloom, Bond, and Reenen (2007); Bloom (2009) and others find that such uncertainty can lead to higher risk premia, lower investment levels, higher volatility, higher correlation levels, and deeper market distortions which last longer. Pastor and Veronesi (2011 2013) show that such political uncertainty can lead to higher return volatility and correlation levels.

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. 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 commodity futures volatility observed during the past decade may in fact be largely in line with high levels of futures volatility observed during 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., Figure 6). In my research framework, I attempt to distinguish between the impact of these various potential hypotheses for which factors explain commodity futures volatility and recent extreme price moves.

The rest of this document is organized as follows. The next section presents the research framework, including the theoretical motivation underpinning this research. 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.

---

Raghuram G. Rajan (Fault Lines: How Hidden Fractures Still Threaten the World Economy (Princeton University Press, 2010)) discusses the risks associated with different political, legal, and financial systems coming into contact with each other, and how this can generate uncertainty and increase the likelihood of financial market crises.

See footnote 1.
Fig. 1. Time series of annualized rolling realized volatility at different horizons for crude oil, natural gas, and gold 1M, 3M, ..., 36M futures (3-day) returns. Here, short-term volatility refers to the standard deviation for the previous month, and long-term volatility refers to the standard deviation for the previous twelve months. The shaded areas highlight NBER recession periods. The dotted line marks January 2004.
Fig. 2. Global copper imports network. The vertex colors identifies the country group: BRIC (red), non-OECD excluding BRIC (yellow), OECD excluding G7 (green), and G7 (blue). The relative size of a country vertex captures its total import value.
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 to the current price of financial assets react to future changes to 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. In order to understand this correspondence for a future on a storable commodity, begin with the no-arbitrage pricing formula for its futures price (Working, 1933, 1949; Kaldor, 1939; Brennan, 1958; Schwartz, 1997),

\[ F_{t,T} = S_t e^{(r-y)(T-t)}, \]

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. \( y \) can be further decomposed into the “benefit” from holding the physical commodity, \( b \), net of the storage (or carry) cost rate \( m \), \( y = b - m \) and \( r = \pi + \psi \), where is \( \pi \) is the inflation rate and \( \psi \) the real interest rate. This decomposition and the analysis that follow can be applied to any type of future, with the interpretation of \( y \) differing depending on the net benefit to holding the underlying asset, e.g., replace \( y \) with dividend yield \( d \) for stock futures or with the foreign currency interest rate \( r^f \) for currency futures.\(^7\)\(^8\)

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,

\[ F_{n,t} = S_t \left( 1 + R_{n,t} \right)^n, \]

\[ (1 + Y_{n,t}) = \left( \frac{1 + B_{n,t}}{1 + M_{n,t}} \right). \]

Denote the log price at time \( t \) of a future expiring in \( n \) periods as \( f_{n,t} \) and the corresponding

\(^{7}\)This decomposition is exact for the forward price. Due to the mark-to-market gains and losses of the corresponding futures contract, there can be a difference between the forward and future prices unless interest rates are deterministic.

\(^{8}\)Several studies investigate the commodity convenience yield. Casassus and Collin-Dufresne (2005) nests several other models (including Gibson and Schwartz (1990) and Schwartz (1997)) and conclude that convenience yield is increasing in the spot price, interest rates, and the extent to which the underlying commodity is used for production purposes.
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, \( p_{n,t} \),

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

\[
p_{n,t} \equiv f_{n,t} - s_t = n(r_{n,t} - y_{n,t}), \quad (4)
\]

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

\[
\delta_{n,t+1} \equiv p_{n-1,t+1} - p_{n,t}, \quad (5)
\]

\[
= (n - 1)(r_{n-1,t+1} - y_{n-1,t+1}) - n(r_{n,t} - y_{n,t}),
\]

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

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

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

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

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

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

\[9\]As discussed further in section 2.3, non-diversifiable risks or market frictions can cause deviation from the no-arbitrage condition such as producer hedging pressure and borrowing constraints (see, for example, Keynes 1930; Cootner 1960; Hirschleifer 1988; 1990; Roon, Nijman, Chris, and Veld 2000; Acharya, Lochstoer, and Ramadorai 2013). \( x_{n,t+1} \) also captures the part of the futures risk premia due to the deviations from the expectations hypothesis in the interest rate term structure, as shown in Appendix A Eq. (39).
Eq. (6) can be substituted into (9) to obtain its unexpected change,

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

Eq. (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 \( n > 0 \), so that 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 Eq. (40) and (41) in Appendix A).

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

\[ x_{n,t+1} - E_t x_{n,t+1} = \eta_{r,t+1} - \eta_{y,t+1} - \eta_{x,t+1}. \]  \hspace{1cm} (11)

Equation (11) shows unexpected changes to the futures risk premium is due to innovations in future expected convenience yields, interest rates, and excess basis returns. The 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 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. Volatility of the excess basis return is driven by unexpected news affecting interest rates, convenience yield, and risk premia. More explicitly, with correlated components,

\[ Var(x_{n,t+1}) = Var(\eta_{r,t}) + Var(\eta_{y,t}) + Var(\eta_{x,t}) - 2Cov(\eta_{r,t}, \eta_{y,t}) - 2Cov(\eta_{r,t}, \eta_{x,t}) + 2Cov(\eta_{y,t}, \eta_{x,t}) \]  \hspace{1cm} (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,

$$x_{n,t+1} - E_t x_{n,t+1} = \sigma_t \epsilon_t,$$

where $$\epsilon_t|\Omega_{t-1} \sim N(0, 1)$$.

(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 return into $$K$$ observable news sources or risk factors which affect expectations of future cash flows to equity and discount rates, i.e., for the unexpected excess equity return, $$e_t - E_{t-1} e_t = \eta_d^t - \eta_r^t - \eta_e^t = \sum_{k=1}^{K} \beta_k \lambda_{k,t}$$. The equivalent for commodities should use the appropriate information set given the decomposition in (11).

2.2. Producer and consumer impact on commodity market volatility

In this section, I illustrate how producer and consumer risks and concentrations can impact commodity market volatility, motivating my empirical approach to analyzing the effect of rapidly growing emerging market demand.

Consider a model where there are $$p = 1, \ldots, P$$ producers and $$c = 1, \ldots, C$$ consumers of a commodity. A producer $$p$$ has market weight $$w_{p,t}$$ and a consumer has market weight $$w_{c,t}$$ with $$\sum_{p=1}^{P} w_{p,t} = 1$$ and $$\sum_{c=1}^{C} w_{c,t} = 1$$. The distribution of weights is power-law distributed, with a handful of consumers (producers) dominating the demand (supply) side. In this case, the idiosyncratic shocks to the trading parties matter in explaining market dynamics. Similar to formulations in Acharya, Lochstöer, and Ramadorai (2013); Ready, Roussanov, and Ward (2013), consumers have a downward-sloping demand curve for the commodity with the price elasticity of demand $$\epsilon$$, and face an idiosyncratic demand shock $$A_{c,t}$$ such that,

$$S_t = A_{c,t} (Q_{c,t})^{-\frac{1}{\epsilon}}.$$

(14)

In the near-term, producers have a price-inelastic supply and face an idiosyncratic supply shock $$B_{p,t}$$, such that $$Q_{p,t} = B_{p,t}$$. Denote the log quantities and prices in lowercase, with $$a_{c,t} \sim$$

See, for example, Gabaix (2011) for an exposition of this principle applied to firm sizes and aggregate volatility.
Given market clearing for the total change in supply and demand in this setting, \( \sum_{p=1}^{P} w_{p,t} q_{p,t} = \sum_{c=1}^{C} w_{c,t} q_{c,t} \), I can derive the impact of consumer and producer concentration on the variance of the commodity, \( \sigma_{s,t}^2 \),

\[
\sigma_{s,t}^2 = \beta_c \sum_{c=1}^{C} w_{c,t}^2 \sigma_{a,c}^2 + \beta_p \sum_{p=1}^{P} w_{p,t}^2 \sigma_{b,p}^2 + \eta_t. \tag{15}
\]

Consider the case where all consumers and producers have the same distribution in their demand shocks, \( a_{c,t} \sim N(0, \sigma_a) \) and supply shocks, \( b_{p,t} \sim N(0, \sigma_b) \), respectively. Then, defining consumer and producer Herfindahls as \( H_{c,t} = \sum_{c=1}^{C} w_{c,t}^2 \) and \( H_{p,t} = \sum_{p=1}^{P} \lambda_{p,t}^2 \), respectively, yields,

\[
\sigma_{s,t}^2 = \beta_{c} \sigma_{a}^2 H_{c,t} + \beta_{p} \sigma_{b}^2 H_{p,t} + \eta_t. \tag{16}
\]

Eq. 16 shows that even with homogeneous shocks to demand and supply, consumer and producer market concentrations can have an impact on market volatility.\(^{11}\)

Heterogeneous consumers and producers may face supply-demand shocks with different variance. When the larger consumers or producers are also riskier and more volatile (experience higher variance shocks), their impact on market volatility is amplified through concentration. This is important when considering the impact of growing emerging market trade on commodity prices.

Developing and emerging market countries have more volatile economies and greater uncertainty.\(^{12}\) We can consider consumers from emerging market, non-OECD countries have demand shocks, \( a_{c,t} \sim N(0, \sigma_{aEM}) \), while all others have demand shocks, \( a_{c,t} \sim N(0, \sigma_a) \), with \( \sigma_{aEM} > \sigma_a \). All idiosyncratic supply shocks remain uniform, \( b_{p,t} \sim N(0, \sigma_b) \). Starting with Eq. 15, with \( H_{EM} = \sum_{c \in EM} w_{c,t}^2 \),

\[
\sigma_{s,t}^2 = \beta_c \sum_{c=1}^{C} w_{c,t}^2 \sigma_{a,c}^2 + \beta_p \sum_{p=1}^{P} w_{p,t}^2 \sigma_{b,p}^2 + \eta_t,
\]

\[
= \beta_c \sigma_{a}^2 H_{c,t} + \beta_p \sigma_{b}^2 H_{p,t} + \eta_t.
\]

\(11\)This analysis is similar to those on the granular origins of aggregate volatility (see, for example, Gabaix (2011) on the impact of power-law distributed firm sizes on aggregate volatility and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012); Kelly, Lustig, and Nieuwerburgh (2013) on the supplier-customer network and size effects on volatility).
In the near-term, producers are price-inelastic and have essentially fixed supply and no unanticipated shocks ($\sigma_{bp} = 0$). Under these conditions, the second term in Eq. (15) drops out, and producer concentration has no effect on commodity market volatility.

\[ \sigma_{s,t}^2 = \beta_c \sum_{c=1}^{C} w_{c,t}^2 \sigma_{ac}^2 + \beta_p \sum_{p=1}^{P} w_{p,t}^2 \sigma_{bp}^2 + \eta_t, \]

\[ \sigma_{s,t}^2 = \beta_c \sum_{c=1}^{C} w_{c,t}^2 \sigma_{ac}^2 + \eta_t, \]  
\[ = \beta_c \sigma_a^2 H_{c,t} + \beta_c (\sigma_a^{EM} - \sigma_a) H_{c,t}^{EM} + \eta_t. \]

With this discussion in mind, I empirically test several hypotheses related to consumer and producer impact on commodity volatility. These hypotheses capture the impact of concentration and risks of commodity producers and consumers.

Denote the trade weights in commodity $i$ of country $j$ as,

\[ w_{i,j,t}^I = \frac{ImportValue_{i,j,t}}{\sum_{j=1}^{N} ImportValue_{i,j,t}}, \]  
\[ w_{i,j,t}^E = \frac{ExportValue_{i,j,t}}{\sum_{j=1}^{N} ExportValue_{i,j,t}} \]

for imports and exports, respectively. Then, the measures of consumer concentration (of all countries and emerging market countries) are captured through Herfindahl indexes and defined as,

\[ H_{i,t}^C = \left( \sum_{j=1}^{N} (w_{i,j,t}^I)^2 \right)^{1/2}, \]
\[ H_{i,t}^{C,EM} = \left( \sum_{j \in EM} (w_{i,j,t}^I)^2 \right)^{1/2}, \]

respectively. The corresponding Herfindahl indexes for producers, $H_{i,t}^P$ and $H_{i,t}^{P,EM}$, are similarly defined using export weights. For notational simplicity, define $\lambda_{i,j,t}^E = (w_{i,j,t}^E)^2$ and $\lambda_{i,j,t}^I = (w_{i,j,t}^I)^2$.

**Hypothesis 2.1.** Concentration in the importing countries of commodity $i$ impacts its futures
volatility, $Vol_{i,t}$,

$$Vol_{i,t} = \mu_i + \beta_1 H_{i,t}^P + \beta_2 H_{i,t}^C + z_t' \theta + \eta_{i,t}, \tag{24}$$

where, $z_t$ is a state vector of relevant control variables and $\theta$ a vector of coefficients. $\beta_2 > 0$ in the specification in Eq. (24).

**Hypothesis 2.2.** Shocks to the major importers of commodity $i$ impact its futures volatility, $Vol_{i,t}$,

$$Vol_{i,t} = \mu_i + \beta_1 \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_2 \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + z_t' \theta + \eta_{i,t}, \tag{25}$$

$\beta_2 > 0$ in the specification in Eq. (25).

Greater uncertainty and lower financial market development reduces the ability of commodity market participants (producers, consumers, and other investors) to insure against the risks of developing and emerging market countries [Bekaert and Harvey 1995, 2000; di Giovanni and Levchenko 2009; Pastor and Veronesi 2011, 2013; Bloom 2014].

**Hypothesis 2.3.** The relationship in hypotheses 2.1 and 2.2 is more significant for imports from countries that have higher policy uncertainty and lower financial openness (denoted EM countries).

$$Vol_{i,t} = \mu_i + \beta_1 H_{i,t}^P + \beta_2 H_{i,t}^C + \beta_3 H_{i,t}^{P,EM} + \beta_4 H_{i,t}^{C,EM} + z_t' \theta + \eta_{i,t}, \tag{26}$$

$$Vol_{i,t} = \mu_i + \beta_1 \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_2 \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + \beta_3 \sum_{j=1}^{N} I_{[j \in EM]} w_{i,j,t}^E \sigma_{j,t} + \beta_4 \sum_{j=1}^{N} I_{[j \in EM]} w_{i,j,t}^I \sigma_{j,t} + z_t' \theta + \eta_{i,t}, \tag{27}$$

$\beta_4 > 0$ in the specification in Eq. (26) and (27).

**Hypothesis 2.4.** In the short-term, producers hedge, have fixed supply, and have no unanticipated supply shocks affecting commodity markets.

$\beta_1 = 0$ in the specifications in Eq. (24) and (25). $\beta_3 = 0$ in the specifications in Eq. (26) and (27).

The sensitivity of commodity futures to consumer and producer shocks will be highest when
there is scarcity or a glut in the commodity. Such periods would be captured by periods of high absolute values of the futures basis ($\text{HIGH\_BASIS}$). Additionally, the information of demand-side or supply-side pressure should be captured by the gamma coefficient of a GJR-GARCH(1,1) fit of commodity futures daily returns (see Eq. 50 and related discussion in Appendix B).

**Hypothesis 2.5.** The impact of shocks to the major importers of commodity $i$ on its futures volatility, $Vol_{i,t}$, should be highest when the futures term structure exhibits high basis.

$$Vol_{i,t} = \mu_i + \beta_1 \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_2 \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + \epsilon_{i,t} + \beta_3 I_{[t-1 \in \text{HIGH\_BASIS}]}$$

$$+ \beta_4 I_{[t-1 \in \text{HIGH\_BASIS}]} \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_5 I_{[t-1 \in \text{HIGH\_BASIS}]} \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + \eta_{i,t}, \quad (28)$$

where, $\text{HIGH\_BASIS} = 1$ during periods when the absolute value of the futures basis is highest (i.e., when the basis is in the top or bottom quintile), and 0 otherwise.

$\beta_4 > 0$ and $\beta_5 > 0$ in the specification in Eq. (28).

**Hypothesis 2.6.** The impact of shocks to the major importers of commodity $i$ on its futures volatility, $Vol_{i,t}$, should be highest when the asymmetric relationship between commodity volatility and returns is highest.

$$Vol_{i,t} = \mu_i + \beta_1 \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_2 \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + \epsilon_{i,t} + \beta_3 I_{[t-1 \in \text{HIGH\_GAMMA}]}$$

$$+ \beta_4 I_{[t-1 \in \text{HIGH\_GAMMA}]} \sum_{j=1}^{N} w_{i,j,t}^E \sigma_{j,t} + \beta_5 I_{[t-1 \in \text{HIGH\_GAMMA}]} \sum_{j=1}^{N} w_{i,j,t}^I \sigma_{j,t} + \eta_{i,t}, \quad (29)$$

where, $\text{HIGH\_GAMMA} = 1$ during periods when the absolute value of the gamma coefficient of conditional GJR-GARCH(1,1) fits of the commodity futures returns is highest (i.e., when the gamma coefficient is in the top or bottom quintile), and 0 otherwise.

$\beta_4 > 0$ and $\beta_5 > 0$ in the specification in Eq. (29).

I further examine the impact of demand and supply shocks on the conditional variation in the asymmetric relationship between commodity volatility and returns.

**Hypothesis 2.7.**

$$Vol_{i,t} = \mu + \alpha |r_{i,t-1}| + \beta Vol_{i,t-1} + \gamma_{i,t-1} I_{[t-1 \in \text{HIGH\_BASIS}]} |r_{i,t-1}| + \epsilon_{i,t-1}' \theta + \eta_{i,t},$$

$$\gamma_{i,t-1} = \kappa_1 + \kappa_2 a_{i,t-1} + \kappa_3 b_{i,t-1}, \quad (30)$$
where, \( f_{i,t-1}^{(+)} = 1 \) when \( r_{i,t-1} > 0 \), and 0 otherwise. \( a_{i,t-1} \) and \( b_{i,t-1} \) are demand and supply shocks, respectively. \( \kappa = [\kappa_1 \, \kappa_2 \, \kappa_3] \) denote regression coefficients. \( \kappa_2 > 0 \) and \( \kappa_3 > 0 \) in the specification in Eq. (30).

### 2.3. 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. Such conditions can cause Eq. (1) to no longer hold exactly for all investors in the market. In Eq. (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 Poteshman (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 and Scholes (1973), Merton (1973)).

Basak and Pavlova (2013a) model the impact to a stock market from institutional investors (2013b) on the scarcity premium in commodity futures prices and Jovanovic (2013) on the possibility of bubbles in the prices of exhaustible commodities.
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 index stock prices and the correlations between them, and increases overall market volatility. They further find that this tilting is done via increased leverage, so imposing leverage limits, while resulting in reduced riskiness of institutional investor portfolios, will also decrease index stock prices.

The term financialization, in the context of commodities trading, is generally used to describe 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. Implicit in the financialization argument is the assumption that there are binding constraints on investors or other significant frictions such as information asymmetries that cause market inefficiencies to persist 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; 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.\textsuperscript{13}

Several recent studies look at 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 non-energy commodities have become increasingly correlated with oil prices, and that this relationship is stronger for constituent commodities of the SP-GSCI and DJ-UBS indexes. They link this trend to an increased financialization, (mainly via the increased investment in popular commodity indexes 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

\textsuperscript{13}See, for example, Cass (1992); Bhamra and Uppal (2006); Basak, Cass, Licari, and Pavlova (2008); Gromb and Vayanos (2010).
finds futures open interest has important predictive power for crude oil prices, confirming the finding in Hong and Yogo (2012).

Acharya, Lochstoer, and Ramadorai (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 find that when speculator activity is constrained or reduced, the impact of hedging demand increases, i.e., unconstrained speculator activity will help to absorb producer demand shocks. Etula (2013) also finds that the risk-bearing capacity of broker-dealers is predictive of commodity risk premia.

With this discussion in context, 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 after 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 mispricings that exited previously due to limited participation and illiquidity.

If financialization increased access to commodity futures markets by participants such as hedge funds, comovement between futures returns, and large-scale trading activity and portfolio shocks of hedge funds may have increased, to an extent that cannot be absorbed by other market participants due to borrowing constraints, illiquidity, or other market friction that introduces limits to arbitrage.

**Hypothesis 2.8.** Shocks to hedge funds during the financialization period are associated with higher commodity futures volatility.

\[
Vol_{i,t} = \mu_i + \beta_1 HF_{RISK} t_{-1} + z_{t-1}'\theta + \beta_2 I_{[t-1 \in \text{Index Period}]} + I_{[t-1 \in \text{Index Period}]} \ast z_{t-1}'\theta^{INDEX} + \beta_3 I_{[t-1 \in \text{Index Period}]} HF_{RISK} t_{-1} + \eta_{i,t},
\]

where \( HF_{RISK} t \) denotes a proxy capturing hedge fund return activity. \( \beta_3 > 0 \) in the specification in (31).

If there are fewer barriers to trading in commodity futures markets, uninformed trader activity in equity and commodity futures markets become more correlated. Then, given previous work on equity market trading volume and uninformed equity trading (e.g., Campbell,
Following financialization in commodity markets, high trading volume in equity markets should be associated with higher noise trading and higher volatility in commodity futures markets.

**Hypothesis 2.9.** Increased noise trading in equity markets during the financialization period is associated with higher commodity futures volatility.

\[
\text{Vol}_{i,t} = \mu_i + \beta_1 \text{EQTURN}_{t-1} + \mathbf{z}_{t-1}' \Theta + \beta_2 I_{[t-1 \in \text{IndexPeriod}]} \\
+ I_{[t-1 \in \text{IndexPeriod}]} \cdot \mathbf{z}_{t-1}' \Theta^{\text{INDEX}} + \beta_3 I_{[t-1 \in \text{IndexPeriod}]} \text{EQTURN}_{t-1} + \eta_{i,t}, \tag{32}
\]

\(\beta_3 > 0\) in the specification in (32).

### 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, Hong and Yogo, 2012; Engle and Rangel, 2008; Bali, Brown, and Caglayan, 2014). 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 shocks to 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. The 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 prior to 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.
Table 1: Commodity Derivative Contract and Trade Classification Information

This table shows the 19 underlying commodities in the dataset categorized into four market groupings (Energy, Metal, Grain, and Soft). For the commodities for which it is available, daily options price data starts on January 2, 2006. Options on commodities in the Grain grouping are sparsely traded in the first couple of years of the dataset. The naming contract for a futures contract is [Contract code][Expiry month code][Last digit of expiry year], e.g., WTI Crude Oil future expiring in December 2008 is CLZ8. The SITC Rev 2. Code shows the key used to obtain trade flow data from UN Comtrade, matched as the goods flows most closely corresponding to the underlying commodity of the futures contract.

<table>
<thead>
<tr>
<th>Contract Code</th>
<th>Exchange Code</th>
<th>Traded Contract Months</th>
<th>Commodity Name</th>
<th>Futures Data Start</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>NYMEX</td>
<td>All months</td>
<td>Crude Oil</td>
<td>03/30/1983</td>
</tr>
<tr>
<td>HO</td>
<td>NYMEX</td>
<td>All months</td>
<td>Heating Oil</td>
<td>11/14/1978</td>
</tr>
<tr>
<td>NG</td>
<td>NYMEX</td>
<td>All months</td>
<td>Natural Gas</td>
<td>04/04/1990</td>
</tr>
<tr>
<td>PN</td>
<td>NYMEX</td>
<td>All months</td>
<td>Propane</td>
<td>01/03/2005</td>
</tr>
<tr>
<td>Metal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>COMEX</td>
<td>G</td>
<td>J</td>
<td>M</td>
</tr>
<tr>
<td>SI</td>
<td>COMEX</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>HG</td>
<td>COMEX</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>PA</td>
<td>NYMEX</td>
<td>H</td>
<td>M</td>
<td>U</td>
</tr>
<tr>
<td>PL</td>
<td>NYMEX</td>
<td>F</td>
<td>J</td>
<td>N</td>
</tr>
<tr>
<td>Grain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>CBOT</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>C</td>
<td>CBOT</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>O</td>
<td>CBOT</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>S</td>
<td>CBOT</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
<tr>
<td>SM</td>
<td>CBOT</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
<tr>
<td>BO</td>
<td>CBOT</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
<tr>
<td>RR</td>
<td>CBOT</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
<tr>
<td>Soft</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>NYCE</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>OJ</td>
<td>NYCE</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
<tr>
<td>KC</td>
<td>CSCE</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>SB</td>
<td>CSCE</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>CC</td>
<td>CSCE</td>
<td>H</td>
<td>K</td>
<td>N</td>
</tr>
<tr>
<td>LB</td>
<td>CME</td>
<td>F</td>
<td>H</td>
<td>K</td>
</tr>
</tbody>
</table>
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, 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 decom-
pose 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, 2006) study the cross-sectional variation in risk premia and idiosyn-
cratic volatility and find a significant positive relationship between the two. Campbell, Giglio.
Polk, and Turley (2012) 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&P500 index, individual US stocks, and currency markets for a variety of reasons in-
cluding easy access to the relevant data, long time periods, liquidity, and 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 $R_{fFD_{d,t}}$. For
example, the 5-day rolling return of the (fixed-term) 3-month commodity future at time $t$ is
denoted $R_{3F5D}^{5D_{t},f}$. 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, $\sigma_{t,t} \approx \sqrt{\text{Vol}_{fFD_{d,t}}^{mM}}$, 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 as 5, 21, and 252 trading days,
respectively. The baseline panel regressions use the (non-overlapping) end-of-month (EOM) volatility of daily returns of the 1-month future, $Vol_{i,t} = Vol_{1,t}^{F1DEOM}$ as the dependent variable, except where I explicitly state 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 A.III in the online Appendix). The baseline predictive regressions take the form,

$$Vol_{i,t} = \mu + \alpha |r_{i,t-1}| + \beta Vol_{i,t-1} + z_{i,t-1}'\theta + \eta_{i,t}, \quad (35)$$

where $r_{i,t} = R_{1,t}^{F21D}$, $z_{i,t}$ is a vector of $K$ (non-negative) explanatory variables, $\alpha$, $\beta$, and the vector $\theta$ denote regression coefficients. In Appendix B 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. In the online Appendix, I show results for the corresponding regressions with the natural logarithm of $Vol_{i,t}$ in place of the level.

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 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 5 for crude oil, copper, gold, natural gas, wheat, and lumber.

Relative to commodities in energy, grain and softs, precious metals broadly show little variation in 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 and 5). 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

\[Vol_{i,t}^{F21D} = \sqrt{\frac{252}{d} \left[ \frac{1}{N} \sum_{p=1}^{N} \left( R_{i,t-N+p}^{F21D} \right)^2 - \left( \frac{1}{N} \sum_{p=1}^{N} R_{i,t-N+p}^{F21D} \right)^2 \right]^2}, \quad (33)\]

\[Vol_{i,t}^{F1D1M} = \sqrt{\frac{252}{21} \left[ \frac{1}{21} \sum_{p=1}^{21} \left( R_{i,t-21+p}^{F1D} \right)^2 - \left( \frac{1}{21} \sum_{p=1}^{21} R_{i,t-21+p}^{F1D} \right)^2 \right]^2}, \quad (34)\]

where $N = m * \frac{21}{d}$ is the number of $d$-day return periods in $m$ months. The factor $\frac{252}{d}$ annualizes the volatility.
Table 2: Commodity Futures Volatility Summary Statistics

This table 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 the online Appendix, we show the same summary statistics by decade.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1M</td>
<td>3M</td>
</tr>
<tr>
<td>Copper</td>
<td>25.295</td>
<td>24.794</td>
</tr>
<tr>
<td>Sugar</td>
<td>38.880</td>
<td>35.493</td>
</tr>
<tr>
<td>Lumber</td>
<td>27.263</td>
<td>24.579</td>
</tr>
</tbody>
</table>

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 uncertainty in supply-demand 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, whereas 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 in Bekaert and Wu (2000); Bollerslev and Todorov (2011) and elsewhere, 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 more 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
This table 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 S&P 500 index, FTSE 100 index, and the US-GBP exchange rate. Energy, Grain, Metal, Soft and All correspond to equal-weighted indexes of the constituent commodity futures as shown in Table I. The LRT column shows the likelihood ratio test statistic (K=1, critical value (P = 0.05) =3.841). The series cover the period from September 1988 to December 2011, yielding 5,636 observations of daily log returns.

<table>
<thead>
<tr>
<th></th>
<th>Omega</th>
<th>Alpha</th>
<th>Beta</th>
<th>LL</th>
<th>Omega</th>
<th>Alpha</th>
<th>Beta</th>
<th>Gamma</th>
<th>LL</th>
<th>LRT</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>0.010</td>
<td>0.068</td>
<td>0.925</td>
<td>-7,692.3</td>
<td>0.017</td>
<td>0.010</td>
<td>0.918</td>
<td>0.117</td>
<td>-7,618.7</td>
<td>147.2</td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.018</td>
<td>0.086</td>
<td>0.904</td>
<td>-8,474.5</td>
<td>0.022</td>
<td>0.028</td>
<td>0.913</td>
<td>0.090</td>
<td>-8,434.9</td>
<td>79.2</td>
</tr>
<tr>
<td>USD-GBP</td>
<td>0.003</td>
<td>0.038</td>
<td>0.954</td>
<td>-4,890.1</td>
<td>0.003</td>
<td>0.032</td>
<td>0.953</td>
<td>0.011</td>
<td>-4,888.5</td>
<td>3.2</td>
</tr>
<tr>
<td>All Commodities</td>
<td>0.003</td>
<td>0.039</td>
<td>0.957</td>
<td>-6,339.4</td>
<td>0.002</td>
<td>0.045</td>
<td>0.959</td>
<td>-0.012</td>
<td>-6,337.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Grain</td>
<td>0.017</td>
<td>0.069</td>
<td>0.923</td>
<td>-9,056.4</td>
<td>0.014</td>
<td>0.081</td>
<td>0.929</td>
<td>-0.032</td>
<td>-9,048.8</td>
<td>15.2</td>
</tr>
<tr>
<td>Metal</td>
<td>0.011</td>
<td>0.055</td>
<td>0.937</td>
<td>-8,107.8</td>
<td>0.010</td>
<td>0.062</td>
<td>0.938</td>
<td>-0.014</td>
<td>-8,106.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Energy</td>
<td>0.052</td>
<td>0.073</td>
<td>0.917</td>
<td>-11,678.0</td>
<td>0.051</td>
<td>0.076</td>
<td>0.919</td>
<td>-0.009</td>
<td>-11,677.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Softs</td>
<td>0.035</td>
<td>0.057</td>
<td>0.913</td>
<td>-8,177.7</td>
<td>0.035</td>
<td>0.059</td>
<td>0.912</td>
<td>-0.003</td>
<td>-8,177.7</td>
<td>0.0</td>
</tr>
</tbody>
</table>

3.3. Commodity supply-demand

The main source of data used for global commodity trade flows is UN Comtrade. I match each of the commodity futures contracts in the study to a particular commodity code in UN Comtrade, determined to be most closely related to the underlying commodity. The commodity product classification used is the Standard International Trade Classification (SITC) Revision 2, in order to obtain the longest possible time series (Table I). The global bilateral trade flow information for the matched commodity is the proxy used for the aggregate supply and demand for the commodity underlying the futures contracts. I document the changes to global commodity supply and demand since 1973 for these matched commodities.

Figure illustrates corresponding trends in trade by emerging countries in several commodities markets using annual UN Comtrade from 1978 to 2010. The figures show the trade value as a percentage of total world trade for BRIC, non-OECD excluding BRIC, OECD excluding G7, and G7 countries, in wheat, lumber, cotton, crude oil, copper, and aluminum. For most commodities in the dataset, the percentage of total world trade value for both imports and exports increases for BRIC countries, and decreases for G7 countries during the past decade.

I construct measures of concentration and economic uncertainty for commodity consumers
Fig. 3. Import value (by country group) as a percentage of total world imports. The four groups shown are, starting from the bottom (darkest to lightest shading): BRIC, non-OECD excluding BRIC, OECD excluding G7, and G7 countries, respectively.
and producers for each commodity as defined in Eq. (22) and Eq (23). It is common that both the supply and demand side of global trade in a commodity are dominated by a handful of countries. As a result, it is possible to take the set of largest exporters and largest importers for each year to characterize the global supply and demand dynamics for each commodity. In constructing trade-weighted indexes for producers (consumers) for a particular commodity as in (23) and (36), I take the set of (minimum five) countries that constitute at least 50% of the total global exports (imports) of that commodity when constructing the producer (consumer) index. In this case, \( C_{\text{HHI},i,t} \approx H_{i,t}^C \), \( C_{\text{HHI},EM,i,t} \approx H_{i,t}^{C,EM} \), \( P_{\text{HHI},i,t} \approx H_{i,t}^P \), and \( P_{\text{HHI},EM,i,t} \approx H_{i,t}^{P,EM} \). The empirical results are robust to the choice of these levels.

I obtain seasonally adjusted quarterly GDP growth rate series using raw data from IMF International Financial Statistics (IFS). I discard any country without at least nine quarterly GDP observations and obtain a dataset of 81 countries. The GDP volatility for producers \( \sigma_{i,t}^2 \) and consumers \( \sigma_{i,t}^2 \) for commodity \( i \) at time \( t \) is constructed by averaging over the squared absolute value of the innovations from an AR(1) fit of all exporters and importers, respectively. For a country \( j \), the trade weights are as defined in Eqs. (20) and (21).

\[
\Delta \log(\text{GDP})_{j,t} = \mu_j + \rho_j \Delta \log(\text{GDP})_{j,t-1} + \epsilon_{j,t},
\]

\[
\sigma_{j,t}^2 = \frac{1}{4} \sum_{k=t-3}^{t} |\epsilon_{j,t}|^2,
\]

\[
C_{VOL_{i,t}} = \left[ \sum_{j=1}^{N} (w_{i,j,t}^I)^2 \sigma_{j,t}^2 \right]^{\frac{1}{2}}.
\]  

Building on findings that link commodity currency returns to commodity futures returns, I construct producer and consumer FX volatility series an explanatory variable: \( P_{FX_{i,t}} = \left[ \sum_{j=1}^{N} (w_{i,j,t}^P)^2 x_{j,t}^2 \right]^{\frac{1}{2}} \), and the corresponding series for \( C_{FX_{i,t}} \) for importers, where \( x_{j,t} \) is the return at time \( t \) of the US dollar exchange rate of the country \( j \) currency. All exchange rate data are collated from Datastream and the Federal Reserve Board to obtain the longest available time series.

\[ \text{Chen, Rogoff, and Rossi (2010) find that the exchange rates of countries which are the major exporters of commodities strongly predict world commodity prices, while the reverse relationship is not as strong. They find some evidence "commodity currency" returns Granger-cause global commodity futures returns. This has implications in terms of commodity price hedging, especially for commodities whose forward markets have reduced horizon and depth.} \]
3.4. Market activity

I obtain information on the evolution of different types of traders (classified as commercial (hedger), non-commercial, spread, or non-reporting (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 4 shows the variation of 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) have increased substantially. Moreover, the imbalance in commercial positions generally appear to be opposite the imbalance in non-commercial positions.

The set of variables identified from previous work that look into the impact of speculator activity on commodity futures returns (Hong and Yogo, 2012; Acharya, Lochstoer, and Ramadorai, 2013) used as explanatory variables in $z_{i,t}$ include changes to open interest and demand imbalance, e.g., using commercial (“hedger”) position values collated by the CFTC,

$$HEDGER\text{IMB}_{i,t} = \frac{\text{ShortOI}_{i,t} - \text{LongOI}_{i,t}}{\text{ShortOI}_{i,t} + \text{LongOI}_{i,t}}$$

I use an indicator for the period beginning January 2004, commonly cited in previous work as the period showing index “financialization” (see, for example, Tang and Xiong, 2012; Singleton, 2014), as $\text{IndexPeriod}_t$, 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 ($HF\text{RET}_i$) using hedge fund data collated from the Lipper-TASS, BarclayHedge, Morningstar, HFR and CISDM databases.

3.5. Macroeconomic uncertainty indicators

I use the IMF World Economic Outlook Database for aggregate economic variables, and IMF Direction of Trade Statistics for country-to-country aggregate import/export data. Both these sources provide data at an annual frequency. All interest rates and exchange rates are from the Global Financial Database (GFD) and Datastream. Wherever necessary, World Bank classifications are used to group world economies. US GDP and CPI (quarterly) forecast statistics

---

16 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.

Fig. 4. The evolution of contract positions for Crude Oil futures broken down by trader type in CFTC Commitment of Traders (COT) reports.
are from Philadelphia Federal Reserve Bank’s Survey of Professional Forecasts. Economic forecasts for any other countries are from analyst forecasts collated in Bloomberg. US recession period data are from NBER.


Other than CFNAI (May 1967), TED (January 1971) and VXO (January 1986), these variables are available from January 1960 to the end of the sample period. 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. Consumer and producer impact

Table 4 shows results from regressions using as explanatory variables consumer and producer trade-weighted indexes that capture supply-demand uncertainty and vulnerability to shocks. Panel A and B show results with year-over-year changes and level of the Herfindahl indexes, respectively. Panel C shows results with trade-weighted volatility indexes of consumer and producer shocks. In Panel A, columns 1 to 3 show results from regressions including the change in the Herfindahl index of all major trading countries (HHI_ALL), without separating out non-OECD countries. Under all three regression specifications, only the consumer Herfindahl index has a positive significant coefficient with a t-statistic of 3.57, while the coefficient for producer concentration is not significant. This is in line with the predictions set forth in section 2.2. There
is a significant impact on futures volatility from consumer concentration. Next, I consider the heterogeneity in shocks between the two groups, OECD and non-OECD.

Columns 4 to 6 in Table 4 show the same regressions with only non-OECD countries (HHI_EM), with the weights of OECD countries replaced with zero in the index. Again, the coefficient on the non-OECD consumer concentration index (CONS_HHI_EM) is the only one that is positive and significant, with a t-statistic of 6.38. The final three columns show the results when all four indexes are included. The coefficient for CONS_HHI_EM remains essentially unchanged, with a t-statistic of 4.71. These results imply that a 1% gain in market concentration by developing country consumers is associated with a 1.19% gain in commodity futures volatility in this period. Regardless of the idiosyncratic variation within the two groups of consumers, controlling for the heterogeneity across the two groups allows us to capture the differential impact of emerging market countries on commodity volatility. These findings are in agreement with previous work that find emerging markets pose greater uncertainty (Bloom, 2014). These results show how this uncertainty may affect commodity futures.

These results showing the significance of non-OECD consumers are robust to using the level or change in HHI indexes and the inclusion of year fixed effects.

Next, using conditional GJR-GARCH fits of commodity futures returns, I examine the time-variation in the asymmetric relationship between returns and volatility, and analyze how this relates to the commodity basis and sensitivity to consumer and producer shocks (hypothesis 2.6).

4.2. Macroeconomic uncertainty

Tables 5 to 8 show results for 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 (35). The volatility series are at a non-overlapping monthly frequency. The results shown are for the commodity groups: Energy, Metal, Grain, Softs, 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 italics below each coefficient estimate. All regressions for a particular dependent variable include observations on the same dates, allowing for the comparison of information criterion.\(^{18}\)

\(^{18}\) In the discussion of regression results that follow, model fit is considered using likelihood ratio tests, denoted LRT (Neyman and Pearson 1933; Wilks 1938; Engle 1984), and Schwarz Bayesian information criteria (also know as the Bayesian information criterion), denoted BIC (Schwarz 1978). LRT can be used to compare two
### Table 4: Commodity Futures VolatilityProducer and Consumer Uncertainty

This table shows results for the balanced panel regressions of 1-month volatility of the front-month futures return, Vol(t), as the dependent variable in regressions 1 through 9 in Panels A, B, and C. The regressions shown in Panel A include year-over-year changes to producer (exporter) and consumer (importer) concentration indexes as the independent variables, whereas Panel B regressions use the level of the concentration indexes. The possible values of the HHI concentration indexes range from 0 to 1, so that the change in the concentration index is between -1 and 1. Panel C regressions include the trade-weighted volatility indexes for producer and consumer country shocks (to quarterly GDP). The results reported here are for all commodities over the entire period of the sample (262 months) for 4,454 commodity-month observations. All regressions include commodity and season (month) fixed effects. Return variables are in percentage. T-statistics clustered by month are shown in italics below each coefficient estimate.

#### Panel A: Changes to producer and consumer concentrations in global trade

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \text{PROD}_{HHI, \text{ALL}}(t) )</td>
<td>-0.138</td>
<td>-0.129</td>
<td>-0.235**</td>
<td>-0.228*</td>
<td>(-1.415)</td>
<td>(-1.303)</td>
<td>(-2.032)</td>
<td>(-1.949)</td>
<td>(-1.605)</td>
</tr>
<tr>
<td>( \Delta \text{PROD}_{HHI, \text{EM}}(t) )</td>
<td>0.045</td>
<td>0.050</td>
<td>0.252</td>
<td>0.251</td>
<td>(0.336)</td>
<td>(0.369)</td>
<td>(1.617)</td>
<td>(1.605)</td>
<td>(1.605)</td>
</tr>
<tr>
<td>( \Delta \text{CONS}_{HHI, \text{ALL}}(t) )</td>
<td>0.365***</td>
<td>0.360***</td>
<td>(3.571)</td>
<td>(3.496)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta \text{CONS}_{HHI, \text{EM}}(t) )</td>
<td>1.192***</td>
<td>1.192***</td>
<td>(6.383)</td>
<td>(6.391)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adjusted R-squared: 0.162
BIC: 34,943.1

#### Panel B: Level of producer and consumer concentrations in global trade

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{PROD}_{HHI, \text{ALL}}(t) )</td>
<td>-0.010</td>
<td>-0.010</td>
<td>-0.122</td>
<td>-0.216</td>
<td>(-0.188)</td>
<td>(-0.195)</td>
<td>(-1.995)**</td>
<td>(-3.518)**</td>
<td>(-6.630)</td>
</tr>
<tr>
<td>( \text{PROD}_{HHI, \text{EM}}(t) )</td>
<td>0.170</td>
<td>-0.199</td>
<td>0.295</td>
<td>0.087</td>
<td>(1.594)</td>
<td>(-1.743)*</td>
<td>(2.267)**</td>
<td>(-0.630)</td>
<td></td>
</tr>
<tr>
<td>( \text{CONS}_{HHI, \text{ALL}}(t) )</td>
<td>-0.061</td>
<td>-0.061</td>
<td>-0.378</td>
<td>-0.422</td>
<td>(-1.417)</td>
<td>(-1.415)</td>
<td>(-5.820)**</td>
<td>(-6.431)**</td>
<td>(-7.952)**</td>
</tr>
<tr>
<td>( \text{CONS}_{HHI, \text{EM}}(t) )</td>
<td>0.476</td>
<td>0.525</td>
<td>0.671</td>
<td>0.773</td>
<td>(6.901)**</td>
<td>(6.702)**</td>
<td>(7.677)**</td>
<td>(7.952)**</td>
<td></td>
</tr>
</tbody>
</table>

Adjusted R-squared: 0.161
BIC: 34,824.1

#### Panel C: Trade-weighted volatility indexes of producer and consumer shocks

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{PROD}_{VOL, \text{ALL}}(t) )</td>
<td>0.094*</td>
<td>0.072</td>
<td>0.023</td>
<td>0.028</td>
<td>(1.750)</td>
<td>(1.417)</td>
<td>(0.291)</td>
<td>(0.365)</td>
<td>(0.291)</td>
</tr>
<tr>
<td>( \text{PROD}_{VOL, \text{EM}}(t) )</td>
<td>0.099**</td>
<td>0.082**</td>
<td>0.083</td>
<td>0.054</td>
<td>(2.518)</td>
<td>(2.215)</td>
<td>(1.410)</td>
<td>(0.929)</td>
<td>(2.495)</td>
</tr>
<tr>
<td>( \text{CONS}_{VOL, \text{ALL}}(t) )</td>
<td>0.165***</td>
<td>0.158***</td>
<td>0.113**</td>
<td>0.105**</td>
<td>(3.589)</td>
<td>(3.614)</td>
<td>(2.567)</td>
<td>(2.495)</td>
<td>(2.495)</td>
</tr>
<tr>
<td>( \text{CONS}_{VOL, \text{EM}}(t) )</td>
<td>0.462***</td>
<td>0.451***</td>
<td>0.405***</td>
<td>0.401***</td>
<td>(6.327)</td>
<td>(6.375)</td>
<td>(6.190)</td>
<td>(6.150)</td>
<td>(6.150)</td>
</tr>
</tbody>
</table>

Adjusted R-squared: 0.163
BIC: 34,938.2

---

31
In Table 5, 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 Eq. (50)) and allows for the capture of any asymmetric affect on volatility from the direction of the lagged return. Similar to the model fits in Table 5, these results also show that unlike in the case of equities (Bekaert and 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 and Wu 2000; Bollerslev and Todorov 2011; Garleanu, Pedersen, and Poteshman 2009), the conditional variation in this relationship bears further study in the commodities space. In later analysis, I examine the impact of demand and supply shocks on the conditional variation in this relationship using the specification in Eq. (30).

In Table 6, I add the variables capturing macroeconomic uncertainty. This results in an adjusted R-squared gain of over 10% (for the Energy group) from the baseline specification in Table 5, 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 adjusted R-squared 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. These results are in agreement with the implications of the derivation in section 2.1, which show that variation in commodity futures volatility arise from 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 and Veronesi 2013). Economic activity is related nested models. More generally, 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).

19See Table A.VII of the online Appendix for results of Granger causality tests that show the direction of predictive causality is from the economic uncertainty variables included here to commodity futures volatility, and not the other way around.
## Table 5: Commodity Futures Volatility - Panel Regression Results

This table shows results for balanced panel regressions of (time, t) 1-month volatility of the front-month futures return, $\text{Vol}(t)$, over lagged (time, t-1) volatility and (abs) return. PostiveReturn is the absolute return series of the front month future with negative return months set to 0. The results reported here are for the commodity groups Energy, Metal, Grain, Soft, and All (all 19 commodities in sample, see Table I). All regressions include commodity and season (month) fixed effects. Return variables are in percentage. T-statistics clustered by month are shown in italics below each coefficient estimate.

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Metal</th>
<th>Grain</th>
<th>Soft</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Baseline panel regression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\text{Intercept})$</td>
<td>20.671</td>
<td>6.765</td>
<td>10.213</td>
<td>16.249</td>
<td>11.335</td>
</tr>
<tr>
<td></td>
<td>(4.571)</td>
<td>(5.084)</td>
<td>(7.897)</td>
<td>(8.419)</td>
<td>(10.582)</td>
</tr>
<tr>
<td>$\text{Return}_{(t-1)}$</td>
<td>0.617</td>
<td>0.416</td>
<td>0.227</td>
<td>0.371</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>(4.074)</td>
<td>(3.321)</td>
<td>(3.137)</td>
<td>(4.119)</td>
<td>(6.781)</td>
</tr>
<tr>
<td>$\text{Vol}_{(t-1)}$</td>
<td>0.412</td>
<td>0.478</td>
<td>0.520</td>
<td>0.336</td>
<td>0.428</td>
</tr>
<tr>
<td></td>
<td>(4.872)</td>
<td>(8.795)</td>
<td>(12.203)</td>
<td>(5.458)</td>
<td>(10.647)</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.343</td>
<td>0.437</td>
<td>0.419</td>
<td>0.252</td>
<td>0.395</td>
</tr>
<tr>
<td>Number of commodity-months</td>
<td>524</td>
<td>1,310</td>
<td>1,310</td>
<td>1,310</td>
<td>4,454</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Metal</th>
<th>Grain</th>
<th>Soft</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel B: Baseline panel regression allowing for asymmetric return effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(\text{Intercept})$</td>
<td>20.739</td>
<td>6.768</td>
<td>10.227</td>
<td>16.260</td>
<td>11.349</td>
</tr>
<tr>
<td></td>
<td>(4.606)</td>
<td>(5.087)</td>
<td>(7.892)</td>
<td>(8.567)</td>
<td>(10.546)</td>
</tr>
<tr>
<td>$\text{Return}_{(t-1)}$</td>
<td>0.721</td>
<td>0.399</td>
<td>0.208</td>
<td>0.297</td>
<td>0.362</td>
</tr>
<tr>
<td></td>
<td>(4.125)</td>
<td>(3.040)</td>
<td>(2.409)</td>
<td>(2.329)</td>
<td>(4.490)</td>
</tr>
<tr>
<td>$I[\text{Return} &gt; 0] \times \text{Return}_{(t-1)}$</td>
<td>-0.141</td>
<td>0.025</td>
<td>0.031</td>
<td>0.118</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(-0.733)</td>
<td>(0.145)</td>
<td>(0.313)</td>
<td>(1.094)</td>
<td>(0.525)</td>
</tr>
<tr>
<td>$\text{Vol}_{(t-1)}$</td>
<td>0.400</td>
<td>0.479</td>
<td>0.520</td>
<td>0.338</td>
<td>0.429</td>
</tr>
<tr>
<td></td>
<td>(4.851)</td>
<td>(8.628)</td>
<td>(12.159)</td>
<td>(5.629)</td>
<td>(10.833)</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.343</td>
<td>0.436</td>
<td>0.419</td>
<td>0.253</td>
<td>0.395</td>
</tr>
<tr>
<td>Number of commodity-months</td>
<td>524</td>
<td>1,310</td>
<td>1,310</td>
<td>1,310</td>
<td>4,454</td>
</tr>
<tr>
<td>LRT statistic</td>
<td>1.60</td>
<td>0.20</td>
<td>0.20</td>
<td>3.20</td>
<td>1.40</td>
</tr>
<tr>
<td>$[K - K = 1</td>
<td>(P = 0.05) = 3.841]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
to the convenience yield (Casassus and Collin-Dufresne, 2005). The results once I control for recession periods (section 4.4) also capture the variation in risk premia associated with the business cycle. Moreover, these results broadly confirm observations on the effects of uncertainty in other markets (Bloom, 2014).

Table 6: Commodity Futures Volatility and Macroeconomic Uncertainty

This table shows results for balanced panel regressions of (time, t) 1-month volatility of the front-month futures return 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 (Table A.V and A.VI in the Online Appendix contains details of this PCA). The results reported here are for the commodity groups Energy, Metal, Grain, Soft, and All (all 19 commodities in sample, see Table I). All regressions include commodity and season (month) fixed effects. Return variables are in percentage. T-statistics clustered by month are shown in italics below each coefficient estimate. The LRT row shows the likelihood ratio test statistic comparing the fit shown in Table 5.A to the current results.

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Metal</th>
<th>Grain</th>
<th>Softs</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCP1(_{(t-1)})</td>
<td>-2.166***</td>
<td>-1.030***</td>
<td>-0.890***</td>
<td>-0.630***</td>
<td>-0.952***</td>
</tr>
<tr>
<td></td>
<td>(-4.975)</td>
<td>(-4.723)</td>
<td>(-4.991)</td>
<td>(-4.552)</td>
<td>(-6.934)</td>
</tr>
<tr>
<td>PCP2(_{(t-1)})</td>
<td>-1.356**</td>
<td>-0.417</td>
<td>0.037</td>
<td>0.380</td>
<td>-0.155</td>
</tr>
<tr>
<td></td>
<td>(-2.139)</td>
<td>(-1.473)</td>
<td>(0.123)</td>
<td>(1.468)</td>
<td>(-0.759)</td>
</tr>
<tr>
<td>PCP3(_{(t-1)})</td>
<td>2.261***</td>
<td>0.538*</td>
<td>0.291</td>
<td>0.426*</td>
<td>0.586***</td>
</tr>
<tr>
<td></td>
<td>(4.616)</td>
<td>(1.903)</td>
<td>(1.114)</td>
<td>(1.749)</td>
<td>(3.968)</td>
</tr>
<tr>
<td>PCP4(_{(t-1)})</td>
<td>1.188**</td>
<td>1.690***</td>
<td>1.513***</td>
<td>-0.320</td>
<td>0.977***</td>
</tr>
<tr>
<td></td>
<td>(2.005)</td>
<td>(5.011)</td>
<td>(3.981)</td>
<td>(-0.997)</td>
<td>(4.489)</td>
</tr>
</tbody>
</table>

All predictors in Table 5 | Yes | Yes | Yes | Yes | Yes
Adjusted R-squared | 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

\[ K - K = 4 | P = 0.05 = 9.488 \]

It is difficult to contemporaneously explain let alone predict financial asset volatility using economic factors (see, for example, Roll (1984); Schwert (1989); Engle and Rangel (2008); Engle, Ghysels, and Sohn (2013)), even when model results and economic intuition posit a relationship between economic conditions and volatility. As such, the results in Table 6 are 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, for example, Campbell and Thompson (2008); Inoue and Kilian (2004) and Moskowitz, Ooi, and Pedersen (2012) for further discussion on the value of time series predictability). An adjusted R-squared gain over the baseline model is useful for investors who have a non-zero “vega” exposure in their portfolio (\( \frac{\delta V}{\delta \sigma} \neq 0 \) 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 A.IV of the online Appendix presents details of the principal component analysis. The panels in Table A.V show regressions results with the included number of principal components varied.\footnote{See Bai and Ng (2002); Stock and Watson (2002) on selecting the appropriate number of factors.}

In future work, I include uncertainty proxies directly based on work in Jurado, Ludvigson, and Ng (2014).

4.3. Hedging and trading activity

Table 7 shows results once variables capturing momentum and hedging activity (Hong and Yogo, 2012; Acharya, Lochstoer, and Ramadorai, 2013) is added to the specification in Table VI, Panel B. While there is some improvement, there is no consistent in gain in predictive power. Table A.VI of the online Appendix shows the results without including the macroeconomic controls. The regressions including only economic uncertainty variables in Table VI perform better on the dimensions of adjusted R-squared and information criterion measures of model fit.

Table 8 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 7. 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 and 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. Time-variation

Table 9 shows comparison of model fits, in addition to results with interactions for different time periods added to the specification in Table 7. In column 5, I interact for the NBER recession periods as in regression specification (37) and in column 6, I show results from interacting...
Table 7: Commodity Futures Volatility and Commodity Market Risk Factors

This table shows results for balanced panel regressions of (time, t) 1-month volatility of the front-month futures return over lagged (time, t-1) commodity market variables in addition to the macroeconomic uncertainty factors included in Table VLB. The results reported here are for the commodity groups Energy, Metal, Grain, Soft, and All (all 19 commodities in sample, see Table I). All regressions include commodity and season (month) fixed effects. Return variables are in percentage. T-statistics clustered by month are shown in italics below each coefficient estimate. The LRT row shows the likelihood ratio test statistic comparing the fit shown in Table 6 to the current results.

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Metal</th>
<th>Grain</th>
<th>Softs</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOM$_{(t-1)}$</td>
<td>0.804</td>
<td>4.009***</td>
<td>0.503</td>
<td>-0.615</td>
<td>1.767***</td>
</tr>
<tr>
<td></td>
<td>(0.681)</td>
<td>(3.538)</td>
<td>(0.561)</td>
<td>(-0.665)</td>
<td>(2.882)</td>
</tr>
<tr>
<td>HEDGER$_{OIG}$(t-1)</td>
<td>1.235</td>
<td>-0.177</td>
<td>0.282</td>
<td>0.774</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>(1.245)</td>
<td>(-0.388)</td>
<td>(0.863)</td>
<td>(1.201)</td>
<td>(1.149)</td>
</tr>
<tr>
<td>HEDGER$_{IMB}$(t-1)</td>
<td>-0.153</td>
<td>-0.021</td>
<td>-0.069</td>
<td>-0.114</td>
<td>-0.089**</td>
</tr>
<tr>
<td></td>
<td>(-0.806)</td>
<td>(-0.472)</td>
<td>(-0.786)</td>
<td>(-1.338)</td>
<td>(-2.328)</td>
</tr>
<tr>
<td>All predictors in Table 6</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.449</td>
<td>0.483</td>
<td>0.46</td>
<td>0.264</td>
<td>0.428</td>
</tr>
<tr>
<td>Number of commodity-months</td>
<td>524</td>
<td>1,310</td>
<td>1,310</td>
<td>1,310</td>
<td>4,454</td>
</tr>
<tr>
<td>LRT statistic</td>
<td>9.6</td>
<td>27.2</td>
<td>5.0</td>
<td>4.4</td>
<td>41.2</td>
</tr>
</tbody>
</table>

[\sqrt{K-K} = 3|P = 0.05) = 7.815]

Table 8: Commodity Futures Volatility and Hedge Fund Activity

This table shows results for balanced panel regressions of (time, t) 1-month volatility of the front-month futures return over lagged (time, t-1) absolute value of the hedge fund industry mean return, in addition to all explanatory variables included in Table 7. The results reported here are for the commodity groups Energy, Metal, Grain, Soft, and All (all 19 commodities in sample, see Table I) 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 italics below each coefficient estimate.

<table>
<thead>
<tr>
<th></th>
<th>Energy</th>
<th>Metal</th>
<th>Grain</th>
<th>Softs</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H F_{RET}$(t-1)</td>
<td>0.684</td>
<td>0.220</td>
<td>0.864***</td>
<td>0.106</td>
<td>0.421*</td>
</tr>
<tr>
<td></td>
<td>(1.435)</td>
<td>(0.611)</td>
<td>(2.585)</td>
<td>(0.428)</td>
<td>(1.798)</td>
</tr>
<tr>
<td>$Index_{Period}(t) * H F_{RET}$(t-1)</td>
<td>1.085***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.627)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Recession_{(t)} * H F_{RET}$(t-1)</td>
<td>1.615**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.983)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All predictors in Table 7</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.451</td>
<td>0.473</td>
<td>0.440</td>
<td>0.250</td>
<td>0.416</td>
</tr>
<tr>
<td>Number of commodity-months</td>
<td>376</td>
<td>940</td>
<td>940</td>
<td>940</td>
<td>3,196</td>
</tr>
<tr>
<td>Number of commodity-months</td>
<td>376</td>
<td>940</td>
<td>940</td>
<td>940</td>
<td>3,196</td>
</tr>
</tbody>
</table>
with \textit{IndexPeriod} as in regression specification (38).

\begin{align*}
\text{Vol}_t & = \mu_i + NBER_{Recession} + z_{t-1}'\theta + NBER_{Recession} * z_{t-1}'\theta^{REC} + \eta_t, \\
\text{Vol}_t & = \mu_i + \text{IndexPeriod} + z_{t-1}'\theta + \text{IndexPeriod} * z_{t-1}'\theta^{INDEX} + \eta_t,
\end{align*}

(37) \quad (38)

Interacting for \textit{NBER\_Recession} increases model fit for all groups relative to the specification without the interaction with up to a 13.6\% adjusted R-squared gain for Energy commodities. Commodities in the Grain and Softs groups shows a better fit under \textit{IndexPeriod} interactions. Metal commodities show no significant difference between the two specifications, while Energy commodities have less explanatory power under the interaction with \textit{IndexPeriod}. Next, I analyze the coefficients from rolling regressions in order to investigate the time-variation in commodity futures volatility.

5. Conclusion

This paper conducts a systematic analysis in order 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 is 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, while controlling for the impact of any frictions due to trading activity.

Using data for major commodity futures markets and global bilateral commodity trade, I analyze the extent to which commodity volatility is related to fundamentals that impact convenience yield and interest rates such as increased emerging market demand and inflation uncertainty, as well as financial frictions introduced by changing market structure and commodity index trading. Higher concentration in emerging market importers of a commodity is associated with higher futures volatility. I find significant predictability in commodity futures volatility using variables capturing macroeconomic uncertainty.

Such explanatory power can be economically significant for market participants (Campbell and Thompson 2008; Inoue and Kilian 2004). Investors who have a volatility-sensitivity (non-
Table 9: Commodity Futures Volatility during Different Time Periods

This table 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-squareds, Panel B shows the Bayesian information criterion values, and Panel C shows the likelihood ratio statistic of the different fits. Column 1 shows fit measures from the regression specification in Table VI.A (Baseline), column 2 shows those from VII.A (Macro uncertainty), column 3 shows those from VII.B (PCP4), and columns 4 shows those from VIII (Commodity market). Column 5 shows the results with all explanatory variables in specification (4) included, together with interactions for NBER recession periods; The last column shows the results with all explanatory variables in specification (4) included, together with interactions with IndexPeriod (the indicator for the period subsequent to January 2004). The results reported here are for the commodity groups Energy, Metal, Grain, Soft, and All (all 17 commodities in sample, see Table I). All regressions include commodity and season (month) fixed effects.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Macro uncertainty</th>
<th>PCP 1-4</th>
<th>Commodity market</th>
<th>Interaction for Recession periods</th>
<th>Interaction for Index period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Energy</td>
<td>34.3</td>
<td>48.9</td>
<td>44.2</td>
<td>44.9</td>
<td>58.5</td>
<td>51.1</td>
</tr>
<tr>
<td>Metal</td>
<td>43.7</td>
<td>47.3</td>
<td>47.4</td>
<td>48.3</td>
<td>49.2</td>
<td>49.3</td>
</tr>
<tr>
<td>Grain</td>
<td>41.9</td>
<td>47.8</td>
<td>45.9</td>
<td>46.0</td>
<td>48.6</td>
<td>49.1</td>
</tr>
<tr>
<td>Softs</td>
<td>25.2</td>
<td>26.3</td>
<td>26.3</td>
<td>26.4</td>
<td>26.2</td>
<td>26.7</td>
</tr>
<tr>
<td>All</td>
<td>39.5</td>
<td>42.4</td>
<td>42.3</td>
<td>42.8</td>
<td>43.9</td>
<td>43.1</td>
</tr>
</tbody>
</table>

Panel A: Adjusted R-squared

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

Panel B: Bayesian information criterion (BIC)

Panel C: Likelihood ratio comparison (LRT)

\[
K - K' \quad \text{K2- K1=10} \quad \text{K3- K1=4} \quad \text{K4- K2=3} \quad \text{K5- K4=10} \quad \text{K6- K4=10}
\]

Critical value

(P = 0.05) 18.307 9.488 7.815 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
zero “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 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.

This work builds on results discussed in [Bloom (2014)], that show emerging markets and recessionary periods are strongly associated with economic uncertainty, and adapts work that studies the granular origins of volatility [Gabaix (2011)] and shows how the same principle can affect volatility in global markets. It is difficult to contemporaneously explain let alone predict financial asset volatility using factors reflecting economic conditions [Roll (1984), Schwert (1989), Engle and Rangel (2008)], even when model results and economic intuition posit such a relationship. As such, the results in this paper are a step forward in our understanding of the factors that drive volatility. As global markets become increasingly interlinked, it is imperative to understand the impact of increased concentration and emerging market participation in commodity trade, and the manner and extent to which shocks propagate between markets.
Appendix A. Variance decomposition of the futures term structure

We can further decompose the excess basis return, \( x_{n,t+1} \), in Eq. (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:

\[
x_{n,t+1} \equiv x^y_{n,t+1} - x^r_{n,t+1},
\]

\[
x^y_{n,t+1} - E_t x^y_{n,t+1} = (E_{t+1} - E_t) \left\{ -\sum_{i=1}^{n-1} y_{1,t+i} - \sum_{i=1}^{n-1} x^y_{n-i,t+i+1} \right\},
\]

\[
x^r_{n,t+1} - E_t x^r_{n,t+1} = (E_{t+1} - E_t) \left\{ -\sum_{i=1}^{n-1} \pi_{1,t+i} - \sum_{i=1}^{n-1} \psi_{1,t+i} - \sum_{i=1}^{n-1} x^r_{n-i,t+i+1} \right\},
\]

where, \( \pi_{1,t} \) is the 1-period inflation rate and \( \psi_{1,t} \) is the 1-period real interest rate at time \( t \). The derivation of (41) 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 \( s_{n,t} \equiv y^*_n - y^*_1 \), where \( y^*_n = y_{n,t} - r_{n,t} \). Given \( p_{n,t} = n(r_{n,t} - y_{n,t}) \) and \( p_{n,t} = -E_t \sum_{i=0}^{n-1} \delta_{n-i,t+i+1} \),

\[
y^*_n = \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} \delta_{n-i,t+i+1}
= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + y^*_1). \tag{42}
\]

Given Eq. (42) and the definition of futures term spread \( s_{n,t} \), the unexpected excess return can be decomposed into the unanticipated change in \( y^*_1 \) and the unexpected change in futures term spread:

\[
x_{n,t+1} - E_t x_{n,t+1} = -(n-1)(y^*_{n-1,t+1} - E_t y^*_n) = -(n-1)(E_{t+1} - E_t) \left[ y^*_{1,t+1} + s_{n-1,t+1} \right]. \tag{43}
\]

Given the definition of \( y^*_1 \), it is straightforward to show that \( y^*_{1,t+1} - E_t y^*_1 = (E_{t+1} - E_t) [y_{1,t+1} - r_{1,t+1}] \). To derive the unexpected change in futures term spread, I start with Eq.
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,

$$s_{n,t} \equiv y_{n,t}^* - y_{1,t}^*$$

$$= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + y_{1,t+i}^* - ny_{1,t}^*)$$

$$= \left( \frac{1}{n} \right) \left( E_t \sum_{i=0}^{n-1} x_{n-i,t+i+1} \right) + E_t [-ny_{1,t}^* + y_{1,t}^* + (n-1)y_{1,t+1}^* - (n-2)y_{1,t+1}^*$$

$$+ \ldots + 3y_{1,t+n-3}^* - 2y_{1,t+n-2}^* + 2y_{1,t+n-2}^* - y_{1,t+n-2}^* + y_{1,t+n-1}]$$

$$= \left( \frac{1}{n} \right) E_t \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + (n - i - 1)\Delta y_{1,t+i+1}^*). \quad (44)$$

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

$$s_{n-1,t+1} - E_t s_{n-1,t+1} = \left( \frac{E_{t+1} - E_t}{n - 1} \right) \sum_{i=0}^{n-1} (x_{n-i,t+i+1} + (n - i - 1)\Delta y_{1,t+i+1}^*)$$

$$= \nu_t^{r+1} + \nu_t^y - \nu_t^{r+1}. \quad (45)$$

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.
Appendix B. Volatility models and extensions

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

B.1. GARCH-type models

Drawing from previous work on equity market volatility (Engle and Lee 1999; Engle and Gallo 2008; Engle and 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 (Engle 1982; Bollerslev 1986) for a particular asset is defined as,

\[ r_t = \mu_t + \sigma_t \epsilon_t, \quad (46) \]
\[ h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}, \quad (47) \]
\[ \sigma_t = \sqrt{h_t}. \quad (48) \]

It follows that the unconditional variance in the model will be 
\[ E[(r_t - E_{t-1} r_t)^2] = E[(r_t - \mu_t)^2] = \frac{\omega}{1-\alpha-\beta}. \]

In its simplest form, extensions to the standard GARCH(1,1) process that include (weakly) exogenous lagged explanatory variables in \( z_t \), with \( \xi_t = \frac{z_t}{E[z_t]} \), takes the form of GARCH-X(1,1),

\[ g_t = \omega + \alpha \epsilon_{t-1}^2 + \beta g_{t-1} + \xi_t' \theta, \]
\[ \sigma_t = \sqrt{g_t}. \quad (49) \]

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

Note that unlike with equity (Bekaert and Wu 2000; Bollerslev and 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\(^{21}\) may be useful in learning about the conditional demand-side or supply-side pressures in a commodity market. There is no asymmetric volatility

\(^{21}\)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 (47) by the specification of \( h_t \),

\[ h_t = \omega + \alpha \epsilon_{t-1}^2 + \sum_{i=1}^{K} \gamma_i I_{t-1}^{(-)} \epsilon_{t-1}^2 + \beta h_{t-1}, \quad (50) \]

where, \( I_{t-1}^{(-)} \) is 1 when \( \epsilon_{t-1} < 0, \) and 0 otherwise.
effect when controlling solely for the sign of lagged commodity futures return (see Table III and VIII.B).

B.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); Engle, Ghysels, and Sohn (2013), differing only in the definition of the slow-moving component of volatility.

\[
\begin{align*}
    r_t &= \mu_t + \sigma_t \varepsilon_t, \\
    h_t &= (1 - \alpha - \beta) + \alpha \frac{\varepsilon_{t-1}^2}{\tau_{t-1}} + \beta h_{t-1}, \\
    \sigma_t &= \sqrt{\tau_t h_t},
\end{align*}
\]

where \( \tau_t \) represents the long-term volatility component and for a set of \( K \) lagged explanatory variables in \( z_t \), is defined,

\[
\log \tau_t = m + z_t' \theta,
\]

The size of the set of estimated parameters in the model, \( \Theta = \{ \mu, \alpha, \beta, m, \gamma_1, \ldots, \gamma_K \} \), 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

\[
E\left[(r_t - E_{t-1}r_t)^2\right] = \tau_t E[h_t] = \tau_t.
\]

Engle, Ghysels, and Sohn (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 (52) 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 (52)), estimate \( \tau \) non-parametrically using an exponential quadratic
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, demarked by \(\{t_0 = 0, t_1, \ldots, t_k = T\}\). The estimated time series of the slow-moving component (\(\tau\)) is subsequently used as the dependent variable in an independent regression, with up to eleven 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 \(z\) that can capture the impact of macroeconomic uncertainty, supply-demand shocks, and trading activity.
Appendix C. Commodity Market Variation through Time

This section presents evidence of significant time-variation in volatility, correlations, and trends in emerging market commodities trade throughout the history of futures trading since the 1950s, which motivates the analysis presented in the rest of this paper. In the online Appendix I show plots of the daily returns and cumulative daily returns of the crude oil futures contract. There are extreme price movements over the entire term structure during the Global Financial Crisis, leading to sharp increases in the volatility exhibited during that period. From late 2001 to early 2007, there is a steady increase in cumulative return across all maturities. Figures 1 and 5 illustrate the difference in the estimated time-varying volatility depending on the term structure and holding period. The figures show the rolling 1-month and 12-month realized volatility for 1M, 3M, ..., 36M futures (3-day) returns for crude oil, natural gas, gold, copper, wheat, and lumber.


Figure 7 shows Chilean exports of copper for the period covered by the OECD STAN Bilateral Trade database. From the early 2000s, there is a sharp uptick in the percentage of exports to China, and the fraction of exports to G7 countries declines over the same period. Starting from near-zero, within less than two decades, the fraction of exports to China rises to 35.64% at the end of 2009, surpassing exports to all G7 countries combined.
Fig. 5. Time series of annualized rolling realized volatility at different horizons for copper, wheat, and lumber 1M, 3M, ..., 36M futures (3-day) returns. Here, short-term volatility refers to the standard deviation for the previous month, and long-term volatility refers to the standard deviation for the previous twelve months. The shaded areas highlight NBER recession periods. The dotted line marks January 2004.
Fig. 6. Time series of the rolling 12-month pairwise correlation between returns of the crude oil 3-month future and gold, copper, coffee, cotton, wheat, and lumber futures returns, respectively. Each date shows the corresponding correlation for the previous 12-month period calculated on 3-day returns. The shaded areas highlight NBER recession periods. The dotted line marks January 2004.
Fig. 7. Breakdown of importers of Basic Metals from Chile from 1990 to 2009. Chile exports close to a third of the world’s copper, a key raw material in manufacturing. The four series displayed add up to 100% of Chilean exports of Basic Metals in a given year.
Fig. 8. Global crude oil trade network. The vertex colors identifies the country group: BRIC (red), non-OECD excluding BRIC (yellow), OECD excluding G7 (green), and G7 (blue). The relative size of a country vertex captures its total import value.
References


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


50


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.
