

OTDK-dolgozat

2025

Budapesti Corvinus Egyetem

**FINANCIALIZATION ÉS BIZONYTALANSÁG AZ ÁRUPIACON: EGY
BAYESI MEGKÖZELÍTÉS**

**FINANCIALIZATION AND UNCERTAINTY IN THE COMMODITY
MARKET: A BAYESIAN APPROACH**

Szilárd Pálma Bernadett

Konzulens: Badics Milán Csaba

Kézirat lezárása: 2025. január 15.

Acknowledgement

I would like to express my gratitude to Milán Csaba Badics for being my research advisor in the last three years in the Empirical Finance Research Program. During this time, he helped me to learn analysing research papers, develop my critical thinking skills, write more effectively for academic audiences, and acquire a deeper knowledge of econometrics and programming. He always encouraged me to pursue my thoughts and interests in the field of empirical financial research. His support, guidance and encouragement contributed to the completion of this project.

Table of Contents

1	Introduction	1
2	Literature review	2
3	Methodology	8
3.1	Factor model with stochastic volatility and time-varying parameters	9
3.1.1	Priors and starting values	11
3.1.2	Estimation algorithms	13
3.1.3	Posteriors	15
3.1.4	Variance decomposition	19
3.2	Nonlinearity in the factor model	20
3.2.1	Priors and starting values	22
3.2.2	Posteriors	24
3.2.3	Generalized impulse response functions	25
4	Data	27
4.1	Commodity futures dataset	27
4.2	Financialization dataset	30
4.3	U.S. economy dataset	32
5	Empirical results	34
5.1	Financialization	34
5.2	Commodity uncertainty	39
6	Robustness tests	42
7	Conclusion	46
A	Appendix	55
A.1	Condition number of a matrix	55
A.2	Financialization indices in existing literature	56
A.3	Impulse-response functions to energy commodity uncertainty shock	61

List of Tables

1	Summary of the parameters prior distributions in model 1.	12
---	---	----

2	Summary of the parameters prior distributions in model 2.	22
3	Summary statistics on commodity futures.	29
4	Summary statistics on financialization dataset.	32
5	Description of the U.S. economy dataset.	33
6	Correlation of the financialization factors.	38
7	Robustness tests of the variance decomposition	44
8	Robustness tests of second model specification.	45
A1	Individual correlation between financialization time series.	60

List of Figures

1	Variance decomposition of commodity returns.	36
2	Financialization indices in the existing literature.	37
3	Impulse response functions of energy commodity futures.	40
4	Impulse response functions of macroeconomic price level time series.	41
5	Impulse response functions of macroeconomic activity and employment time series.	42
A1	R^2 decomposition of commodity returns.	57
A2	Co-movement between equity and commodity futures.	58
A3	Index Investment Data.	59
A4	Impulse response functions of energy commodity futures and financial variables.	61
A5	Impulse response functions of financial and macroeconomic variables.	62
A6	Impulse response functions of macroeconomic variables and commodity futures.	63

1 Introduction

Commodity markets have a pivotal role in the economy. Energy and raw material costs affect global economic activity, and consumer prices through transportation, heating and production costs among others. Precious metals other than being used in industrial production offer a safe haven in times of financial crisis (Ji et al., 2020). Furthermore, agricultural commodities provide a large share of foreign currency income through exports for developing countries.

I analyse two pivotal phenomena in the commodity market: financialization and commodity uncertainty. The naming financialization describes the integratedness of commodities in the financial markets.

Although the commodity spot market is significant, commodities are mostly traded on the derivatives market. Traditionally, the commodity derivatives market provided the opportunity for commercial investors to hedge their exposure in the spot market. Since the 2000s investment in commodity futures increased rapidly, and between 1998 and 2008, the total notional value of the OTC contracts on commodities increased 28 times (Domanski and Heath, 2007). Existing literature attributed the large influx of investment to index investors who sought to diversify their exposure to financial risks by investing in the commodity market. As a result, connectedness between the financial markets and commodity futures increased through the trading activity and portfolio re-balancing of index investors (Büyüksahin and Robe, 2014).

Financialization is often defined as a fundamental change in commodity futures markets inasmuch as commodities began to be more integrated with financial markets (Silvennoinen and Thorp, 2013), resemble a positive correlation with equities (Nguyen et al., 2020), and volatility spillovers originating outside the commodity market became more frequent (Adams and Glück, 2015). I categorize the definitions and methodologies scholars applied into three groups. The first approach identifies financialization as the increased correlation between commodity futures and equities (Nguyen et al., 2020; Ordu et al., 2018), the second estimation is based on the net long position of a subgroup of commodity investors published by the Commodity Futures Trading Commission (Irwin and Sanders, 2012; Tang and Xiong, 2012), and the last methodology is measuring the importance of financial variables in explaining commodity returns variation (Adams et al., 2020).

I unite these three definitions and measure financialization with a Bayesian factor model. The methodology is most closely related to the R^2 decomposition of Adams et al. (2020). However, as I measure financialization by a common factor of commodity futures and financial

variables it also relates to the group of scholars that define financialization as the increased correlation between the two markets. Furthermore, by applying OTC derivatives contract data my results also relate to the second group of financialization estimation.

Second, I calculate three components of commodity uncertainties: energy, precious metal and commodity real activity that include agricultural and industrial metal futures. I estimate their impact on the U.S. economy with a Bayesian factor model similar to the one applied in the first step to compute uncertainties. Estimating uncertainty indices and their impact in two different models has several disadvantages. First, it suffers from inconsistency as the uncertainty measure is usually treated as unobservable in the first model but present in the main equation and observable in the second model (Creal and Wu, 2017). Furthermore, the two models usually employ a different subset of variables; therefore, possibly suffer from omitted variable bias (Carriero et al., 2018).

Compared to studies that estimate uncertainty and its impact in a single model I introduce a factor level that reduces the number of variables and enables the estimation of a small VAR model. Similar to Carriero et al. (2018) and Mumtaz and Theodoridis (2020) I set the uncertainty to enter the mean equation and allow the interaction of uncertainty measures within a VAR model. To the best of my knowledge uncertainty has not been estimated in a model that includes these three features. Furthermore, as the existing literature on the impact of commodity uncertainty is scarce, I contribute to it with empirical analysis.

To summarize, my contribution to the existing literature is twofold. First, I propose a measure of financialization that incorporates the definitions applied by existing literature and unifies them. Second, I compute energy, precious metal and commodity real activity uncertainty and estimate their impact on the U.S. macroeconomic, financial and monetary policy time series. To the best of my knowledge, no study has estimated commodity, and economic uncertainties jointly and analysed their impact on the economy in a one-step estimation.

2 Literature review

Existing literature attributes financialization to the increased presence of index traders in the commodity futures market (Büyüksahin and Robe, 2014; Ma et al., 2021; Tang and Xiong, 2012). Based on the Commodity Futures Trading Commission staff report the index-related instruments held by institutional investors increased from \$15 billion in 2003 to \$200 billion in

2008¹ when co-movement between commodity and stock market and the impact of financialization peaked (Ma et al., 2021).

Traditionally, in the commodity market commercial investors sought to hedge their exposure to production or consumption. Earning a risk premium, specialised investors provided this insurance, usually by offsetting their positions in the spot market (Domanski and Heath, 2007). Based on the theory of normal backwardation, investors provide a futures price below the expected spot price prevailing at maturity, therefore earning a risk premium (Keynes, 1930).

In an efficient market, the commodity futures prices should be linked to the demand for production, the capacity of suppliers, and the marginal cost of production. These fundamentals indicate that the commodity market was segmented from the financial markets and partially from each other. Academic papers began to emphasise the negative correlation between commodity and stock returns² as well as the fact that historically returns on commodities and equities with similar risk features are comparable. Moreover, in the early recession and late expansion periods, when bond and stock returns are below their average, commodity returns can be positive (Gorton and Rouwenhorst, 2006). As a result, the passive investment strategy in a basket of commodities providing diversification benefits attracted long-only institutional investors, and commodity index traders between 2000 and 2010 more than quadrupled (Cheng et al., 2015), leading to the financialization of the commodity futures market.

In the existing literature, there are different explanations for financialization. Ma et al. (2021) analysed co-movements between 21 commodity futures, concluding that in the pre-crisis period between 2002 and 2008, positive co-movement was driven by the growing demand of emerging markets, mainly BRIC countries. Nevertheless, Tang and Xiong (2012) found only a moderate correlation between the S&P GSCI and the MSCI Emerging Markets Index, concluding that it cannot be the main driving force behind financialization. Similarly, the authors tested for the global financial crisis, inflation, and the adoption of biofuel, concluding that none of them explain the phenomenon. Only the weekly investment flow from index traders showed a positive correlation with commodity futures, but this does not confirm causality³.

¹The report is available at <https://www.cftc.gov/PressRoom/PressReleases/5542-08>.

²Historically, the negative correlation only holds for horizons longer than a month. Furthermore, commodity returns are positively correlated with inflation, the relation increasing over a longer horizon, from 1 to 5 years (Gorton and Rouwenhorst, 2006).

³Tang and Xiong (2012) argue that the Granger causality test adopted in the literature of financialization only discovers the lead-and-lag relationship between time series. Although index investment flows Granger cause commodity returns, the results might be driven by reverse causality. For instance, index traders could predict commodity returns. Adams et al. (2020) also points out the endogeneity between the flow of investment and commodity prices, as positive prices attract investment and the buying pressure of financial investors leads to

Ordu et al. (2018) analysing the correlation of agricultural commodities and equities concluded that commodity index traders' (CITs) explanatory power although significant, but dependent on the liquidity of a given commodity market. Alquist et al. (2020) propose that co-movements can be explained by a common factor affecting macroeconomic fundamentals ⁴. Ordu et al. (2018) confirm that this factor could be business cycle fluctuations that explain positive co-movements between stock and commodity market; however, it does not change the significance of CITs.

Similar to the cause of financialization, there is no consent in the existing literature on the definition and estimation of it. I intend to present scholar's works and findings by grouping them into three categories based on their methodology and analysed dataset.

Several studies analysed trading activities, in particular, index investment following the testimony of Masters (2008) in the US Senate, stating that the sharp increase of commodity prices particularly oil futures in 2008 was caused by index investment. To form his hypothesis, he analysed the Supplemental Commitment of Traders (SCOT) data published by the Commodity Futures Trading Commission ⁵. Early papers following the statement found little evidence to support the hypothesis (Brunetti and Buyuksahin, 2009; Stoll and Whaley, 2010). For instance, Irwin and Sanders (2012) analysed the relation of futures returns and volatility with the CFTC's Index Investment Data (IID), SCOT and Disaggregated Commitment of Traders by regression and Granger causality and found no support for Master's hypothesis. On the other hand, conditional on the VIX, Cheng and Xiong (2014) concluded that CFTC commodity index trading data correlates positively and significantly with commodity futures prices. The authors argued that as index traders sell their positions in a crisis, conditioning on the VIX speculators and hedgers' trading activity can be distinguished. Furthermore, Chen et al. (2021) concluded that commodity trading advisors' activity predicts the co-movement between the stock- and commodity market.

In this research, I do not intend to take sides on the matter. Instead, I rely on the studies to use the CFTC data to proxy financialization. Applying CFTC data to estimate the impact of financialization was applied by Adams and Glück (2015); Tang and Xiong (2012) and Ordu et al. (2018) among others. As detailed by Irwin and Sanders (2012) the Index Investment Database captures the best non-commercial, index investment activity; henceforth, I use that database in my research to compare the results.

By other scholars, financialization was defined as the heightened co-movement between increased commodity prices.

⁴It is worth noting the findings of Adams et al. (2020) indicate that in the period of financialization, economic fundamentals no longer explain the total variation in commodity prices.

⁵Henceforth CFTC.

equities and commodities and the connectedness among commodities (Balli et al., 2019; Ma et al., 2021; Nguyen et al., 2020; Tang and Xiong, 2012). Historically, Gorton and Rouwenhorst (2006) demonstrated that the correlation between the S&P500 and commodity returns was minor. Computing an ARMA-filter-based correlation between commodities and the S&P500 Nguyen et al. (2020) concluded that a positive correlation between equities and commodities became largely significant especially, from 2000 ⁶. By the same token, Ordu et al. (2018) confirmed an increased correlation between equities and commodities since 2008.

In the analysis, I adopt their approaches and calculate a rolling window correlation of commodity futures with the S&P500 to compare the estimated index of financialization to their definition.

The last approach to quantify financialization was described by Adams et al. (2020). The authors argued that if commodities were no longer segmented from stock markets, variables explaining movements in the financial markets should explain variations in commodity futures return and volatility. The study applied LMG decomposition of rolling window linear regressions and concluded that the influence of financial variables greatly increased since the global financial crisis.

The model I develop to estimate financialization is closely related to the decomposition of Adams et al. (2020), but methodologically more complex and therefore overcomes the limitations of the regression and the LMG R^2 decomposition ⁷. I employ a dynamic factor model with a Bayesian estimation that to the best of my knowledge has never been applied to estimate financialization. The model allows in a one-step estimation to decompose the variation of commodity returns and include a large set of financial and economic variables. The flexible specification of factor models enables calculating the dynamic impact of financialization without selecting an arbitrary window size to roll over the data. Rolling window estimations, applied by two out of the three approaches to measure financialization, suffer from the drawback that the window contains a large share of past observations that might camouflage sudden jumps in the data.

Dynamic latent factor models were proposed by Otrok and Whiteman (1998) to analyse the economic data of Iowa. Kose et al. (2003) further developed the model by adding region- and country-specific factors to study the common macroeconomic components of 60 counties. With

⁶The only exception is the gold futures that correlated negatively with equities.

⁷In a linear regression if the explanatory variables are correlated, LMG R^2 decomposition allocates a minimum weight to variables with 0 coefficient value (Grömping, 2007).

this improvement, one can estimate pure global factors and separate region-specific variations that in this case are commodity classes, such as agricultural or energy futures. Del Negro and Otrok (2008) developed an estimation of time-varying factor loading so that countries will have different exposure to the global and regional factors in the sample period. In the models, it is crucial to allow time-varying factor loadings as I expect larger exposure of the commodities to the global financial factor during the period of financialization and moderate exposure before.

Furthermore, I address the problem of order-dependent covariance matrices in the vector autoregressive model that describes the estimation of the interaction between financial and economic factors. With the growing application of large VAR models (Bańbura et al., 2010; Carriero et al., 2018), several articles were published to address or solve the order-dependent covariance matrix problem in Bayesian estimations. A covariance matrix is order-invariant if changes in the order of the variables leave unaffected the posterior distribution of the covariance matrix (Chan et al., 2023). I follow Chen et al. (2023) and allow the scales of stochastic volatilities to be a full matrix instead of the lower-triangular specification of Cogley and Sargent (2005) and Primiceri (2005). Henceforth, my models are order-invariant.

My research also relates to the fast-growing literature on measuring uncertainty and its impact on the economy. Recently, scholars began to measure specific components of uncertainty, such as the economic policy uncertainty (Baker et al., 2016), the consumer's perceived uncertainty (Leduc and Liu, 2016) or macroeconomic uncertainty (Jurado et al., 2015).

Based on the applied methodology uncertainty measures can be grouped into four categories (Cascaldi-Garcia et al., 2023). News-based measures usually arrange a dictionary related to the uncertainty component and count the number of words in a group of newspapers for a given frequency. For instance, the geopolitical risk index of Caldara and Iacoviello (2022) measures geopolitical act and threat uncertainties by spotting words such as war, embargo, nuclear weapon etc. More sophisticated dictionaries allow the combination of nouns and adjectives, relating more to sentiment analysis in order to determine the sign of uncertainty. For instance, to analyse monetary policy statements Apel and Grimaldi (2014) categorized adjective-noun combinations to hawkish and dovish monetary policy signals and develop an index as a weighted sum of the two signs in a given policy statement.

The second methodology relies on surveys conducted with consumers, professionals and other market participants. To summarize the results into a single time series Leduc and Liu (2016) averaged the responses while Rossi and Sekhposyan (2015) computed the U.S. GDP

professional forecasters uncertainty index from the tail distribution of the answers. One advantage of this methodology is measuring uncertainty directly from a specified market segment compared to the news-based approach where determining the specific newspapers might be provisional. Its disadvantage comes from the quality of the responses.

The market-based methodology estimates uncertainty from the volatility of observable market data. A prominent example is the CBOE VIX derived from the implied volatility of the options on the S&P500. A disadvantage of these measures is reflecting solely a small set of investors' opinions trading in the options market of a specific asset.

The last methodology is the econometric-based approach that usually defines uncertainty as an unforecastable component of a large dataset (Jurado et al., 2015) or identifies it with the stochastic volatility in a large VAR model (Carriero et al., 2018). The advantage of this approach is that the horizon to which the uncertainty measure relates is well-defined (Cascaldi-Garcia et al., 2023) and as opposed to the market-based methodology a large set of data can be applied to measure a single time series. For instance, (Jurado et al., 2015) uses more than 200 economic and firm-level data to measure macroeconomic uncertainty as the average volatility of forecast error of each time series.

I also estimate commodity uncertainty with an econometric-based methodology given that it is the only approach that enables the calculation of uncertainty and its impact within the same model.

Measuring uncertainty and its effect in separate models brings several disadvantages and inconsistencies between the models. Reviewing the impact of uncertainty measures, Creal and Wu (2017) described an important drawback of the two-step uncertainty measures: in the estimation process uncertainty is treated as a non-observable variable and in a second model ⁸ it is included as an observable time series leading to a contradictory overall model specification. Carriero et al. (2018) also argues that the second step can suffer from omitted variable bias given that uncertainty measures are usually obtained from a large data frame and the analysis of the second step is conducted with a small subset of that data. The authors suggest measuring the impact of uncertainty in a VAR model and quantifying its impact in the same model in order to overcome the limitations outlined above. In the second model, I follow Mumtaz and Zanetti (2013) and Carriero et al. (2018) and estimate the uncertainties in a VAR model with stochastic volatilities.

⁸Usually in a small, homoscedastic VAR model.

3 Methodology

In the research, I measure financialization and estimate commodity uncertainty time series. I choose a research framework that offers a parsimonious and flexible way to carry out the estimations in separate steps ⁹. I chose to estimate a dynamic factor model for two purposes. First, factor models offer an efficient way to estimate and analyse a large dataset by dimension reduction, this way I broaden the scope of the analysis. Second, the specification of the factor dynamics offers the flexibility to enable interaction within the factors and between the factors and uncertainties.

Factor models are mostly estimated with the Bayesian approach given that as the number of parameters increases the likelihood estimation becomes cumbersome.

In statistics, there are two approaches to estimate models: frequentist and Bayesian. In the frequentist approach, I treat the input data as stochastic and estimate “fixed” parameters meanwhile, in the Bayesian approach parameters are stochastic. Bayesian estimation therefore enables the incorporation of prior belief of the data at hand and also provides information on “how confident we are on the parameters realizations”. The estimation procedure relies on the Bayes’ theorem. Given our prior belief before seeing the data, the prior density $p(\theta)$, and the likelihood function, $p(y | \theta)$ the posterior distribution by the Bayes’ theorem is:

$$p(\theta | y) = \frac{p(y | \theta) p(\theta)}{\int p(y | \theta) p(\theta) d\theta} \propto p(y | \theta) p(\theta) \quad (1)$$

meaning that the posterior distribution is proportional to the likelihood times the prior density.

Bayesian estimations sometimes come under criticism concerning the prior distribution specifications. To address this concern several priors are estimated from a training sample that is discarded from the sample I use to estimate the posteriors. All other prior distributions are specified based on previous studies. Furthermore, I conduct robustness tests that involve modification of the prior distribution parameters.

In the following two subsections, I present the dynamic linear factor model and its modification that enables nonlinearity. In the first part, I introduce a time-varying parameter factor model with stochastic volatility. In the second subsection, I only describe the modifications of that model to estimate commodity, macroeconomic and financial uncertainty. Both subsections provide a detailed description of the sampling algorithm and link the model components to the empirical

⁹It might be possible to use a one-step estimation; however, the two research questions employ different datasets.

findings of existing literature.

3.1 Factor model with stochastic volatility and time-varying parameters

To estimate a financialization index by variance decomposition of a commodity-financial and commodity-macro factor I employ a dynamic factor model with time-varying factor loadings, stochastic volatility and VAR dynamics of the global factors. In the model specifications and prior selections I rely on the studies of Blake and Mumtaz (2012); Mumtaz and Theodoridis (2020); Mumtaz and Zanetti (2013), Mumtaz and Musso (2021) and Chan et al. (2023) ¹⁰.

Let Y_{it} denote a $T \times n$ matrix of time series that consists of $n = \{k, m, f\}$ number of commodity, macroeconomic and financial variables. The observation equation of the factor model is defined as:

$$Y_{it} = B_{i,t}^G F_t^G + B_{i,t}^R F_t^R + v_{i,t} \quad (2)$$

where F_t^G denotes the global factors that are different from existing literature. Usually, one global factor is extracted from the whole dataset and additional regional factors capture the variation of the subset of the data frame ¹¹ (Le Pen and Sévi, 2018; Mumtaz and Musso, 2021). I estimate two global factors with a different subset of the dataset. On the first global factor commodity and macroeconomic data have loadings, on the second commodity and financial data loads. This way I create a parsimonious one-step estimation to measure financialization of all the commodity data defined by the variance decomposition. The region-specific factors are denoted by F_t^R and their definition is consistent with the literature on factor models.

As described in Del Negro and Otrok (2008) I allow time-varying factor loadings that are crucial to estimate financialization as the exposure of each commodity future to the macroeconomic but more importantly to the financial factor. The time-varying parameters ($B_{i,t} = [B_{i,t}^G; B_{i,t}^R]$) follow a random walk:

$$B_{i,t} = B_{i,t-1} + (Q_i^B)^{1/2} \epsilon_t^B, \quad \epsilon_t^B \sim \mathcal{N}(0, 1) \quad (3)$$

Noting at this point that the global, regional and country indices G, R, C denote the number of global, regional and idiosyncratic factors.

¹⁰The replication code of the model is available at: Github

¹¹From a frequentist point of view a global factor can be interpreted as the first principal component of all the variables and a region-specific factor as the first principal component of the regional data after extracting the global factor from the dataset. For instance, in the case of commodities, a regional-specific factor is estimated from energy futures.

The global factors follow a time-varying parameter VAR process:

$$F_t^G = \alpha_t^G + \sum_{i=1}^P \beta_{i,t}^G F_{t-i}^G + \Omega_t^{1/2} \epsilon_t, \quad \epsilon_t^G \sim \mathcal{N}(0, I_T) \quad (4)$$

with $P = 2$ following Mumtaz and Musso (2021). $\beta_{i,t}^G, \Omega_t$ denote the VAR coefficients and covariance matrix respectively.

Following Cogley and Sargent (2005) the covariance matrix of the residual in the VAR model is defined as:

$$\Omega_t = A^{-1} H_t A^{-1^T} \quad (5)$$

The most widely used specifications described in Primiceri (2005) and Cogley and Sargent (2005) define A as a lower triangular matrix. This would lead to an order-dependent covariance matrix as noted by the authors¹². Henceforth, as discussed in the literature review I set A as an unrestricted square matrix following Chan et al. (2023) to specify an order-invariant covariance matrix.

The matrix H_t is diagonal with elements:

$$H_t = \text{diag}(h_t^1, h_t^2, \dots, h_t^G) \quad (6)$$

with G denoting the number of global factors, in this case $G = 2$. The VAR coefficients, ($\bar{\beta}_t^G = \{\alpha_t^G, \beta_{1,t}^G, \dots, \beta_{P,t}^G\}$) follow random walk as described in Blake and Mumtaz (2012):

$$\bar{\beta}_t^G = \bar{\beta}_{t-1}^G + Q_{\beta^G}^{1/2} e_t^G \quad e_t^G \sim \mathcal{N}(0, 1) \quad (7)$$

and the covariance Q_{β^G} has an inverse-Wishart distribution and is a full matrix.

The regional-specific factors follow an AR process:

$$F_t^R = \alpha_t^R + \sum_{i=1}^q \beta_{i,t}^R F_{t-i}^R + (h_t^R)^{1/2} \epsilon_t^R \quad \epsilon_t^R \sim \mathcal{N}(0, 1) \quad (8)$$

with $q = 2$ following Blake and Mumtaz (2012). Similarly to the VAR model of the global

¹²A covariance matrix is order-dependent if the order of the variables in F_t^G affects the posterior distribution or the covariance matrix Ω_t .

factor, the parameters $(\bar{\beta}_t^R = \{\alpha_t^R, \beta_{1,t}^R, \dots, \beta_{q,t}^R\})$ follow a random walk:

$$\bar{\beta}_t^R = \bar{\beta}_{t-1}^R + Q_{\beta^R}^{1/2} e_t^R \quad e_t^R \sim \mathcal{N}(0, 1) \quad (9)$$

The error term in the factor observation equation 2 represents country (idiosyncratic) factors. These factors follow an AR(1) process as in (Blake and Mumtaz, 2012):

$$v_{i,t} = \beta_{i,t}^C v_{i,t-1} + (h_t^C)^{1/2} \epsilon_{i,t}^C \quad \epsilon_{i,t}^C \sim \mathcal{N}(0, 1) \quad (10)$$

Similarly to the global and regional factors the parameters are time-varying and their dynamics are defined by a random walk:

$$\beta_t^C = \beta_{t-1}^C + Q_{\beta^C}^{1/2} e_t^C \quad e_t^C \sim \mathcal{N}(0, 1) \quad (11)$$

The stochastic volatilities of the global, regional and idiosyncratic factors follow a random walk as described in Blake and Mumtaz (2012):

$$\ln h_t^J = \ln h_{t-1}^J + (g^J)^{1/2} e_t^J \quad e_t^J \sim \mathcal{N}(0, 1) \quad (12)$$

with $J = \{G, R, C\}$ denoting the global, regional and idiosyncratic factors respectively.

3.1.1 Priors and starting values

In Bayesian estimation, parameters are treated as random variables, and their prior distribution must be specified. The starting values are obtained using the first $S_0 = 40$ observations as a training sample that I discard in the posterior estimation of the parameters.

The index $J = \{G, R, C\}$ stands for the global, regional and idiosyncratic factor indices altogether. Table 1 summarizes the priors of the parameters.

I set priors based on the following estimations.

Parameter	Equation(s)	Prior distribution	Starting value
F_t^J	2,4, 8, 10	-	Principal components
$B_{i,t}$	2,3	$\mathcal{N}(B_{i,0}, V_{B_i})$	$B_{i,0}$, PCA loadings V_{B_i} , PCA residual covariance
Q^B	3	$\mathcal{IW}(Q_0^B, V_{B_i})$	$Q_0^B = V_{B_i} \times S_0 \times 3.5 \times 10^{-4}$
Q^J	7, 9, 11	$\mathcal{IW}(Q_0^{\beta^J}, V_{\beta^J})$	$Q_0^{\beta^J} = V_{\beta^J} \times S_0 \times 3.5 \times 10^{-4}$ V_{β^J} , coefficients covariance matrix
$\bar{\beta}_t^J$	7,9, 11	$\mathcal{N}(\beta_0, \Sigma_0)$	$\beta_0 = (X_t' X_t)^{-1} (X_t' Y_t)$, OLS estimate $\Sigma_0 = \frac{(Y_t - \beta_t X_t)^{-1} (Y_t - \beta_t X_t)'}{S_0 - n}$ n = num. of variables in X_t Starting values: β_0
A	5	$a_i \sim \mathcal{N}(0, I)$	$I \in R^{2 \times 2}$
h_t^J	12	-	$(F_t^J - \alpha_t^J - \sum_{i=1}^{P_J} \beta_{i,t}^J F_{t-i}^J)^2 + 0.001$
g^J	12	$\mathcal{IG}(0.01, 1)$	

Table 1: Summary of the parameters prior distributions in model 1.

Note. The table summarizes the prior distributions and starting values of the model in this section. The parameter $S_0 = 40$ denotes the training sample size, and $J = \{G, R, C\}$ the global, regional and idiosyncratic factors.

- (1) Factors, loadings and their covariance matrices, $F_t^J, B_{i,t}, Q^B, Q_{\beta^G}, Q_{\beta^R}, q_{\beta^C}$:

The starting values of all the factors are the first principal components of the full sample
13.

I assume that the factor loadings are normally distributed: $B_{i,t}^G, B_{i,t}^R \sim \mathcal{N}(B_{i,0}, V_{B_i})$. The prior mean, $B_{i,0}$ equals the loadings of the principal components and the prior variance is set equals the covariance of the residual that remains after the principal components are extracted from the training sample.

The covariance matrices of the time-varying parameters random walk equation, 3, 7, 9, 11 have an inverse-Wishart prior distribution: $Q_i \sim \mathcal{IW}(Q_{i,0}, S_0)$. The $Q_{i,0}$ prior has a pivotal role as larger values of the scale matrix would imply larger fluctuations of the $B_{i,t}, \beta^J$ parameters. Following Mumtaz and Musso (2021) I set $Q_{i,0}$ based on the training sample data ¹⁴: $Q_0^B = V_{B_i} \times S_0 \times 3.5 \times 10^{-4}$. Parameters of the other inverse-Wishart distributions are set similarly.

- (2) VAR coefficients, $\bar{\beta}_t^G$:

¹³It is necessary to use the full sample to obtain a vector and loadings with dimensions of the full sample period. Estimating the prior factor and loadings is standard in the existing literature (Blake and Mumtaz, 2012; Mumtaz and Musso, 2021)

¹⁴In section 6 I conduct robustness test to the scaling factor $\times 3.5 \times 10^{-4}$.

A standard VAR model is estimated on the training sample with OLS. The starting values of VAR coefficients equals this estimate: $\beta_0 = (X_t' X_t)^{-1} (X_t' Y_t)$.

(3) AR coefficients, $\bar{\beta}_t^R, \bar{\beta}_t^C$:

On the training sample, an OLS regression is estimated. The initial value of the AR coefficients: $\beta_0 = (X_t' X_t)^{-1} (X_t' y_t)$. The prior of the coefficients are therefore normally distributed: $\beta_t^{R,C} \sim \mathcal{N}(\beta_0^{R,C}, \sigma_0^{R,C})$ with $\beta_0^{R,C}, \sigma_0^{R,C}$ denoting the coefficient and the covariance matrix OLS estimates on the training sample.

(4) Scale of the VAR model covariance matrix, A :

Following Chan et al. (2023) I use independent priors for each row of A denoted by a_i , $i = 1, 2$ assuming normal distribution: $a_i \sim \mathcal{N}(a_0, V_{a_0})$. Setting $a_0 = 1$ and $V_{a_0} = 1 \in R^{2 \times 2}$.

(5) Stochastic volatilities, $h_t^J, J = \{G, R, C\}$:

In the training sample I estimate the residuals of the VAR and the AR models for the global, regional and country factors as specified by equation 4, 8, 10 respectively: $u_t = F_t^J - \alpha_t^J - \sum_{i=1}^P \beta_{i,t}^J F_{t-i}^J$. The starting values of the stochastic volatilities are set to $h_{0,t}^J = u_t^2 + 0.001$ as described in Blake and Mumtaz (2012).

(6) Covariance of stochastic volatilities, g^G, g^R, g^C :

Following Blake and Mumtaz (2012) I assume inverse-gamma priors for these variances: $g^J \sim \mathcal{IG}(g_0, V_g)$, setting $g_0 = 0.01$ and $V_g = 1$.

3.1.2 Estimation algorithms

To sample from the posterior distribution of the parameters I use Markov Chain Monte Carlo method, namely the Gibbs sampler. This algorithm offers an easy solution when the joint posterior distribution is not known explicitly or has such large dimensions that it would be cumbersome to draw from it directly, but the conditional distributions of the parameters are simply estimated. Denoting with θ all the parameters, y the input data, x_i the i th variable ($i \in [1, \dots, n]$) and θ_{-i} all variables except for the i th, the algorithm has the following steps:

1. Start by setting initial values for each variable as outlined above, denoted by θ^0 .
2. Draw from $p(x_1 | \theta_{-1}^0, y)$, given that the conditional distribution is known and we can evaluate it with the starting points: $p(x_1 | \theta_{-1}^0, y) = p(x_1 | \theta_{-1} = \hat{\theta}_{-1}^0, y)$. The obtained

sample is denoted by x_1^1 .

3. Draw from $p(x_2 | \theta_{-1,2}^0, x_1^1, y)$. The obtained sample is denoted by x_2^1 .
4. ...
5. Draw from $p(x_n | x_1^1, x_2^1 \dots x_{n-1}^1, y)$.
6. Repeat steps 1. – 5. and after discarding a “burn-in” sample save the draws.

The obtained samples are correlated as we draw them one after the other and they generate a Markov chain (Smith and Roberts, 1993). To have independent samples researchers only keep every 5th (Carriero et al., 2018), 10th (Mumtaz and Musso, 2021) or 50th draw (Moench et al., 2013). In the estimation, I choose a burn-in sample of 10 000 draws and keep every 10th draw of the following 10 000 draws.

In some steps of the Gibbs sampling I transform the equations to a linear state-space model and draw the variable of interest with the Carter and Kohn (1994) algorithm. In the description of the algorithm, I rely on Blake and Mumtaz (2012); Kim and Nelson (1999). Consider the following state-space model:

$$y_t = A x_t + H \beta_t + e_t \quad e_t \sim \mathcal{N}(0, R) \quad (13)$$

$$\beta_t = \mu + F \beta_{t-1} + v_t \quad v_t \sim \mathcal{N}(0, Q) \quad (14)$$

we want to estimate the state variable $\hat{\beta}_T = \{\beta_1, \dots, \beta_T\}$ given the values of the other parameters $A, H, R, \mu, F, Q, e_t, v_t$ ¹⁵. I use the notation for all the variables

$\theta = \{y_1, \dots, y_T, A, H, R, \mu, F, Q, e_t, v_t\}$. To obtain the conditional posterior distribution Carter and Kohn (1994) proposed to use the following factorization:

$$p(\hat{\beta}_T | \theta) = p(\beta_T | \theta) \prod_{t=1}^{T-1} p(\beta_t | \beta_{t+1}, \theta) \quad (15)$$

Using that the state variables follow a first-order AR(1) process and their distribution only depends on the following β_{t+1} term (Kim and Nelson, 1999).

Knowing that the error terms in the state-space model are Gaussian the distribution of the first term becomes: $p(\beta_T | \theta) \sim \mathcal{N}(\beta_{T|T}, P_{T|T})$ and the distributions in the product are also normally distributed: $p(\beta_t | \beta_{t+1}, \theta) \sim \mathcal{N}(\hat{\beta}_{t|t}, \hat{P}_{t|t})$. To estimate the parameters of the

¹⁵We are able to sample these parameters conditional on the others with a regression.

first distribution $\beta_{T|T}, P_{T|T}$ we run the Kalman filter and then draw β_T from the corresponding distribution, denoted by $\bar{\beta}_T$. Followingly, as proposed by Carter and Kohn (1994) we iterate backwards in steps $T - 1, T - 2, \dots, 1$ and with the $\beta_{t|t}, P_{t|t}$ estimates and $\bar{\beta}_T$ draws obtained in the Kalman filter and we calculate:

$$\hat{\beta}_{t|t} = \beta_{t|t} + \underbrace{(P_{t|t} F' (F P_{t|t} F' + Q)^{-1})}_{\text{gain matrix}} \underbrace{(\bar{\beta}_T - \mu + F \beta_{t|t})}_{\text{state equation}} \quad (16)$$

$$\hat{P}_{t|t} = P_{t|t} - \underbrace{(F P_{t|t} F' + Q)^{-1}}_{\text{gain matrix}} F P_{t|t} \quad (17)$$

and draw β_t from the normal distribution: $p(\beta_t | \beta_{t+1}, \theta) \sim \mathcal{N}(\hat{\beta}_{t|t}, \hat{P}_{t|t})$.

Lastly, I briefly review an algorithm that can be used when sampling directly even from the conditional distribution is not possible. The Metropolis-Hastings algorithm offers the solution when we can evaluate the density function of the conditional distribution for a given, θ^* value: $p(\theta = \theta^*)$.

To draw from the target density, $p(\theta)$, we have to specify a candidate density, proportional to p . In each step I first draw from the candidate density, then calculate an acceptance probability. If that probability is larger than a threshold value I keep the sample in the chain, otherwise we draw again from the candidate density. As I use the independence Metropolis-Hastings algorithm, the steps, the candidate density and the acceptance probability are described in the posterior draws of the stochastic volatilities.

3.1.3 Posteriors

I repeat the following steps in the Gibbs algorithm to draw from the posterior distributions of the variables ¹⁶

(1) Factors: $F_t^J | B_{i,t}, \alpha_t^J, \beta_{i,t}^J, h_{i,t}^J$,

The Kalman filter and the Carter and Kohn (1994) algorithm is used to draw the factors.

To illustrate the observation and state equations consider an example with 1 global factor,

¹⁶In this description $J = \{G, R, C\}$ denote the global, regional and country-specific indices and i the number of factors.

2 regional factors and 4 data points. The observation equation:

$$\begin{bmatrix} Y_{1,t} - \beta_{1,t}^C Y_{2,t-1} \\ Y_{2,t} - \beta_{2,t}^C Y_{3,t-1} \\ Y_{3,t} - \beta_{3,t}^C Y_{4,t-1} \\ Y_{4,t} - \beta_{4,t}^C Y_{5,t-1} \end{bmatrix} = \begin{bmatrix} B_{1,t}^G & B_{1,t}^{R,1} & 0 & -\beta_{1,t}^C B_{1,t}^G & -\beta_{1,t}^C B_{1,t}^{R,1} & 0 \\ B_{2,t}^G & B_{2,t}^{R,1} & 0 & -\beta_{2,t}^C B_{2,t}^G & -\beta_{2,t}^C B_{2,t}^{R,1} & 0 \\ B_{3,t}^G & 0 & B_{3,t}^{R,2} & -\beta_{3,t}^C B_{3,t}^G & 0 & -\beta_{3,t}^C B_{3,t}^{R,2} \\ B_{4,t}^G & 0 & B_{4,t}^{R,2} & -\beta_{4,t}^C B_{4,t}^G & 0 & -\beta_{4,t}^C B_{4,t}^{R,2} \end{bmatrix} \begin{bmatrix} F_t^G \\ F_t^{R,1} \\ F_t^{R,2} \\ F_t^G \\ F_t^{R,1} \\ F_t^{R,2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \cdot (h_{1,t}^C)^{1/2} \\ \epsilon_{2,t} \cdot (h_{2,t}^C)^{1/2} \\ \epsilon_{3,t} \cdot (h_{3,t}^C)^{1/2} \\ \epsilon_{4,t} \cdot (h_{4,t}^C)^{1/2} \end{bmatrix}$$

With normalized error terms:

$$\epsilon_{1,t}, \epsilon_{2,t}, \epsilon_{3,t}, \epsilon_{4,t} \sim \mathcal{N}(0, 1)$$

The state equation becomes:

$$\begin{bmatrix} F_t^G \\ F_t^{R,1} \\ F_t^{R,2} \\ F_t^G \\ F_t^{R,1} \\ F_t^{R,2} \end{bmatrix} = \begin{bmatrix} \beta_{1,t}^G & 0 & 0 & \beta_{2,t}^G & 0 & 0 \\ 0 & \beta_{1,t}^{R,1} & 0 & 0 & \beta_{2,t}^{R,1} & 0 \\ 0 & 0 & \beta_{1,t}^{R,2} & 0 & 0 & \beta_{2,t}^{R,1} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} F_{t-1}^G \\ F_{t-1}^{R,1} \\ F_{t-1}^{R,2} \\ F_{t-2}^G \\ F_{t-2}^{R,1} \\ F_{t-2}^{R,2} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \cdot (A^{-1} \cdot h_{1,t}^G \cdot A^{-1^T})^{1/2} \\ \epsilon_{2,t} \cdot (h_{1,t}^R)^{1/2} \\ \epsilon_{3,t} \cdot (h_{2,t}^R)^{1/2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(2) Factor loadings: $B_{i,t}^G, B_{i,t}^R \mid F_t^J, v_{i,t}, \beta_{i,t}^C, h_{i,t}^C, Q_i^B$

Consider the following two equations:

$$\begin{aligned} Y_{i,t} &= B_{i,t}^G F_t^G + B_{i,t}^R F_t^R + v_{i,t} \\ v_{i,t} &= \beta_{i,t}^C v_{i,t-1} + (h_{i,t}^C)^{1/2} \epsilon_t^C \end{aligned}$$

The time-varying factor loadings are estimated equation-by-equation with a time-varying parameter OLS. The heteroscedasticity is removed by dividing the variables in equation 2 with the stochastic volatility and remove the correlation (Del Negro and Otrok, 2008).

This leads to the equation:

$$\begin{aligned} \frac{Y_{i,t} - \beta_{i,t}^C Y_{i,t-1}}{(h_{i,t}^C)^{1/2}} &= \beta_{i,t}^G \frac{F_t^G - \beta_{i,t}^C F_{t-1}^G}{(h_{i,t}^C)^{1/2}} + \beta_{i,t}^R \frac{F_t^R - \beta_{i,t}^C F_{t-1}^R}{(h_{i,t}^C)^{1/2}} + \epsilon_{i,t}^C \\ B_{i,t} &= B_{i,t-1} + (Q_i^B)^{1/2} \epsilon_t \end{aligned}$$

which is a linear state-space model for each i :

$$\begin{aligned} Y_{i,t} &= \beta_{i,t}^G F_t^G + \beta_{i,t}^R F_t^R + \epsilon_{i,t}^C \\ B_{i,t} &= B_{i,t-1} + (Q_i^B)^{1/2} \epsilon_t \end{aligned}$$

and can be estimated by applying the Kalman filter and the Carter and Kohn (1994) algorithm.

(3) Covariance matrix of the factor loadings: $Q_i^B | B_{i,t}$

Conditional on the draws of the time-varying parameters the posterior distribution is inverse-Wishart (Mumtaz and Musso, 2021): $p(Q_i^B | B_{i,t}) \sim \mathcal{IW}(B_{i,t} - B_{i,t-1} + Q_{i,0}^B, T + T_0)$. With T denoting the total sample size after discarding the training sample.

(4) Time-varying parameters of the VAR and model: $\beta_t^G | F_t^G, h_t^G, A$

Similar to the time-varying parameters in the factor model the sampling is carried out by first estimating the linear state-space model with the Kalman filter:

$$\begin{aligned} F_t^G &= \beta_t^G F_{t-1}^G + \Omega_t^{1/2} \epsilon_t^G \\ \beta_t^G &= \beta_{t-1}^G + Q_{\beta^G}^{1/2} e_t^G \end{aligned}$$

and then using the Carter and Kohn (1994) algorithm.

(5) Time-varying parameters of the AR models: $\beta_t^R, \beta_t^C | F_t^R, v_{i,t}, h_t^R, h_t^C$

The coefficients of the autoregressive model are sampled similar to the TVP-VAR parameters with the Kalman filter and the Carter and Kohn (1994) algorithm. The difference between the two models comes in the form of the state-space equation. In the case of the AR model specification, an equation-by-equation TVP-regression can be estimated, while for the VAR model, the parameters are sampled in one step. The state-space model has the form:

$$\begin{aligned} F_t^K &= \beta_t^K F_{t-1}^K + h_t^{K^{1/2}} \epsilon_t^K \\ \beta_t^K &= \beta_{t-1}^K + Q_{\beta^K}^{1/2} e_t^K \end{aligned}$$

with $K = \{R, C\}$.

(6) Stochastic volatilities: $h_t^G, h_t^R, h_t^C \mid g^G, g^R, g^C, \epsilon_t^G, \epsilon_t^R, \epsilon_t^C, e_t^G, e_t^R, e_t^C$

With the residuals of the factors VAR and AR models (equation 4, 8, 10) and the transition equations of the stochastic volatilities, each h_t is drawn with an independence Metropolis-Hastings algorithm from the following non-linear state-space model (Jacquier et al., 2002)

¹⁷:

$$r_t = (h_t^J)^{1/2} \epsilon_t^J$$

$$\ln h_t^J = \ln h_{t-1}^J + (g^J)^{1/2} e_t^J$$

Denoting the stochastic volatilities with $h_T = \{h_1, \dots, h_T\}$, the author simplifies the conditional distribution $p(h_t \mid h_T, r_t) = p(h_t \mid h_{t+1}, h_{t-1}, r_t)$ as the transition equation follows a random walk. The density of the simplified distribution is a product of a normal and a lognormal density and Jacquier et al. (2002) applies the log-normal density in the distribution as the candidate density in the Metropolis-Hastings algorithm. Formally:

$$p(h_t \mid h_{t+1}, h_{t-1}, r_t) = \underbrace{h_t^{-1/2} \exp\left(\frac{-r_t^2}{2h_t}\right)}_{\text{normal density}} \underbrace{h_t^{-1} \exp\left(\frac{-(\ln h_t - \mu)^2}{2\sigma_h}\right)}_{=: p_c, \text{lognormal density}}$$

with $\mu = \ln h_t - \ln h_{t-1}$. The independence Metropolis-Hastings algorithm acceptance probability is calculated as (Blake and Mumtaz, 2012) ¹⁸:

$$\alpha = \min \left(\frac{h_{t,\text{draw}}^{-1/2} \exp\left(\frac{-r_t^2}{2h_{t,\text{draw}}}\right)}{h_t^{-1/2} \exp\left(\frac{-r_t^2}{2h_t}\right)}, 1 \right) \quad (18)$$

The α acceptance rate is compared to a uniformly distributed random variable realization: if α is greater than the realization the $h_{t,\text{draw}}$ draw is accepted in the chain.

(7) Covariance matrix of the VAR model: $A \mid F_t^G, \alpha_t^G, h_t^G$

To draw the elements of the A matrix the steps introduced by Chan et al. (2023) are employed, but are not presented in detail. The main idea is to re-write the VAR equation

¹⁷The observation equation is $r_t = (\exp(\ln h_t^G))^{1/2} \epsilon_t^G$ that is simplified.

¹⁸Denoting the parameters $\Theta = \{h_{t-1}, h_{t+1}, r_t\}$, the acceptance rate in the Metropolis-Hastings algorithm:

$$\alpha = \min \left(\frac{p(\Theta_{\text{draw}}) / p_c(\Theta_{\text{draw}} \mid \Theta)}{p(\Theta) / p_c(\Theta \mid \Theta_{\text{draw}})}, 1 \right)$$

which can be simplified in the case of the independence Metropolis-Hastings algorithm.

and sample A by rows a_i ($i = 1, \dots, n$). Consider the reduced-form VAR:

$$\begin{aligned} \mathbf{F}_t^G &= \alpha_t + \beta_t^G \mathbf{F}_{t-1}^G + A^{-1} \epsilon_t^G & \epsilon_t^G &\sim \mathcal{N}(0, H_t) \\ \mathbf{F}_t^G - \alpha_t + \beta_t^G \mathbf{F}_{t-1}^G &= A^{-1} \epsilon_t^G & \epsilon_t^G &\sim \mathcal{N}(0, H_t) \\ (\mathbf{F}_t^G - \alpha_t + \beta_t^G \mathbf{F}_{t-1}^G) a_i &= \epsilon_t^G & \epsilon_t^G &\sim \mathcal{N}(0, h_t) \end{aligned}$$

It is unfeasible to sample from the conditional distribution of a_i . However, it has the same distribution as a vector with $n - 1$ variables with normal and 1 variables with absolute normal distribution. To sample from the absolute normal distribution a normal mixture approximation with 2 components is applied (Villani, 2009).

(8) Covariance of the stochastic volatilities: $g^J | h_t^J$:

The conditional posterior distribution of the covariances is inverse-gamma: $p(g^J | h_t^J) \sim \mathcal{IG}((\ln h_t^J - \ln h_{t-1}^J)^2 + g_0, V_g + T)$.

3.1.4 Variance decomposition

The main result of the model is the variance decomposition that determines the importance of the factors by calculating the share they explain from the total variation of the input data. In the research, it is applied to identify the extent to which each commodity is financialized. In the model the formula for the variance decomposition is the following:

$$Var(Y_{it}) = Var(B_{i,t}^G) Var(F_t^G) + Var(B_{i,t}^R) Var(F_t^R) + Var(v_{it}) \quad (19)$$

To calculate the variance of one factor the following formula is applied:

$$\begin{aligned} Var(F_t^J) &= Var(\beta_t^J F_{t-1}^J) + Var(\Omega_t) = Var(\beta_t^J (\beta_{t-1}^J F_{t-2}^J + \Omega_{t-1})) + Var(\Omega_t) = \dots \\ &= \underbrace{Var\left(\prod_{i=1}^T \beta_i^J F_0^J\right)}_{=0} + \underbrace{Var\left(\sum_{i=1}^T \left(\prod_{j=1}^i \beta_{t-j}^J\right) \Omega_i\right)}_{V_{F^J}} \end{aligned}$$

The variable of interest is the share of the variance explained by each global factor:

$$V_{F,G,t}^K = \frac{(B_{k,t}^G)^2 V_{F^K}}{(B_{-k,t}^G)^2 V_{F^{-k}} + (B_{i,t}^R)^2 V_{F^R} + Var(v_{it})} \quad (20)$$

where $K \in G$ denotes one of the global factors. The decompositions are time-varying due to the time-variation of the factor loadings.

3.2 Nonlinearity in the factor model

I present a modification of the model that allows us to measure the uncertainties in the commodity market, and estimate their impact on the U.S. macroeconomic, financial and monetary policy variables ¹⁹.

The first level of the model is the same dynamic factor model with time-varying parameters as described by equation 2 and 3 with one world factor and six regional factors that are the energy futures, monetary policy rates, financial variables, macroeconomic time series, agricultural- and industrial metals, and precious metals ²⁰:

$$Y_{it} = B_{i,t}^G F_t^G + B_{i,t}^R F_t^R + v_{i,t} \quad (21)$$

$$B_{i,t} = B_{i,t-1} + (Q_i^B)^{1/2} \epsilon_t^B, \quad (22)$$

$$\epsilon_t^B \sim \mathcal{N}(0, 1) \quad (23)$$

The factors dynamics are specified similarly to the previous model. The only world factor follows an AR(1) model ²¹. The main focus of the model is the regional factors and their interaction described by a VAR model. In contrast to the previous model, the stochastic volatility component is allowed to enter the mean equation of the VAR model as in Carriero et al. (2018); Chan (2017) and Mumtaz and Theodoridis (2020). Described first by Koopman and Hol Uspensky (2002) stochastic volatility in mean represents the empirical feature that the first momentum of return time series is dependent on the volatility. Mumtaz and Zanetti (2013) shows that disregarding the effect of volatility in mean would lead to the spurious conclusion that monetary policy shocks have insignificant impacts on macroeconomic data. For the estimation of uncertainty Carriero et al. (2018) points out the relevance of including the volatility in the mean equation. If the estimation of the uncertainty lacks the volatility-in-mean feature then the following analysis would be inconsistent given that in the second step, one usually includes the estimated uncertainty time-series in the mean equation to measure its beta.

¹⁹The replication code of the model is available at: Github

²⁰The regions are described in detail in the commodity and U.S. economic dataset of section 4.

²¹The world factor was necessitated in the model for reasons of stability to capture the small common variation of the 70 time series.

The VAR equation henceforth is specified in the following way:

$$F_t^R = \alpha + \sum_{i=1}^p \beta_i^R F_{t-i}^R + \sum_{j=0}^q \gamma_j \ln(h_{j,t}^R) + \Omega_{t,1}^{1/2} \epsilon_t \quad \epsilon_t \sim \mathcal{N}(0, 1) \quad (24)$$

The scale of the covariance matrix, A is kept as a full matrix to preserve the order invariance property:

$$\Omega_t = A^{-1} H_t A^{-1^T} \quad (25)$$

Similarly to the previous model I set $p = 2$ and set $q = 1$ so that the contemporaneous and one lag of the volatility can affect the mean. By allowing the stochastic volatility to have an immediate impact on the mean, the uncertainty becomes endogenous (Carriero et al., 2018).

The stochastic volatilities are allowed to interact in a VAR model, capturing the contagion of uncertainty shocks in times of financial stress or as a response to an external shock, as described by Carriero et al. (2018); Mumtaz and Zanetti (2013). Moreover, the regional factors that are the response variables in the VAR model enter the mean equation of the volatility. This represents a feedback effect of the volatility shocks through the response variables (Carriero et al., 2018; Mumtaz and Theodoridis, 2020):

$$\ln(h_t^R) = c + \sum_{i=1}^r \beta_i^{h,R} \ln(h_{t-i}^R) + \sum_{j=1}^s \delta_j F_{t-j}^R + (Q^{h,R})^{1/2} \eta_t \quad \eta_t \sim \mathcal{N}(0, 1) \quad (26)$$

$$E(\epsilon_t, \eta_t) = 0 \quad \forall t = 1, 2, \dots, T \quad (27)$$

Setting $r = 1$, $s = 1$ based on Mumtaz and Surico (2018).

To compute the impact of uncertainty shocks the covariance matrix $Q^{h,R}$ is specified as a full matrix and identified with the order of the variables.

In the model, a shock to one of the uncertainties affects immediately the response variables F_t^R through the volatility in mean and the covariance matrices Ω_t . In the following time steps, all other stochastic volatilities are also affected by the shock through the VAR relation of stochastic volatilities and the feedback effect of F_t^R in equation 26.

The global and idiosyncratic factors have the same AR(1) dynamics with time-varying

parameters as the idiosyncratic factors of the previous model:

$$\begin{aligned}
F_t^G &= \beta_t^G v_{i,t-1} + (h_t^G)^{1/2} \epsilon_{i,t}^G & \epsilon_t^G &\sim \mathcal{N}(0, 1) \\
v_{i,t} &= \beta_{i,t}^C v_{i,t-1} + (h_t^C)^{1/2} \epsilon_{i,t}^C & \epsilon_t^C &\sim \mathcal{N}(0, 1) \\
\beta_t^J &= \beta_{t-1}^J + Q_{\beta^J}^{1/2} e_{t,\beta}^J & e_{t,\beta}^J &\sim \mathcal{N}(0, 1) \\
\ln h_t^J &= \ln h_{t-1}^J + (g^J)^{1/2} e_{t,h}^J & e_{t,h}^J &\sim \mathcal{N}(0, 1)
\end{aligned}$$

with $J = \{G, C\}$ denoting the global and idiosyncratic factors respectively.

3.2.1 Priors and starting values

Similar to the previous model, the first $S_0 = 40$ observation are employed as a training sample to determine the prior parameters²². All the parameters unchanged compared to the previous model have the same prior distribution. Table 2 summarizes the starting values and prior parameters.

Parameter	Equation(s)	Prior distribution	Starting value
F_t	21	-	Principal components
$B_{i,t}$	21	$\mathcal{N}(B_{i,0}, V_{B_i})$	$B_{i,0}$, PCA loadings V_{B_i} , PCA residual covariance
β^R, γ, δ	24, 26	$\mathcal{N}(\hat{\beta}_0, Q \otimes \mathcal{S})$	$\hat{\beta}_0$ VAR coefficients estimate Q , covariance of the VAR coefficients $\mathcal{S}$, diagonal matrix of shrinkage parameters
A	24	$a_i \sim \mathcal{N}(0, I)$	$I \in R^{6 \times 6}$
$Q^{h,R}$	26	$\mathcal{IW}(0.1, n+1,)$	n = number of variables in the VAR model
$\bar{\beta}_t^G, \bar{\beta}_t^C$	28	$\mathcal{N}(\beta_0, \Sigma_0)$	$\beta_0 = (X_t' X_t)^{-1} (X_t' Y_t)$, OLS estimate $\Sigma_0 = \frac{(Y_t - \beta_t X_t)^{-1} (Y_t - \beta_t X_t)'}{S_0 - n}$
h_t^G, h_t^C	28	-	$(F_t^J - \alpha_t^J - \sum_{i=1}^{P_J} \beta_{i,t}^C F_{t-i}^J)^2 + 0.001$
g^G, g^C	28	$\mathcal{IG}(0.01, 1)$	-

Table 2: Summary of the parameters prior distributions in model 2.

Note. The table summarizes the prior distributions and starting values of the model described in this section. The parameter $S_0 = 40$ denotes the training sample size and $J = \{G, C\}$ the global and idiosyncratic factors respectively.

Importantly, shrinkage priors are specified for both VAR equations 24 and 26 to avoid

²²In the description, I denote the VAR coefficients and the constant with overline.

overfitting. The Minnesota priors include the belief that the coefficients in the VAR model follow a random walk. Furthermore, this prior enables us to add more weights to the recent lags. The parameterization of the prior sets the first two moments of the VAR coefficients ²³. Relying on the study of Bańbura et al. (2010), a normal inverse-Wishart Minnesota prior is implemented via dummy observations. The prior distributions are:

$$p(\beta | Q) = \mathcal{N}(\hat{\beta}_0, Q \otimes \mathcal{S}) \quad (28)$$

$$p(Q) = \mathcal{IW}(\bar{S}_0, \alpha_0) \quad (29)$$

Given that these priors are applied for both VAR equations, the notations stand for $Q = \Omega$, $Q^{h,R}$ and $\beta = \beta^R, \beta^h$. The prior parameters and starting values are estimated the following way:

- (1): Factors, loadings, the covariance of the TVP loadings, the covariance of the VAR model, the global and the idiosyncratic components,

$$F_t^G, F_t^R, B_{i,t}^G, B_{i,t}^R, Q^B, A, \beta_t^J, h_t^J, Q_{\beta^J}, g^J:$$

The initial values and priors are determined similarly to the previous model. The index $J = \{G, C\}$ denotes the global and idiosyncratic factors.

- (2): VAR coefficients, $\bar{\beta}^R, \bar{\beta}^{h,R}, \gamma, \delta$: The tightness of the VAR coefficients are set to 0.1, following Mumtaz and Surico (2018). This parameter defines the rate of shrinkage of the VAR coefficients to a random walk. For the intercept, the shrinkage is set to 1 000, indicating a flat prior. The shrinkage on the stochastic volatilities entering the VAR equation is set to 1 as in Mumtaz and Surico (2018).

For the VAR equation 26 of the stochastic volatilities, the overall tightness is set to $\lambda = 0.1$ and for the intercept to 100. Tightness on the lags of the external variables entering the stochastic volatility VAR is set to 0.05, following Mumtaz and Musso (2021). This way I include the prior belief that the feedback effect of the regional factors has a mild impact.

Larger values indicate less informative priors, while imposing 0 tightness means that the posteriors equal the prior distribution. In order to implement the dummy observations, the VAR coefficients and their scale parameters are estimated from an equation-by-equation AR(1). The scale parameters equal the standard deviation of AR(1) residuals.

²³For a detailed description of the Minnesota priors and their parameters I refer to Blake and Mumtaz (2012).

Starting values of the coefficients are obtained by estimating a VAR model on the training sample with OLS, similar to the previous model.

- (3): Stochastic volatility, h_t : Following Cogley and Sargent (2005) and Mumtaz and Surico (2018) starting values of the stochastic volatilities are set from the VAR residuals, estimated on the training sample. By denoting the diagonal elements of the variances Cholesky decomposition with μ_0 I set $h_0 \sim \mathcal{N}(\ln(\mu_0), \mathbf{I}_{n \times n})$.
- (4): Covariance matrix of the stochastic volatilities, $Q^{h,R}$: For the diagonal covariance matrix an inverse-Wishart prior is assumed with 7 degree of freedom that equals the number of variables (6) in the VAR equation plus one. The prior scale equals 0.1.

3.2.2 Posteriors

The posterior distributions are obtained via Gibbs sampling. The following steps are computed to draw from the posterior distributions of the variables ²⁴:

- (1) **Coefficients of the VAR model of the factors:** $\bar{\beta}^R, \gamma \mid F_t^R, A, h_t^R, v_{i,t}$

Similarly to the VAR(2) specification in the previous model, the VAR model can be transformed into a linear state-space model to draw the normally distributed coefficients with the Kalman filter and the Carter and Kohn (1994) algorithm.

- (2) **Covariance matrix of the VAR model:** $A \mid F_t^R, \bar{\beta}^R, h_t^R$

These parameters are drawn similarly to the previous model.

- (3) **Stochastic volatilities of the VAR model:** $h_t^R \mid F_t^R, \bar{\beta}^h, \gamma, A, h_t^R Q^{h,R}$

Stochastic volatilities are sampled with the ancestor particle Gibbs sampling of Lindsten et al. (2014). The applied algorithm is based on the study of Mumtaz and Theodoridis (2020) and at each step, 40 particles are used.

- (4) **Transition equation of stochastic volatilities:** $\beta^R, \delta \mid h_t, F_t$

As described in the prior specification I set a normal inverse-Wishart (conjugate) prior; therefore, the conditional posterior distribution of the VAR coefficients is also normal: $\beta^h \sim \mathcal{N}((x^T x)^{-1}(x^T y), Q^h \otimes ((x^T x)^{-1}))$. With x, y are denoted the posterior draws of the factors and the stochastic volatilities in the previous step augmented with the dummy observations.

²⁴I denote the VAR coefficients with the constants as: $\bar{\beta} = \{c, \beta_{1,t}, \beta_{2,t}, \dots, \beta_{p,t}\}$.

(5) Covariance matrix of the VAR of the stochastic volatilities: $Q^{h,R} | h_t, F_t, \bar{\beta}^h, \delta$

Similarly to the posteriors of the transition equation due to the conjugate priors the posterior distribution is inverse-Wishart: $Q^{h,R} \sim \mathcal{IW}(\mathcal{S}^*, T^*)$. Where $T^* = T + 7$ with T denoting the observations used to draw the posterior distribution, and the scaling factor equals the squared residuals of the VAR model plus the prior scaling factor.

(6) Factors F_t^G, F_t^R , factor loadings $B_{i,t}$, global and idiosyncratic components, $\bar{\beta}_t^G, \bar{\beta}_t^C, h_t^G, h_t^C, q^G, q^C, g^G, g^C$

The posterior distributions of these parameters are not affected by the changes in this model; henceforth, their posterior distributions are the same as in the previous model.

3.2.3 Generalized impulse response functions

In the model, the covariance matrix of the stochastic volatility VAR model, $Q^{h,R}$ in equation 26 is set to a full matrix and identified by ordering the variables. Considering that no study has analysed the interaction of commodity, macroeconomic, financial and monetary policy uncertainties, I rely on the studies that investigated the interaction of the level variables ²⁵.

Following Wang and McPhail (2014) for the three commodity factors I assume that the energy factor is the most exogenous and is not contemporaneously affected by the others. Then real economic activity and precious metal commodities are ordered. For the other three factors, based on existing literature, monetary policy is ordered before macroeconomic variables (Caldara and Herbst, 2019; Mumtaz and Zanetti, 2013). Relying on the VAR model specification of Park and Ratti (2008) energy commodity variables are ordered before the macroeconomic and financial variables.

To summarize, the variables are in the following order:

1. energy commodities
2. monetary policy
3. financial variables
4. macroeconomic time series
5. real economic activity (agricultural and industrial metal commodities)
6. precious metals.

Other common identifications involve sign restrictions or external instruments that are commonly used for monetary policy shocks (Gertler and Karadi, 2015; Mumtaz and Theodoridis,

²⁵By level shocks and variables I refer to the factor variables and the covariance matrix.

2020). However, to identify with external instruments uncertainty indices for all six variables of interest would be necessitated and as it was noted, existing literature is scarce on commodity uncertainty indices.

To account for the nonlinearity in the model that is the shock of the stochastic volatility in mean, in equation 24, I use the Monte Carlo integration described in Koop et al. (1996) to compute impulse-response functions ²⁶. The main idea is to treat the impulse responses as the realization of the difference between two random variables, namely the expected values of the output variables. Denoting the parameter draws from the Gibbs sampling by θ , the response variable by Y_t and the covariance to shock by $Q^{h,R}$ the formula is:

$$GIRF(Y_t, h_t, \theta) = E(Y_{t+k} | h_t, \theta, Y_t, \eta_t = \hat{s}) - E(Y_{t+k} | h_t, \theta, Y_t) \quad (30)$$

where k denotes the horizon of the impulse response and $\hat{s}$ the magnitude of the shock in the starting period. To evaluate the two conditional expectations Monte Carlo integration is applied that takes the following steps:

- (1) Select a realization (draw of the Gibbs sampling) of θ_t and simulate the covariance $Q_i^{h,R}$: multiplying its Cholesky decomposition by random normal variables over the forecasting horizon.
- (2) To calculate the first conditional expected value, on $t = 0$ shock the selected row of the covariance matrix: $\eta_t = \hat{s}$ and compute the values of $\hat{Y}_{t=1,2,\dots,k}$ by using the starting values $Y_{t=-2,-1,0}$, θ and the simulated draws of $Q_i^{h,R}$.
- (3) To calculate the second conditional expected value, on $t = 0$ simulate again a draw from the covariance matrix and forecast $Y_{t=1,2,\dots,k}$.
- (4) Repeat the steps above m times and apply the same realization of θ in each step.
- (5) Repeat the steps above n times and apply different realization of θ in each step.

By the law of large numbers taking the average of the obtained forecasts for each time horizon $t = 1, 2, \dots, k$ converges to the two conditional expectations.

In the calculation, the horizon is set to 48 months. The θ parameters are the last $n = 1,000$ values of the Gibbs sampling, obtained from 10 000 draws keeping every 10th draw after

²⁶The replication code of the generalized IRF is available at: Github

discarding 10 000 burn-in draws. To further account for the historical dependence, as described in Mumtaz and Theodoridis (2020) responses are estimated for every 12th month except for the last 48 periods that are employed to evaluate the responses at the factor level.

For a given time step and parameter selection, $m = 100$ Monte Carlo repetitions are used. In each period, I re-arrange the input or predicted variables as described in equation 21, 24, 26. At the regional level, a one-standard-deviation shock is added to the stochastic volatility of the energy commodities. If the covariance matrix, $Q^{R,h}$, of a Gibbs draw has a too-large condition number, the covariance matrix is constructed from the scale matrices: $\Omega_t := A^T A$ just as in Mumtaz and Theodoridis (2020). A detailed description of the condition number and its calculation is described in appendix A.1.

To compute the responses at the factor level, I do not apply simulation but use the Gibbs draws to evaluate the response. At the factor level, the uncertainty shocks only have a linear effect through the regional factors. Therefore, computing the factor loadings and the other factors with Monte Carlo integration would only introduce Monte Carlo noise to the results.

4 Data

In the analysis, three databases are employed for the commodity futures, the financialization explanatory variables and the U.S. economy. All databases are at a monthly frequency, and the sample period runs between February 1998 and December 2023 ²⁷.

In the first part of the analysis, the commodity and financialization database is applied and in the second model, the commodity and U.S. economy database are employed.

4.1 Commodity futures dataset

In the commodity database, monthly return data is sampled from the end-of-month prices of the generic first futures contracts. For commodity futures, it's common for traders to focus on the front-month contract. As the expiration date of a contract nears, traders transition to the next closest contract, creating a cyclical pattern in trading volume, open interest, and commodity futures prices (Bessembinder and Seguin, 1992; Holmes and Rougier, 2005).

²⁷The start of the sample period was determined by the availability of the commodity futures data. From the industrial metals the aluminium, lead, nickel and zinc were not available prior to August 1997 in the Bloomberg database.

Bessembinder (2018) defines the “roll yield” as the difference in the price of futures on the rollover day. Although the effect might be small in monthly returns, it could lead to spurious results. Therefore, following Demirer and Kutan (2010) I roll over to the second-month contract one week before the expiry of the first contract.

In the sample, I select commodity futures that have been widely analysed in the literature on financialization. The data is grouped into four classes that are the regional factors of the first model: agricultural, energy, industrial metal and precious metal.

For selecting the agricultural commodities I rely on the study of Nguyen et al. (2020); Ordu et al. (2018); Tang and Xiong (2012) and Ma et al. (2021). In the database, those futures are included that have been analysed by all four studies, namely: coffee, cocoa, corn, cotton and wheat futures. Additionally, due to identification restriction, I include feeder cattle, sugar and soybean futures²⁸. These three futures have been analysed by three of the aforementioned studies²⁹. The energy group is selected by including all energy commodities analysed by Adams and Glück (2015); Balli et al. (2019); Da et al. (2023); Ma et al. (2021); Nguyen et al. (2020); Tang and Xiong (2012) except for the RBOB gasoline as it would greatly curtail the sample period. The industrial and precious metals are selected similarly: including all from the studies of Adams et al. (2020); Adams and Glück (2015); Balli et al. (2019); Ma et al. (2021). Table 3 presents the commodity futures with the tickers and descriptive statistics of the log returns.

In the first model, time series are categorized into the four groups described by table 3. The reason for this is to isolate region-specific variations only to include common variations of commodities in the financial and macroeconomic factors.

In the second model, the commodity time series are grouped differently given that the goal of that model is to estimate uncertainty components. In the first group, agricultural and industrial metal futures are incorporated to represent global economic activity based on existing literature. Alquist et al. (2020) estimated a common factor of 40 agricultural and industrial metal commodities and found co-movement between the global factor and macroeconomic fluctuation. Baumeister et al. (2022) also estimated a real commodity price factor from agricultural and industrial metal commodity prices and concluded that it forecasts crude oil prices just as well as the World Industrial Production Index.

²⁸To estimate the time-varying factor loadings the first three loadings must be restricted to an identity matrix. I restrict soybean to load only on the financial, feeder cattle on the macroeconomic and sugar on the agricultural country factor. All other loadings are unrestricted.

²⁹Feeder cattle, sugar and soybean futures were included in the study of Ma et al. (2021); Ordu et al. (2018); Tang and Xiong (2012).

Futures	Ticker	Mean	Min	Max	Median	Volatility	ADF test
<i>Agricultural</i>							
Coffee	KC1	0.13	-10.53	17.02	-3.84	14.08	0.01
Cocoa	CC1	1.66	-13.34	13.76	3.54	13.76	0.01
Corn	C 1	0.92	-12.53	12.5	0.96	12.92	0.01
Cotton	CT1	0.34	-17.68	10.08	2.76	13.88	0.01
Feeder cattle	FC1	1.75	-9.52	6.49	3.48	6.98	0.01
Sugar	SB1	1.02	-14.19	12.46	4.57	14.88	0.01
Soybean	S 1	1.1	-15.89	8.17	0.84	11.61	0.01
Wheat	W 1	1.04	-11.87	16.57	-0.94	13.38	0.01
<i>Energy</i>							
Brent	CO1	2.64	-24.07	16.77	5.5	15.44	0.01
Gasoil	QS1	2.76	-16.3	12.51	3.95	15.35	0.01
Heating oil	HO1	2.78	-15.53	14.26	6.86	14.75	0.01
Natural gas	NG1	0.18	-20.84	23.51	2.26	24.7	0.01
WTI	CL1	2.39	-28.79	31.66	8.45	17.22	0.01
<i>Industrial metal</i>							
Aluminium	LA1	0.73	-7.39	7.04	-0.23	8.77	0.01
Copper	HG1	2.67	-17.88	14.08	1.7	11.29	0.01
Lead	LL1	2.29	-12.96	10.58	3.31	12.17	0.01
Nickel	LN1	1.88	-11.04	13.93	1.04	15.37	0.01
Zinc	LX1	1.44	-16.35	11.18	2.56	11.89	0.01
<i>Precious metal</i>							
Gold	GC1	3.23	-8.18	6.68	2.04	7.07	0.01
Palladium	PA1	2.6	-16.12	18.18	9.08	15.6	0.01
Platinum	PL1	1.58	-15.35	10.37	4.02	10.02	0.01
Silver	SI1	2.3	-13.22	12.3	0.53	13.35	0.01

Table 3: Summary statistics on commodity futures.

Note. The table reports the annualized mean and median log return, standard deviation (volatility), minimum, and maximum monthly returns expressed in percentages. The last column of the table shows the p-values of the Augmented Dickey-Fuller test. The sample period begins in February 1998 and ends in December 2023.

The second and the third groups are energy and precious metal commodities given that their impact on macroeconomic and financial variables is usually analysed separately from other commodities.

4.2 Financialization dataset

In this set, I gather time series that scholars found to influence financialization and explain commodity returns and equity-commodity co-movement. Summary statistics and the data source are described in table 4.

The economic variables are grouped into two regional factors. The first regional factor includes economic activity and production and the second regional factor represents uncertainty, consumer expectations and real interest rate.

In the first regional economic factor, the most important indicator is the Kilian (2009) index that is widely used to proxy global economic activity and measures demand for the industrial commodity market ³⁰. In the financialization literature, Ordu et al. (2018) and Adams et al. (2020) found that the explanatory power of the Kilian index became larger around the global financial crisis of 2008.

Baumeister et al. (2022) researched indicators that predict best global petroleum consumption and crude oil prices. They found that the World Industrial Production Index, developed by Baumeister and Hamilton (2019) is the most informative indicator; therefore, I also include it in the database as an indicator of economic activity. As another indicator of industrial commodity demand, the U.S. ISM Manufacturing Index is included. To control for the demand from emerging market economies Tang and Xiong (2012) analyses the correlation between commodity returns and the MSCI Emerging Market Index. Although the time series contains stock market data, it is used as a macroeconomic indicator for economic activity and commodity demand of emerging markets as in Tang and Xiong (2012) and Ordu et al. (2018).

In the second regional factor, the macroeconomic uncertainty index of Jurado et al. (2015) is included. The index is calculated as the average volatility of the h step ahead forecasting error. The main idea behind this measure is to capture the unforecastable component of hundreds of macroeconomic time series. Adams et al. (2020) also employs this measure in analysing financialization.

The second indicator is the University of Michigan Consumer Sentiment Index. In analysing the predictability of crude oil prices and petroleum demand, Baumeister et al. (2022) concluded that the index is an appropriate indicator for measuring short- and long-term expectations of households, signalling consumer demand.

³⁰“The index developed in this section is based on dry cargo single voyage ocean freight rates and is explicitly designed to capture shifts in the demand for industrial commodities in global business markets.” (Kilian, 2009).

The last indicator is the 1-month U.S. real interest rate. Adams et al. (2020) proxied real interest rate by extracting the YoY CPI change from the 3-month Treasury Bill rates. However, the so obtained time series are ex-post real interest rates that are not relevant information for economic agents (Carlstrom, 1993). I include the ex-ante 1-month real interest rate indicator of the Federal Reserve Bank of Cleveland. The calculation relies on Fisher's hypothesis, extracting the expected inflation from the nominal interest rate for a given maturity date.

The financial time series are also divided into two regional factors. The first factor contains volatility and financial stress indices while the second regional factor incorporates general stock market activity and price fluctuation indicators that explain financialization and commodity return co-movements.

The first regional financial factor includes the VIX index. In financialization literature, it is used as a proxy for investors' risk appetite, risk expectations and financial stress (Ma et al., 2021; Ordu et al., 2018). The authors found that the VIX significantly explains financialization and return co-movement, although Ma et al. (2021) noted that its impact is only relevant in periods of financial stress. This is supported by Cheng et al. (2015), showing that commodity index investment falls in response to an increase in the VIX. Another proxy for financial stress, the Kansas City Financial Stress Index is also used to measure the impact of financial stress on commodity volatility. The index includes a wider range of financial variables such as swap and yield spreads, and other market-implied volatility indices (Hakkio et al., 2009).

To measure the market liquidity the TED spread is included, which is widely used in the contagion literature and found to predict equity-commodity co-movement (Büyüksahin and Robe, 2014). Furthermore, Ordu et al. (2018) found that the financialization of agricultural commodities depends on the liquidity of financial markets.

The second regional factor incorporated the S&P500, which is widely used to measure co-movement between equity and commodity returns (Ma et al., 2021), and in the period of financialization, it explained a large share of crude oil returns variation (Adams and Glück, 2015). The U.S. Dollar Index is important as the commodity futures are denominated in USD and exchange rate fluctuations could affect commodities demand (Adams et al., 2020) and also the interaction between FX and commodity markets (Ordu et al., 2018). Following scholars the trade-weighted US Dollar Index is included in the database (Ma et al., 2021; Tang and Xiong, 2012) ³¹.

³¹The trade-weighted US Dollar Index is a weighted average of the dollar exchange rate with the currencies of the following countries: Argentina, Australia, Brazil, Canada, Chile, China, Columbia, Euro Area, Hong Kong,

Time series	Mean	Std	ADF test	Data source
<i>Macroeconomic variables</i>				
ISM Manufacturing PMI	-0.02	1.59	0.01	Bloomberg NAPMPMI
Kilian Index	-0.92	253.42	0.01	Dallas Fed
MSCI Emerging Markets	0.14	2.86	0.01	Bloomberg: 9J6X
World Industrial Production	0.08	0.4	0.01	Baumeister's webpage
Macroeconomic uncertainty	0.03	1.18	0.01	Ludvigson's webpage
Michigan Cons. Sent.	-0.06	2.95	0.01	Bloomberg: CONSENT:IND
Real 1-mo. Int.Rate	84.7	955.13	0.01	FRED
<i>Financial variables</i>				
Financial stress	-16.55	395.27	0.01	FRED
TED spread	94.13	208.22	0.01	Bloomberg: NFCINTED
VIX	-0.08	9.17	0.01	Bloomberg: VIX
Moody's BAA-AAA	0.06	3.85	0.01	FRED
USDIDX	0.02	0.54	0.01	Bloomberg USTRGBD
S&P 500	0.22	1.98	0.01	Bloomberg SPX
BIS OTC derivatives	0.31	2.11	0.01	Bloomberg: SPX

Table 4: Summary statistics on financialization dataset.

Note. The table reports the mean and standard deviation in percentage of the macroeconomic and financial variables in the financialization dataset. The 4th column reports the p-values of the Augmented Dickey-Fuller test of the transformed variables. The 5th column describes the data source and the ticker of the time series. The sample period begins in February 1998 and ends in December 2023.

Lastly, the BIS global outstanding OTC derivatives data is incorporated in the database to relate my financial indicators to those studies that measure financialization with commodity index trader (CIT) (Ma et al., 2021) positions, index investment (Irwin and Sanders, 2012) or other trader positions (Büyükaşahin and Robe, 2014). The BIS dataset includes all outstanding OTC swaps, futures and options on commodities in a semi-annual database. To sample monthly observations a cubic spline interpolation is applied.

4.3 U.S. economy dataset

In the second model, I analyse commodity uncertainty and its impact on the economy, relying on the large macroeconomic FERD-MD database of McCracken and Ng (2016) that is publicly available and updated in real-time.

Japan, Korea, Malaysia, Mexico, India, Indonesia, Israel, Philippines, Russia, Saudi Arabia, Singapore, Sweden, Switzerland, Thailand, Taiwan, United Kingdom and Venezuela.

Time series	Ticker	Transformation	ADF test
IP Index	INDPRO	5	0.01
IP: Manufacturing	IPMANSICS	5	0.01
Capacity Utilization: Manufacturing	CUMFNS	2	0.01
Real Personal Income	RPI	5	0.01
CPI: All items	CPIAUCSL	6	0.01
CPI for all urban consumers: less food and energy	CPILFESL	6	0.01
PCE: Chain index	PCEPI	6	0.01
PCE: Nondurable goods	DNDGRG3M086SBEA	6	0.01
PPI: Finished Goods	WPSFD49207	6	0.01
<i>Short run inflation expectation 1-year</i>	MICH	5	0.01
<i>Long run inflation expectation 5-year</i>	—	5	0.01
Civilian Employment	CE16OV	5	0.01
Civilian Unemployment Rate	UNRATE	2	0.01
All Employees: Total nonfarm	PAYEMS	5	0.01
Avg Weekly Hours : Goods-Producing	CES0600000007	2	0.01
Ratio of Help Wanted/No. Unemployed	HWIURATIO	2	0.01
Avg Hourly Earnings: Goods-Producing	CES0600000008	6	0.01
Housing Starts: Total New Privately Owned	HOUST	5	0.01
New Private Housing Permits (SAAR)	PERMIT	5	0.01
ISM Manufacturing: New Orders Index	NAPMNOI	1	0.01
New Orders for Durable Goods	AMDMNOx	5	0.01
Real manufacturing and trade sales	CMRMTSPLx	5	0.01
Capacity utilization in manufacturing	CUMFNS	2	0.01
Real PCE	DPCERA3M086SBEA	5	0.01
<i>National Activity Index</i>	CFNAI	2	0.01
Euro (D-Mark)	EURUSD	5	0.01
British Pound	GBPUSD	5	0.01
Canadian Dollar	EXCAUSx	5	0.01
Swiss Franc	EXSZUSx	5	0.01
Real Broad Effective Exchange Rate for U.S.	RBUSBIS	5	0.01
S&P 500 index	SPX	5	0.01
S&P's Composite Common Stock	SPTSX	5	0.01
<i>Nominal home price index</i>	CSUSHPINSA	5	0.053
<i>SMB FF</i>	SMB	2	0.01
<i>HML FF</i>	HML	2	0.01
<i>Momentum factor</i>	Mom	2	0.01
<i>5 Industry Portfolios: 1st</i>	Cnsmr	2	0.01
<i>5 Industry Portfolios: 2nd</i>	Manuf	2	0.01
<i>5 Industry Portfolios: 3rd</i>	HiTec	2	0.01
<i>5 Industry Portfolios: 4th</i>	Hlth	2	0.01
<i>5 Industry Portfolios: 5th</i>	Other	2	0.01
AAA corporate bond yield - 10Y Treasury Rate	AAA	2	0.01
BAA corporate bond yield - 10Y Treasury Rate	BAA	2	0.01
<i>Excess bond premium</i>	—	2	0.01
3-Month Treasury Bill	TB3MS	2	0.017
1-Year Treasury Rate	GS1	2	0.01
5-Year Treasury Rate	GS5	2	0.01
10-Year Treasury Rate	GS10	2	0.01
<i>Shadow rate</i>	—	2	0.015
Fed Funds rate	FEDFUNDS	2	0.017

Table 5: Description of the U.S. economy dataset.

Note. The table summarizes the U.S. macroeconomic dataset. In the first column, the time series that augment the FRED-MD database are noted with italics. The second column presents the tickers from the FRED-MD database and Bloomberg. If a time series is not from the FRED-MD database or Bloomberg, the hyperlink is included in the second column. The last column reports the p-values of the Augmented Dickey-Fuller test estimated on the transformed time series. The sample period begins in February 1998 and ends in December 2023.

The database is widely used for empirical analysis of uncertainty estimation, monetary policy transmission and crude oil shocks (Alessandri and Mumtaz, 2019; Carriero et al., 2018; Känzig, 2021; Ludvigson et al., 2021; Miranda-Agrippino and Ricco, 2021). The data is grouped into three categories and I estimate from them the following three factors: macroeconomic, financial and monetary policy.

I augment the group of macroeconomic time series with the University of Michigan Inflation Expectation Survey and the National Activity Index of the Fed based on the study of Chang et al. (2023) and Aastveit et al. (2017) respectively. The financial time series are augmented with three Fama-French factors, five industry portfolio factors, the nominal home price index of Case and Shiller (1987), and the excess bond premium of Gilchrist and Zakrajšek (2012). Additionally to the monetary policy time series, the shadow rate of Wu and Xia (2016) is added to represent the impact of monetary policy with a hypothetical interest rate in the zero-lower bound period.

Table 5 provides a description of the time series, their ticker in the FRED-MD database. Variables that augment the FRED-MD database are denoted in italics and their ticker includes the source of data. The third row denotes the transformation suggested by McCracken and Ng (2016) to convert the data to stationary³² and the last row reports the p-values of the Augmented Dickey-Fuller test.

5 Empirical results

5.1 Financialization

In this section, I present the variance decomposition of the dynamic factor model described by equations (2)-(12). Financialization is defined as the share of the variance explained by the financial variables similar to the approach of Adams et al. (2020). Figure 1 illustrates the

³²The transformation codes are the following (McCracken and Ng, 2016):

1. no transformation
2. first difference: Δx_t
3. second difference: $\Delta^2 x_t$
4. logarithmic transformation: $\log(x_t)$
5. first difference of the logarithmic transformation: $\Delta \log(x_t)$
6. second difference of the logarithmic transformation: $\Delta^2 \log(x_t)$
7. return transformation: $(x_t/x_{t-1}) - 1$

variance decomposition for each asset and the average contribution of the three factors to the variation of agricultural, energy, precious- and industrial metals. In alignment with existing literature, the contribution of financial variables peaked between 2007-2008 (Bianchi et al., 2020; Cheng and Xiong, 2014).

Considering agricultural commodities the results indicate that cocoa and cotton have larger exposure to financialization, but coffee and corn were not exposed to financial variables, only during the global financial crisis. In comparison, Ordu et al. (2018) concluded significant financialization of cocoa and further explained that the low level of financialization of the corn³³ and coffee futures might be due to the poor performance of these commodities. This was also confirmed by Adams and Glück (2015). Moreover, Nguyen et al. (2020) found significant financialization of the cotton, cocoa and coffee market.

In the energy market, a strong financialization between 2007-2013 is estimated. These findings are in alignment with the study of Ma et al. (2021); Nguyen et al. (2020), except for the natural gas futures market where financial variables explain little variation. I further compare this result to the R^2 LMG decomposition approach of Adams et al. (2020), illustrated in figure A1, and show that both results find little financialization in the natural gas market.

Regarding the industrial- and precious metals, increased financialization for copper, platinum and palladium is documented. The study of Ma et al. (2021) confirmed that copper is one of the three central nodes in the commodity connectedness network. In the case of copper and platinum Nguyen et al. (2020) found a positive, and significant correlation with the equity market.

The low financialization of gold was described by Adams and Glück (2015) finding a negative correlation between gold and equities and further confirmed by Bianchi et al. (2020). Therefore, I conclude that the results are in line with scholars.

I further compare the results to three other measures of financialization estimated in the existing literature. The first and closest approach to mine is the LMG R^2 decomposition of Adams et al. (2020). For each commodity in the database, I replicate the decomposition with a 60-day rolling window regression and the macroeconomic and financial variables described by the study³⁴. The results for individual time series are illustrated in figure A1.

³³Corn futures market is the largest among the analysed commodities with an average of the net long notional value of 378,99 between November 2006 and December 2015 as reported by the CFTC Index Investment Database; therefore, one would expect significant impact of financialization in the corn market.

³⁴From the four financial variables I exclude the risk premium given that the data is not available in any database.

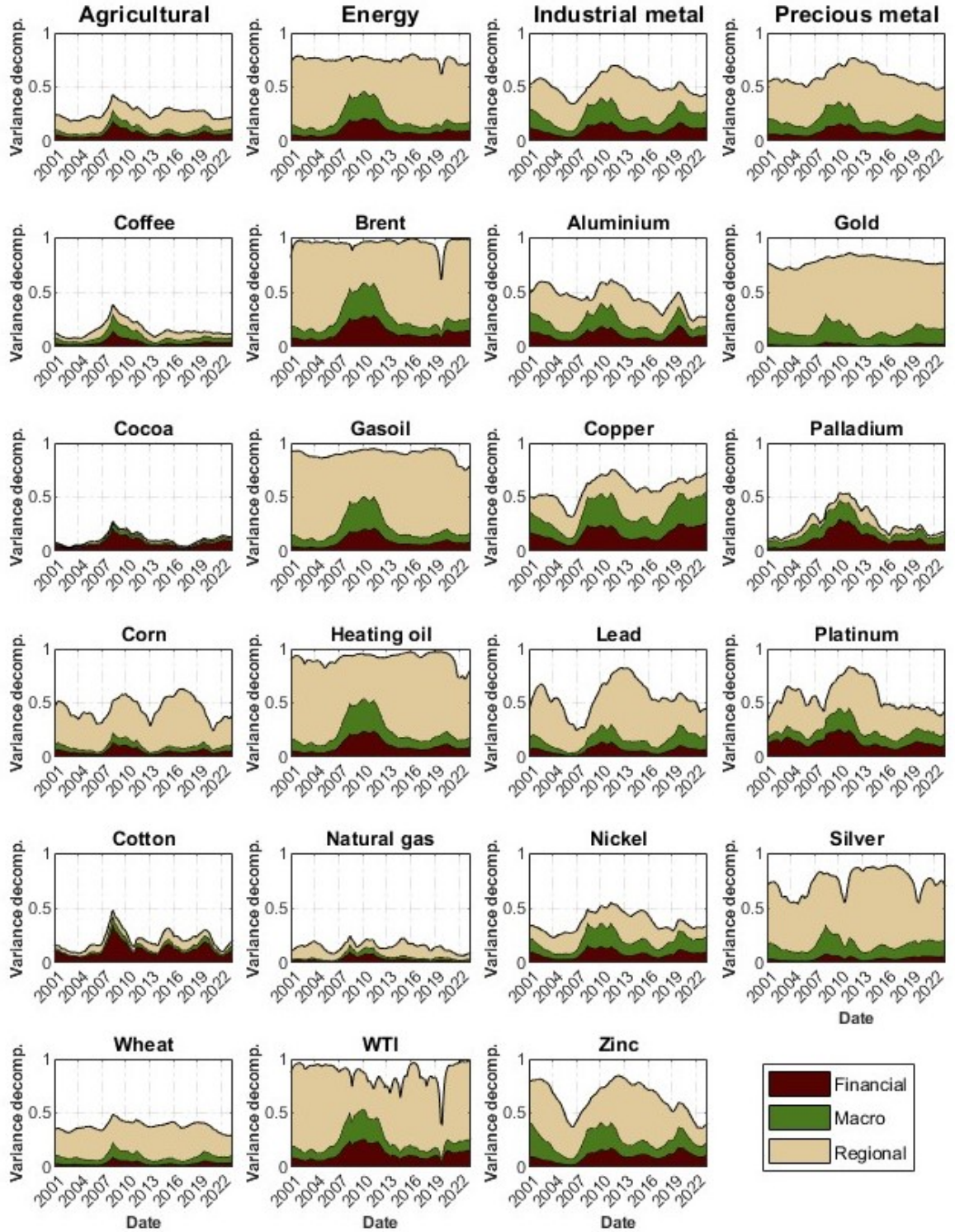


Figure 1: Variance decomposition of commodity returns.

Note. The figure illustrates the contribution of the financial, macroeconomic and regional factors to the variation of commodity returns estimated with the factor model of equations (2)–(12). The top row depicts the average contribution of the 3 factors to the agricultural, energy, industrial- and precious metal commodities respectively. The databases are summarized in table 3 and 4. The sample period starts in May 2001 and ends in December 2023.

Second, financialization was also identified with the conditional- and unconditional correlation

between the equity indices and commodity futures as in Cheng and Xiong (2014); Nguyen et al. (2020); Ordu et al. (2018). To represent this approach I calculate the correlation between the S&P500 and the individual commodities in a 100 day rolling window. Results for individual commodities are depicted A2.

Lastly, I use the notional value of the net long positions in Index Investment Data of the CFTC ³⁵ to compare my results to studies that proxy financialization with trading activity in the commodity futures market (Henderson et al., 2015; Irwin and Sanders, 2012; Ordu et al., 2018). The index investment data available for the commodities in the sample are depicted in figure A3.

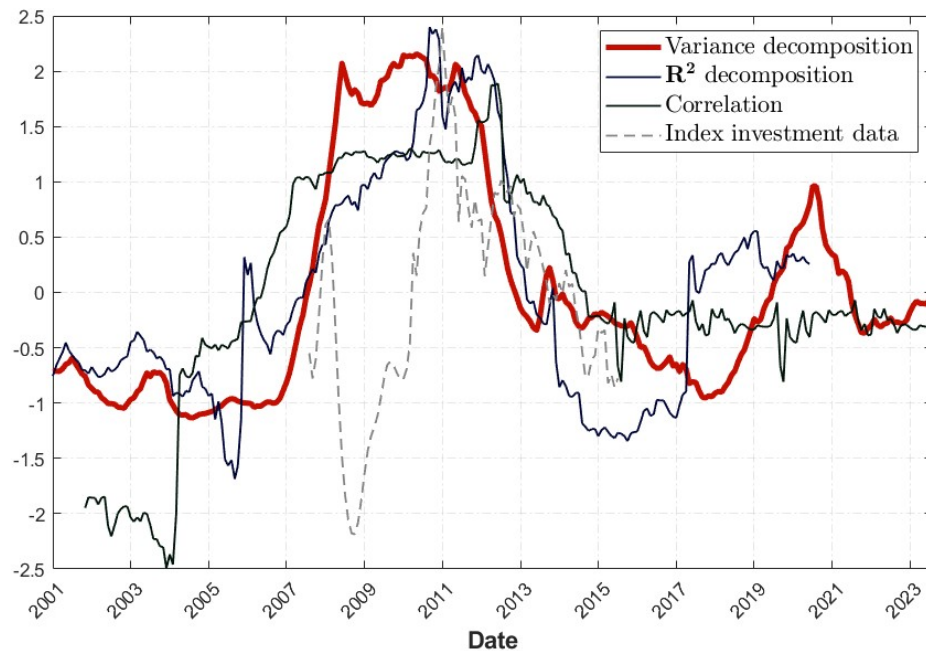


Figure 2: Financialization indices in the existing literature.

Note. The figure illustrates the average of the estimated financialization indices of figure 1 and three other proxies of financialization. The average R^2 decomposition of Adams et al. (2020), the average equity commodity co-movement and the average of the CFTC's index investment data. The time series are standardized in the full sample period. The R^2 decomposition and the equity commodity co-movement are measured in a rolling window their values correspond to the middle point of time in the window. The sample period starts in May 2001 and ends in December 2023.

Figure 2 illustrates the average share of the variance explained by the financial global factor in the factor model and the average of the estimated three indices. I infer from the graph that based on the variance, R^2 decomposition and the correlation-based indices financialization increased from 2007-2008 and dropped rapidly after 2013. Furthermore, the variance and R^2

³⁵Between December 2007 and November 2015 the data is reported in quarterly frequency. To obtain only frequency I use a cubic spline interpolation.

decomposition indicate a moderate increase from 2017 and 2021 respectively. In my index, the second peak originates from the precious metal sector and might capture the drop in stock market and industrial production during the lockdowns of the pandemic. It was a period when the VIX spiked up, indicating financial stress. As described by Cheng et al. (2015) as a response to a higher VIX index investment in the commodities fall. The sudden decline in commodity futures investment and the equity market explains the stronger explanatory power of financial variables. It was also concluded by Adams and Glück (2015) and Ordu et al. (2018) that the VIX significantly explains return co-movement among commodities and with equities in times of financial stress.

	Variance decomposition	R^2 decomposition	Equity commodity co-movement	Index investment
Variance decomposition	1.00	0.88***	0.92***	-0.10
R^2 decomposition	0.88***	1.00	0.91***	-0.32***
Eq. cmdty co-movement	0.92***	0.91***	1.00	-0.05
Index investment	-0.10	-0.32***	-0.05	1.00

Table 6: Correlation of the financialization factors.

Note. The table illustrates the average Pearson correlation between the contribution of financial variables to the variance decomposition and three other proxies of financialization. The R^2 decomposition is the financialization index of Adams et al. (2020). The equity, commodity co-movement the correlation approach of Nguyen et al. (2020); Ordu et al. (2018); Tang and Xiong (2012) among others and the last row is the Index Investment Data provided by the CFTC and used by Irwin and Sanders (2012). The *, **, *** marks denote the p-values at 10%, 5% and 1%.

Table 6 reports the correlation between the time series of financialization and the three other approaches. To obtain a single measure of financialization the unweighted average of the individual time series are applied respectively. The correlation is computed between December 2007 and November 2015 due to the availability of the index investment data.

Given that the R^2 decomposition and the equity commodity co-movement are measured in a rolling window their values correspond to the middle point of time in the window. This way I account for the fact that the CFTC index investment and my measure of financialization do not include large lagged data in the sample.

Based on the average correlation, my measure of financialization correlates strongly with the one of Adams et al. (2020) and the equity commodity co-movement approach. The table indicates that correlations are significant at 1%. The CFTC index investment shows a negative correlation with my index and the two other indices as well. A possible explanation of the

insignificant relation could be the limitations of the available index investment data. First, the “special call” letter did not define the index investment activity, its description was provided by the CFTC staff through discussions with the institutions. Second, time series are available only for 12 out of the analysed 19 commodity futures; therefore, the average financialization index might be biased towards the agricultural commodity futures.

I report the correlation between the estimated financialization index and the three other approaches for each commodity futures asset individually in table A1 appendix A.2.

In conclusion, I find strong financialization in the energy, industrial- and precious metal markets with exception of the natural gas, gold and silver. The results are consistent with two out of the three measures of financialization applied estimated in the existing literature.

5.2 Commodity uncertainty

In this section, I present the results of the factor model that enables non-linearity within the VAR equation of the regional variables, described in subsection 3.2.

The main results in this section are impulse responses of the variables to a one standard deviation shock to the energy uncertainty. Shock to the energy uncertainty immediately affects the regional factors in the VAR model as the stochastic volatilities enter the VAR equation and the response variables are also affected through the factor model structure. Some time steps later, all uncertainties are affected by the shock through their connectedness in the VAR model. As described in section 3.2.3, I calculate the impulse response functions with Monte Carlo integration, 48 months ahead and identify the covariance matrix of the stochastic volatilities by ordering the variables.

First, I analyse the impulse responses in the same region where the uncertainty shock originated. Figure 3 illustrates that the responses are homogeneous among the five variables and after 20 periods the shock becomes insignificant. I would like to refer back to the previous results regarding financialization, where natural gas, unlike all other energy commodities, had little common variation with financial and macroeconomic variables. From this figure, I also infer that the shock propagates less to the U.S. natural gas market confirming that the market is less connected to other energy commodities.

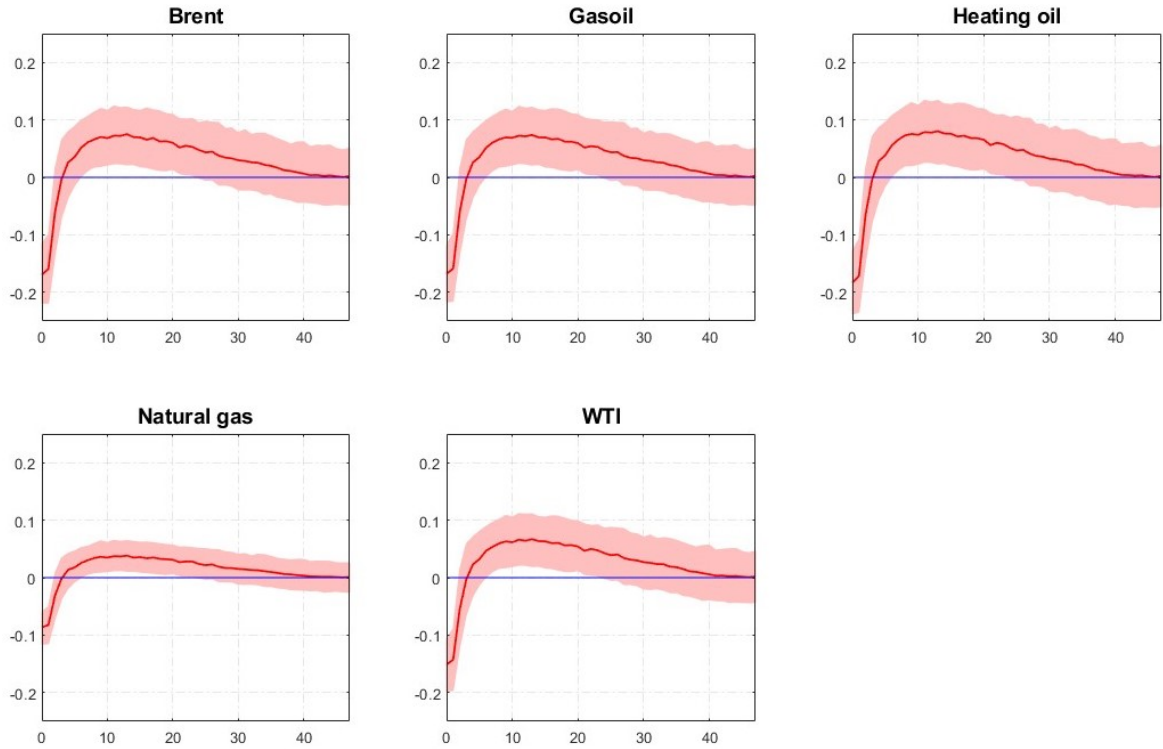


Figure 3: Impulse response functions of energy commodity futures.

Note. The figure illustrates the 48 months ahead average impulse response function of the energy commodity futures to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.

Second, I turn to the analysis of the macroeconomic variables that capture inflation. Figure 4 illustrates the impulse responses of the 7 variables that include CPI, PCE, PPI data and inflation expectations. From the responses I conclude that a positive energy shock leads to a substantial increase in consumer prices; however, the effect wears off within a year. These results are in alignment with the insignificant response of the 5-year inflation expectations and with the almost zero impulse response of the 1-year inflation expectations. Furthermore, I also note that the core CPI that is the CPI for all items less food and energy has no significant response. This confirms that the shock indeed originated from the energy sector and also indicates that the uncertainty shock did not propagate further by increasing the cost of all other consumer items.

Last, I investigate the impulse response of macroeconomic variables that are associated with economic activity. Figure 5 illustrates the industrial production activity in the first row, the employment in the second row and the commercial and national activity in the last row. Industrial production shows a short, countercyclical response. The employment data also reflects a countercyclical shock and the employment and unemployment variables have an almost symmetric response. In the last row, real manufacturing and sales have the largest negative response and

the national activity index shows almost zero response to the increased energy uncertainty.

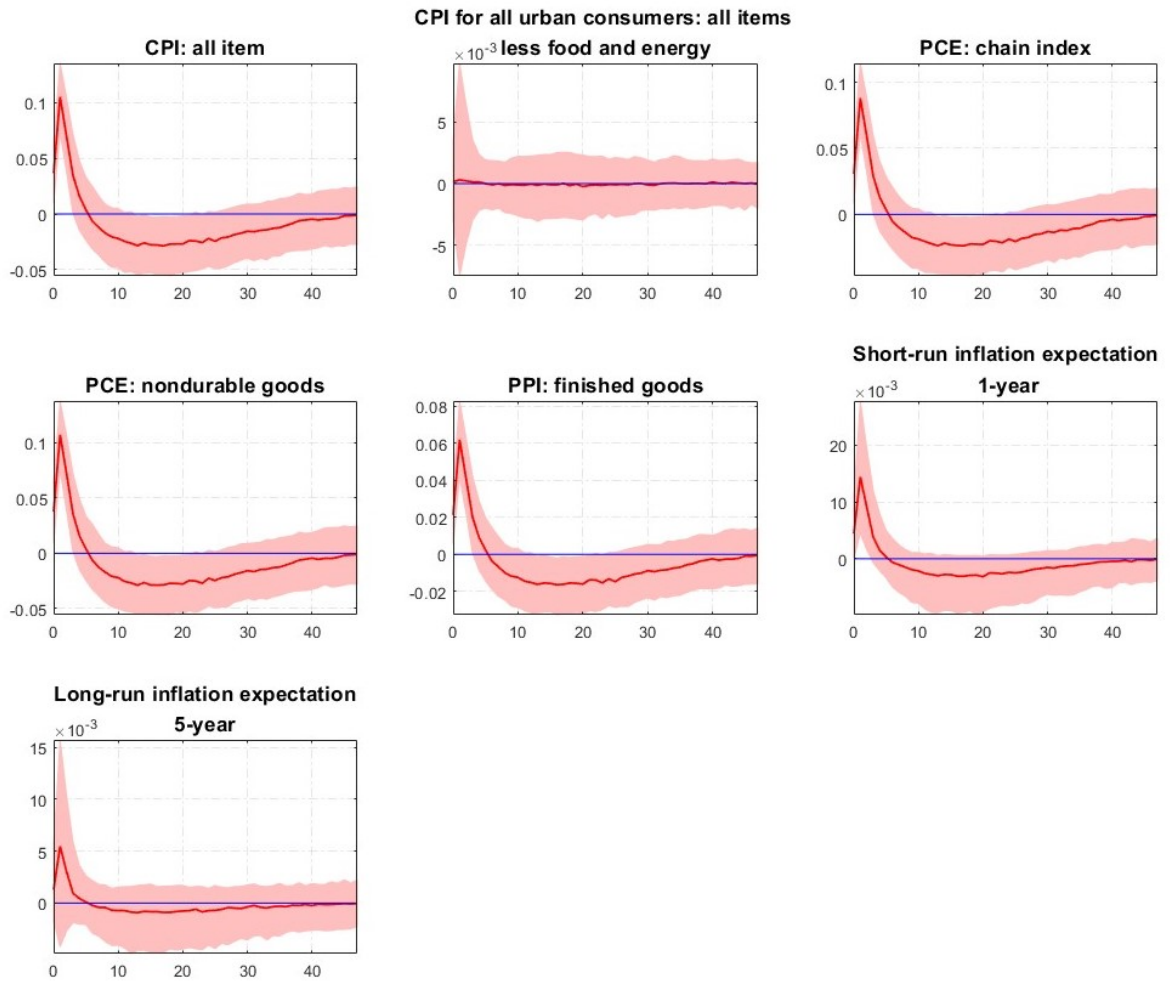


Figure 4: Impulse response functions of macroeconomic price level time series.

Note. The figure illustrates the 48 months ahead average impulse response function of the macroeconomic inflation variables to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.

Altogether, I conclude that the energy uncertainty shock leads to a strong countercyclical effect that re-bounds quickly within a year, in alignment with the impulse responses of the inflation variables. For industrial production, the variables respond almost identically.

All impulse-response functions are presented in appendix A.3 by figure A4, A5 and A6. An important conclusion of those figures is that the energy commodity shocks did not affect significantly the prices of other commodity groups, except for some of the precious metals that are known for their countercyclical fluctuation with inflation.

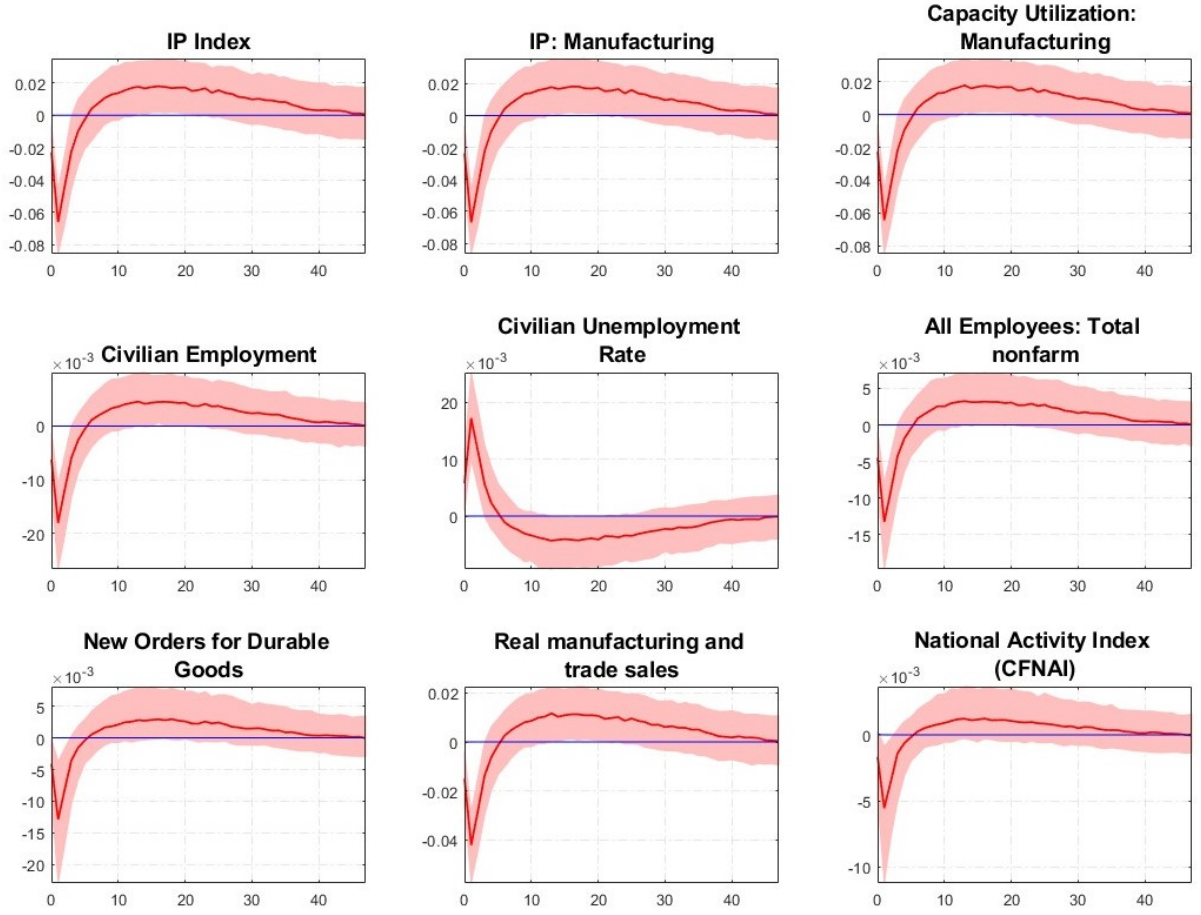


Figure 5: Impulse response functions of macroeconomic activity and employment time series.

Note. The figure illustrates the 48 months ahead average impulse response function of the macroeconomic activity variables to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.

In conclusion, I analyse the response of the commodity and macroeconomic and financial variables to a one standard deviation shock to the energy commodity uncertainty. Considering the macroeconomic data I see a countercyclical impact that dissipates within a year: consumer prices rise, industrial production decreases and unemployment increases. I further conclude that the shock to energy uncertainty homogeneously increases all energy returns and affects precious metal prices as well. Nevertheless, there is no significant contagion between the energy and the real economic activity commodity variables.

6 Robustness tests

This section presents alternative specifications to the factor model we applied to estimate the financialization index to show that results do not depend on the arbitrary choice of data

or priors. To this end, I conduct four robustness checks. I re-estimate three times the model described in section 3.1 with 10 000 draws from which every 10th draw of the last 5 000 samples are kept.

Two robustness checks concern the selection of the macroeconomic and financial variables applied in the variance decomposition to measure financialization. The variables are described in table 4. To illustrate that the estimated variance decomposition does not depend on a single variable one of the variables from each group is left out and the model re-estimated. First, the BIS OTC derivatives contract data from the financial variables and the University of Michigan Consumer Sentiment Index from the macroeconomic data are omitted. These variables were chosen as they have not been used by scholars in the investigation of financialization. I compare the variance explained by the macroeconomic and financial variables for each commodity futures asset. The correlation between the variance decomposition of the benchmark model and the first robustness test is illustrated in the first and fourth rows of table 7. The correlation between the explained shares of the financial and macroeconomic variables is more than 80% for each asset and above 90% in the case of financial variables except for two assets.

In the second robustness check, the Kansas City financial stress index is omitted from the financial variables and the world industrial production index of Baumeister and Hamilton (2019) from the macroeconomic ones. The second and fifth row of table 7 illustrates the results. Again I conclude that the large correlation indicates that removing some variables from the financial and macroeconomic data does not alter the results.

In the third test, I intend to illustrate that the choice of the prior parameters does not limit the results. As outlined in the section 3.1.1 of the methodology, most of the prior parameters are selected based on existing literature or estimated on a training sample. Nevertheless, the priors of the covariance matrices have an important role in determining the variation and sudden changes of the time-varying coefficients.

The covariance matrices involved in the robustness check are:

1. The covariance of the time-varying factor loadings, Q^B in equation 3.
2. The covariance matrix of the time-varying VAR coefficients, $Q_{\beta G}$ in equation 7.
3. The covariance matrices of the regional and the idiosyncratic factors AR coefficients, $Q_{\beta R}, Q_{\beta C}$ in equation 9 and 11 respectively.

Commodity futures	Financial factor			Macroeconomic factor		
	Test 1	Test2	Test 3	Test 1	Test2	Test 3
Coffee	0.97	0.99	0.99	0.80	0.96	0.71
Cocoa	0.97	0.99	0.9	0.95	0.9	0.94
Corn	0.91	0.90	0.93	0.91	0.93	0.74
Cotton	0.95	0.96	0.98	0.97	0.91	0.87
Wheat	0.92	0.88	0.93	0.81	0.90	0.76
Brent	0.99	0.99	0.99	0.96	0.98	0.54
Gasoil	0.99	0.99	0.98	0.92	0.97	0.33
Heating oil	0.99	0.99	0.98	0.92	0.97	0.34
Natural gas	0.98	0.97	0.91	0.97	0.98	0.99
WTI	0.99	0.99	0.97	0.99	0.99	0.69
Aluminium	0.95	0.96	0.87	0.94	0.93	0.94
Copper	0.97	0.98	0.97	0.97	0.97	0.89
Lead	0.98	0.93	0.92	0.96	0.91	0.84
Nickel	0.90	0.92	0.82	0.92	0.98	0.92
Zinc	0.98	0.97	0.85	0.98	0.97	0.87
Gold	0.89	0.97	0.78	0.86	0.94	0.57
Palladium	0.92	0.95	0.97	0.83	0.95	0.53
Platinum	0.89	0.96	0.98	0.87	0.93	0.73
Silver	0.92	0.94	0.93	0.92	0.84	0.74

Table 7: Robustness tests of the variance decomposition

Note. The table shows the Pearson correlation between the variance explained by the financial and macroeconomic factors of the benchmark model and those of the robustness check. Test 1 indicates the model estimated with the exclusion of the BIS OTC derivatives and the University of Michigan consumer sentiment index time series. The model of test 2 excludes the Kansas City financial stress index and the world industrial production index. The thirist model increases prior scale of the coefficients covariance matrices in equation 3, 7, 8, 10 by 10 times. The sample period starts in February 1998 and ends in December 2023.

Each covariance matrix had an inverse-Wishart distribution with a prior scaling factor $V_{B_i} \times S_0 \times 3.5 \times 10^{-4}$ where V_{B_i} denotes the covariance of the regression on the training sample with S_0 observations. Although this prior was applied by Blake and Mumtaz (2012); Mumtaz and Musso (2021) possible concerns involve the ad-hoc re-scaling of the estimated covariance. Therefore, a non-informative prior is specified by increasing the scaling factor 10 times to $V_{B_i} \times S_0 \times 3.5 \times 10^{-3}$.

Time series	Feedback effect	Vol-in-mean effect
Brent	-0.74	-0.18
Gasoil	-0.75	-0.18
Heating oil	-0.74	-0.18
Natural gas	-0.74	-0.18
WTI	-0.75	-0.18
3-Month Treasury Bill	-0.79	0.11
1-Year Treasury Rate	-0.78	0.19
5-Year Treasury Rate	-0.73	0.05
10-Year Treasury Rate	-0.73	0.06
Shadow rate	-0.74	0.05
Fed Funds rate	-0.74	0.02
EURUSD	-0.65	-0.02
GBPUSD	-0.65	-0.02
CADUSD	-0.63	-0.07
CHFUSD	-0.65	0
Real Exchange Rate	-0.64	-0.03
S&P 500	-0.65	-0.02
S&P's CCS	-0.65	-0.01
Nominal home price index	-0.65	0.11
SMB FF	-0.67	-0.12
HML FF	-0.59	-0.3
Momentum factor	-0.65	-0.05
5 Ind. Portfolios 1	-0.65	0.2
5 Ind. Portfolios 2	-0.7	-0.02
5 Ind. Portfolios 3	-0.65	0
5 Ind. Portfolios 4	-0.65	-0.02
5 Ind. Portfolios 5	-0.71	0.31
AAA	-0.65	-0.01
BAA	-0.65	-0.01
Excess bond premium	-0.64	-0.01
IP Index	-0.6	-0.23
IP: Manufacturing	-0.61	-0.23
Capacity Utilization: Manufacturing	-0.6	-0.24
Real Personal Income	-0.74	-0.19
CPI: All items	-0.61	-0.22
CPI: all items less food and energy	0.11	-0.07
PCE: Chain index	-0.61	-0.22
PCE: Nondurable goods	-0.61	-0.21
PPI: Finished Goods	-0.6	-0.22
Short run inflation expectation 1-year	-0.53	-0.35
Long run inflation expectation 5-year	-0.37	0.04
Civilian Employment	-0.62	-0.33
Civilian Unemployment Rate	-0.61	-0.11
All Employees: Total nonfarm	-0.58	-0.2
Avg Weekly Hours : Goods-Producing	-0.61	-0.23
Ratio of Help Wanted/No. Unemployed	-0.58	-0.17
Avg Hourly Earnings: Goods-Producing	-0.47	-0.17
Housing Starts: Total New Privately Owned	-0.61	-0.2
New Private Housing Permits	-0.57	-0.14
ISM Manufacturing: New Orders Index	-0.4	-0.07
New Orders for Durable Goods	-0.58	-0.27
Real manufacturing and trade sales	-0.62	-0.22
Capacity utilization in manufacturing	-0.6	-0.24
Real PCE	-0.61	-0.28
National Activity Index	-0.56	-0.22
Soybean	-0.57	-0.3
Coffee	-0.56	-0.32
Cocoa	-0.55	-0.32
Corn	-0.58	-0.32
Cotton	-0.56	-0.31
Wheat	-0.57	-0.3
Aluminium	-0.57	-0.31
Copper	-0.57	-0.31
Lead	-0.57	-0.31
Nickel	-0.58	-0.3
Zinc	-0.57	-0.31
Gold	-0.75	-0.08
Palladium	-0.74	-0.07
Platinum	-0.74	-0.09
Silver	-0.73	-0.08

Table 8: Robustness tests of second model specification.

Note. The table shows the Pearson correlation between the 36 period ahead impulse response functions of the benchmark model and its modifications. The first row reports the result where $s = 0$ in equation 12 and the second row when $q = 0$ in equation 12. The sample period starts in February 1998 and ends in December 2023.

Rows three and six of table 7 report the correlation between the benchmark results and the variance decomposition of this third robustness check. For almost all of the commodities the correlation between the two financialization factors is more than 80%; therefore, I conclude that the results are robust to large changes in the prior of the scale covariance matrix.

7 Conclusion

I analyse two major phenomena in the commodity market: financialization and uncertainty. To estimate financialization we employ a methodologically more complex model than existing literature that adjusts for the shortcomings of previous models, namely the rolling window estimations. In our model, which is a Bayesian factor model with time-varying parameters, stochastic volatility and VAR dynamic of the factors, we unify three definitions of financialization: co-movement between the financial and commodity markets, growing investment in commodity derivatives and the increased role of financial variables in explaining commodity return variations. In the estimation, the commodity futures, macroeconomic and financial variables are selected based on existing literature on financialization.

The estimated financialization index shows increased financialization between 2007 and 2013 in alignment with the existing literature (Adams et al., 2020; Ma et al., 2021; Tang and Xiong, 2012). I find that the energy market and certain industrial metals are mostly exposed to financialization except for natural gas. Furthermore, we confirm that gold and silver remained intact to financialization as described by Nguyen et al. (2020). The results are robust to the omission of financial and macroeconomic variables and the modification of the prior variance parameters.

Second, I modify the dynamic factor model and enable the regional factors to interact in a vector autoregressive model that enables the volatility to enter the mean equation and the regional variables to have a feedback effect. The energy, precious metal, real economic activity, monetary policy, and financial and macroeconomic uncertainties are estimated within this model. In a one-step estimation, we calculate the energy uncertainty and its impact on the U.S. economy. This methodology overcomes the drawbacks of the two-step estimations of uncertainty time series and their impacts: inconsistencies of the two models and omitted variable bias. Findings indicate that a one-standard-deviation increase of commodity energy uncertainty homogeneously increases all energy prices except for natural gas futures. The shock

leads to a countercyclical response of macroeconomic variables: increasing inflation indicators, unemployment and decreasing industrial production, capacity utilization and manufacturing. The effect of the shock; however, wears off within a year.

An interesting avenue for future research involves estimation of uncertainty by a two-level dimension reduction. Reducing the input data into 10–15 variables in the factor level and then in the VAR model applying a factor volatility structure to further reduce the number of time series. This estimation would enable a rich interaction within the groups of time series we applied in the estimation. For instance, an interaction between inflation, production and unemployment. Furthermore, it would facilitate the interpretation of the regional factors.

References

- Aastveit, K. A., Natvik, G. J., and Sola, S. (2017). Economic uncertainty and the influence of monetary policy. *Journal of International Money and Finance*, 76:50–67.
- Adams, Z., Collot, S., and Kartsakli, M. (2020). Have commodities become a financial asset? evidence from ten years of financialization. *Energy Economics*, 89:104769.
- Adams, Z. and Glück, T. (2015). Financialization in commodity markets: a passing trend or the new normal? *Journal of Banking & Finance*, 60:93–111.
- Alessandri, P. and Mumtaz, H. (2019). Financial regimes and uncertainty shocks. *Journal of Monetary Economics*, 101:31–46.
- Alquist, R., Bhattarai, S., and Coibion, O. (2020). Commodity-price comovement and global economic activity. *Journal of Monetary Economics*, 112:41–56.
- Apel, M. and Grimaldi, M. B. (2014). How informative are central bank minutes? *Review of Economics*, 65(1):53–76.
- Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.
- Balli, F., Naeem, M. A., Shahzad, S. J. H., and de Bruin, A. (2019). Spillover network of commodity uncertainties. *Energy Economics*, 81:914–927.
- Bañbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of applied Econometrics*, 25(1):71–92.
- Baumeister, C. and Hamilton, J. D. (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review*, 109(5):1873–1910.
- Baumeister, C., Korobilis, D., and Lee, T. K. (2022). Energy markets and global economic conditions. *Review of Economics and Statistics*, 104(4):828–844.
- Bessembinder, H. (2018). The “roll yield” myth. *Financial Analysts Journal*, 74(2):41–53.
- Bessembinder, H. and Seguin, P. J. (1992). Futures-trading activity and stock price volatility. *the Journal of Finance*, 47(5):2015–2034.

- Bianchi, R. J., Fan, J. H., and Todorova, N. (2020). Financialization and de-financialization of commodity futures: A quantile regression approach. *International Review of Financial Analysis*, 68:101451.
- Blake, A. P. and Mumtaz, H. (2012). Applied bayesian econometrics for central bankers. *Technical Books*.
- Brunetti, C. and Buyuksahin, B. (2009). Is speculation destabilizing? *Available at SSRN 1393524*.
- Büyüksahin, B. and Robe, M. A. (2014). Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, 42:38–70.
- Caldara, D. and Herbst, E. (2019). Monetary policy, real activity, and credit spreads: Evidence from bayesian proxy svrs. *American Economic Journal: Macroeconomics*, 11(1):157–192.
- Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4):1194–1225.
- Carlstrom, C. T. (1993). Assessing real interest rates. *Economic Commentary*, (8/15/1993).
- Carriero, A., Clark, T. E., and Marcellino, M. (2018). Measuring uncertainty and its impact on the economy. *Review of Economics and Statistics*, 100(5):799–815.
- Carter, C. K. and Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika*, 81(3):541–553.
- Cascaldi-Garcia, D., Sarisoy, C., Londono, J. M., Sun, B., Datta, D. D., Ferreira, T., Grishchenko, O., Jahan-Parvar, M. R., Loria, F., Ma, S., et al. (2023). What is certain about uncertainty? *Journal of Economic Literature*, 61(2):624–654.
- Case, K. E. and Shiller, R. J. (1987). Prices of single family homes since 1970: New indexes for four cities.
- Chan, J. C. (2017). The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Business & Economic Statistics*, 35(1):17–28.
- Chan, J. C., Koop, G., and Yu, X. (2023). Large order-invariant bayesian vars with stochastic volatility. *Journal of Business & Economic Statistics*, pages 1–13.

- Chang, Y., Herrera, A. M., and Pesavento, E. (2023). Oil prices uncertainty, endogenous regime switching, and inflation anchoring. *Journal of Applied Econometrics*, 38(6):820–839.
- Chen, L., Verousis, T., Wang, K., and Zhou, Z. (2023). Financial stress and commodity price volatility. *Energy Economics*, 125:106874.
- Chen, Y., Dai, W., and Sorescu, S. M. (2021). A hiding place? diversification, financialization, and return comovement in commodity markets. *Diversification, Financialization, and Return Comovement in Commodity Markets (July 19, 2021)*.
- Cheng, I.-H., Kirilenko, A., and Xiong, W. (2015). Convective risk flows in commodity futures markets. *Review of Finance*, 19(5):1733–1781.
- Cheng, I.-H. and Xiong, W. (2014). Financialization of commodity markets. *Annu. Rev. Financ. Econ.*, 6(1):419–441.
- Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.
- Creal, D. D. and Wu, J. C. (2017). Monetary policy uncertainty and economic fluctuations. *International Economic Review*, 58(4):1317–1354.
- Da, Z., Tang, K., Tao, Y., and Yang, L. (2023). Financialization and commodity markets serial dependence. *Management Science*.
- Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *FRB of New York Staff Report*, (326).
- Demirer, R. and Kutan, A. M. (2010). The behavior of crude oil spot and futures prices around opec and spr announcements: An event study perspective. *Energy Economics*, 32(6):1467–1476.
- Domanski, D. and Heath, A. (2007). Financial investors and commodity markets. *BIS quarterly review*, March.
- Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

- Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American economic review*, 102(4):1692–1720.
- Gorton, G. and Rouwenhorst, K. G. (2006). Facts and fantasies about commodity futures. *Financial Analysts Journal*, 62(2):47–68.
- Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, 61(2):139–147.
- Hakkio, C. S., Keeton, W. R., et al. (2009). Financial stress: What is it, how can it be measured, and why does it matter. *Economic Review*, 94(2):5–50.
- Henderson, B. J., Pearson, N. D., and Wang, L. (2015). New evidence on the financialization of commodity markets. *The Review of Financial Studies*, 28(5):1285–1311.
- Holmes, P. and Rougier, J. (2005). Trading volume and contract rollover in futures contracts. *Journal of Empirical Finance*, 12(2):317–338.
- Irwin, S. H. and Sanders, D. R. (2012). Testing the masters hypothesis in commodity futures markets. *Energy economics*, 34(1):256–269.
- Jacquier, E., Polson, N. G., and Rossi, P. E. (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20(1):69–87.
- Ji, Q., Zhang, D., and Zhao, Y. (2020). Searching for safe-haven assets during the covid-19 pandemic. *International Review of Financial Analysis*, 71:101526.
- Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.
- Känzig, D. R. (2021). The macroeconomic effects of oil supply news: Evidence from opec announcements. *American Economic Review*, 111(4):1092–1125.
- Keynes, J. M. (1930). Treatise on money: Pure theory of money vol. i.
- Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American economic review*, 99(3):1053–1069.
- Kim, C.-J. and Nelson, C. R. (1999). State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books*, 1.

- Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147.
- Koopman, S. J. and Hol Uspensky, E. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets. *Journal of applied Econometrics*, 17(6):667–689.
- Kose, M. A., Otrok, C., and Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *American economic review*, 93(4):1216–1239.
- Le Pen, Y. and Sévi, B. (2018). Futures trading and the excess co-movement of commodity prices. *Review of Finance*, 22(1):381–418.
- Leduc, S. and Liu, Z. (2016). Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics*, 82:20–35.
- Lindsten, F., Jordan, M. I., and Schon, T. B. (2014). Particle gibbs with ancestor sampling. *Journal of Machine Learning Research*, 15:2145–2184.
- Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and business cycles: exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13(4):369–410.
- Ma, Y.-R., Ji, Q., Wu, F., and Pan, J. (2021). Financialization, idiosyncratic information and commodity co-movements. *Energy Economics*, 94:105083.
- Masters, M. W. (2008). Testimony before the committee on homeland security and governmental affairs. *US Senate, Washington, May, 20*.
- McCracken, M. W. and Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.
- Miranda-Agrippino, S. and Ricco, G. (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics*, 13(3):74–107.
- Moench, E., Ng, S., and Potter, S. (2013). Dynamic hierarchical factor models. *Review of Economics and Statistics*, 95(5):1811–1817.
- Mumtaz, H. and Musso, A. (2021). The evolving impact of global, region-specific, and country-specific uncertainty. *Journal of Business & Economic Statistics*, 39(2):466–481.

- Mumtaz, H. and Surico, P. (2018). Policy uncertainty and aggregate fluctuations. *Journal of Applied Econometrics*, 33(3):319–331.
- Mumtaz, H. and Theodoridis, K. (2020). Dynamic effects of monetary policy shocks on macroeconomic volatility. *Journal of Monetary Economics*, 114:262–282.
- Mumtaz, H. and Zanetti, F. (2013). The impact of the volatility of monetary policy shocks. *Journal of Money, Credit and Banking*, 45(4):535–558.
- Nguyen, D. K., Sensoy, A., Sousa, R. M., and Uddin, G. S. (2020). Us equity and commodity futures markets: Hedging or financialization? *Energy Economics*, 86:104660.
- Ordu, B. M., Oran, A., and Soytas, U. (2018). Is food financialized? yes, but only when liquidity is abundant. *Journal of Banking & Finance*, 95:82–96.
- Otrok, C. and Whiteman, C. H. (1998). Bayesian leading indicators: measuring and predicting economic conditions in iowa. *International Economic Review*, pages 997–1014.
- Park, J. and Ratti, R. A. (2008). Oil price shocks and stock markets in the us and 13 european countries. *Energy economics*, 30(5):2587–2608.
- Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852.
- Rossi, B. and Sekhposyan, T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *American Economic Review*, 105(5):650–655.
- Silvennoinen, A. and Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, 24:42–65.
- Smith, A. F. and Roberts, G. O. (1993). Bayesian computation via the gibbs sampler and related markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(1):3–23.
- Stoll, H. R. and Whaley, R. (2010). Commodity index investing and commodity futures prices. *Journal of Applied Finance (Formerly Financial Practice and Education)*, 20(1).
- Tang, K. and Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(6):54–74.

- Villani, M. (2009). Steady-state priors for vector autoregressions. *Journal of Applied Econometrics*, 24(4):630–650.
- Wang, S. L. and McPhail, L. (2014). Impacts of energy shocks on us agricultural productivity growth and commodity prices—a structural var analysis. *Energy Economics*, 46:435–444.
- Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

A Appendix

A.1 Condition number of a matrix

The condition number of a matrix measures among others the sensitivity in a linear system. For instance, it measures how much the solution of the $Ax = b$ solution varies for small perturbations to x . Due to the floating point number representation of real numbers such small changes may occur unintentionally in the calculation.

The Δ -perturbed linear system is:

$$A(x + \Delta x) = b + \Delta b$$

The main focus of interest is to determine a lower and an upper bound of the relative errors:

$$\frac{\|\Delta x\|}{\|x\|}, \quad \frac{\|\Delta b\|}{\|b\|} \quad (31)$$

We separate the perturbed linear system into two equations, take the norm and transform it by applying the sub-multiplicativity³⁶ of the matrix norm:

$$\begin{aligned} b &= Ax & \Delta b &= A\Delta x \\ \|b\| &= \|Ax\| \leq \|A\| \|x\| & \|\Delta b\| &= \|A\Delta x\| \leq \|A\| \|\Delta x\| \\ \frac{\|b\|}{\|Ax\|} &\leq \|x\| & \frac{\|\Delta b\|}{\|A\Delta x\|} &\leq \|\Delta x\| \end{aligned}$$

Similarly, applying the matrix sub-multiplicativity to the equation $A^{-1}b = x$, we get the following two inequalities:

$$\begin{aligned} x &= A^{-1}b & \Delta x &= A^{-1}\Delta b \\ \|x\| &= \|A^{-1}b\| & \leq \|A^{-1}\| \|b\| & \|\Delta x\| &= \|A^{-1}\Delta b\| \leq \|A^{-1}\| \|\Delta b\| \end{aligned}$$

Combining these two results leads us to an estimate of the upper bound of the relative error

³⁶If an induced norm is used we can also rely on the definition of the norm induced by a vector norm: $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v}$.

of x computed by increasing the numerator and decreasing the denominator of the fraction:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \|\Delta b\|}{\frac{\|b\|}{\|A\|}} = \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}$$

The lower bound is computed reversely by decreasing the numerator and increasing the denominator:

$$\frac{\|\Delta x\|}{\|x\|} \geq \frac{\frac{\|\Delta b\|}{\|A\|}}{\|A^{-1}\| \|b\|} = \frac{1}{\|A\| \|A^{-1}\|} \frac{\|\Delta b\|}{\|b\|}$$

The condition number of a matrix is defined as

$$\kappa(A) := \|A\| \|A^{-1}\| \quad (32)$$

The interpretation of the condition number is intuitive if the following equation is considered:

$$\kappa(A)^{-1} \delta_b \leq \delta_x \leq \kappa(A) \delta_b$$

where δ_x, δ_b denote the relative errors of x and b respectively. From the inequality, we can infer that the condition number shows that if the left side of the equation, b changes by Δ , then the solution x changes at least by $1/\kappa(A)$ times the change and at most by $\kappa(A)$.

In this application if the covariance matrix has a condition number:

$$\frac{1}{\kappa(A)} < 10^{-13} \quad (33)$$

I substitute the covariance matrix by the scale matrices: $A^T A$.

A.2 Financialization indices in existing literature

This subsection illustrates the time series of financialization applied in the existing literature. First, the approach of Adams et al. (2020) that involves the LMG R^2 decomposition of a rolling window regression of each commodity returns on three financial and four macroeconomic variables. Second, the correlation between the S&P500 and the commodity futures returns. The correlation is also calculated in a rolling window. Lastly, the Commodity Futures Trading Commission Index Investment Data, which is only available for a subset of the analysed commodities.

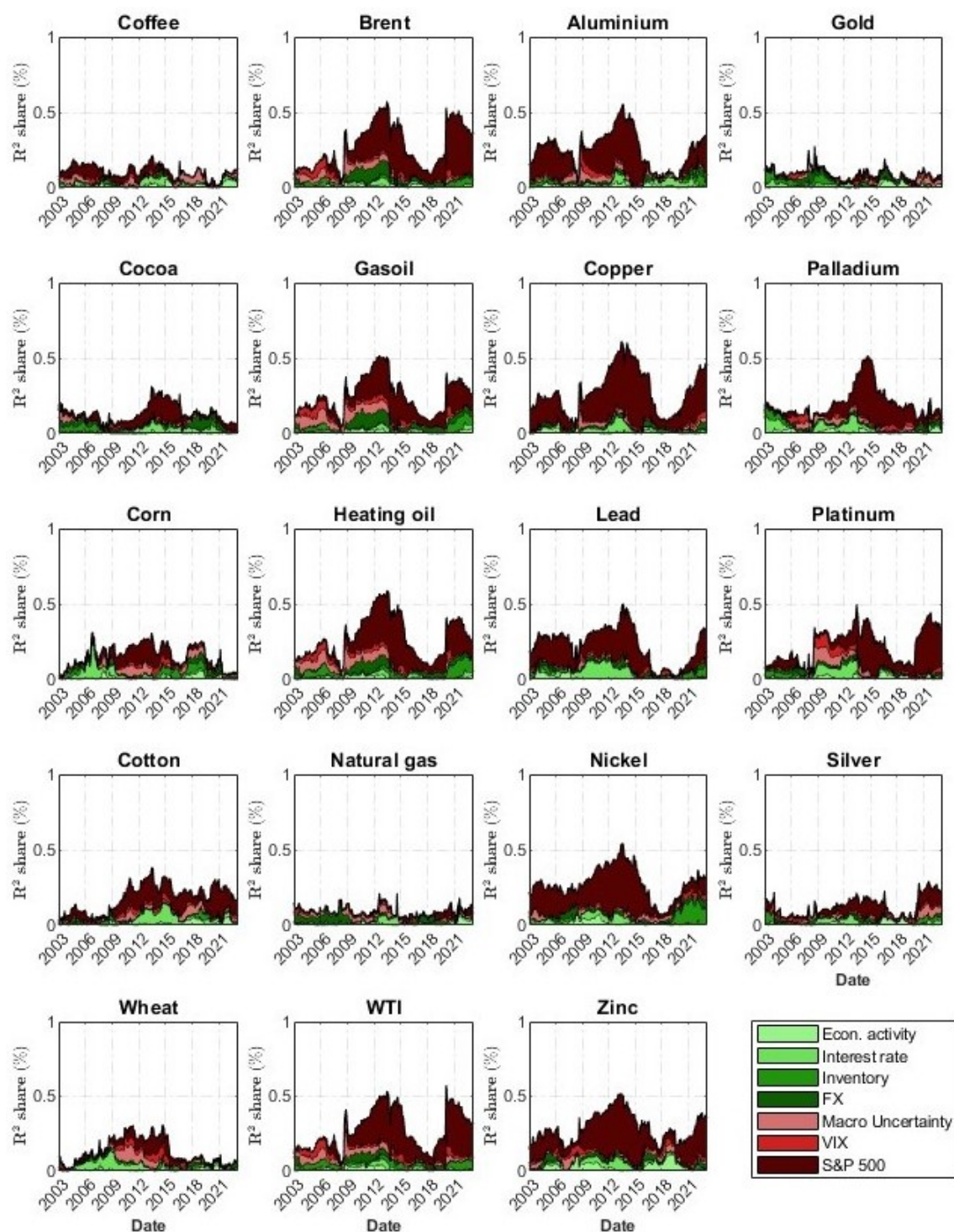


Figure A1: R^2 decomposition of commodity returns.

Note. The figure illustrates the LMG R^2 decomposition of the commodities in table 3. The regressions are estimated in a 60-day rolling window as described by Adams et al. (2020). The macroeconomic regressors are the Kilian index (Kilian, 2009) (economic activity), the difference between the 3-month Treasury Bill rate and the change of the YoY CPI (interest rate), the first difference of the oil inventory reported by the U.S. Energy Information Administration (inventory) and first difference of the trade-weighted US Dollar Index. The financial regressors are the macroeconomic uncertainty index of Jurado et al. (2015), the VIX and the first difference of the S&P500. The sample period starts in February 1998 and ends in April 2023.

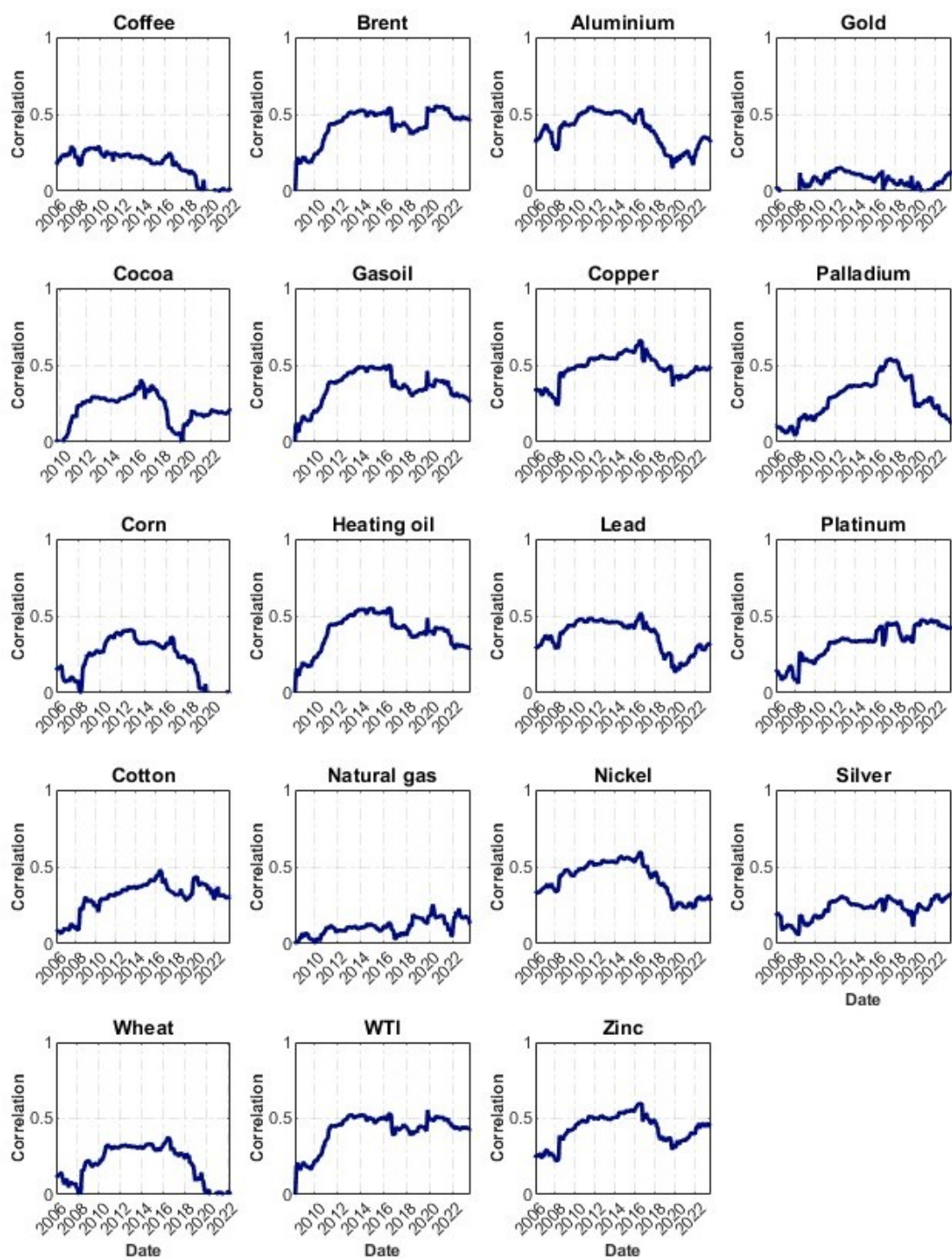


Figure A2: Co-movement between equity and commodity futures.

Note. The figure illustrates the Pearson correlation between the monthly returns of the S&P500 and the commodity futures described in table 3. The correlation is measured in a 100-observations rolling window. The sample period starts in February 1998 and ends in December 2023.

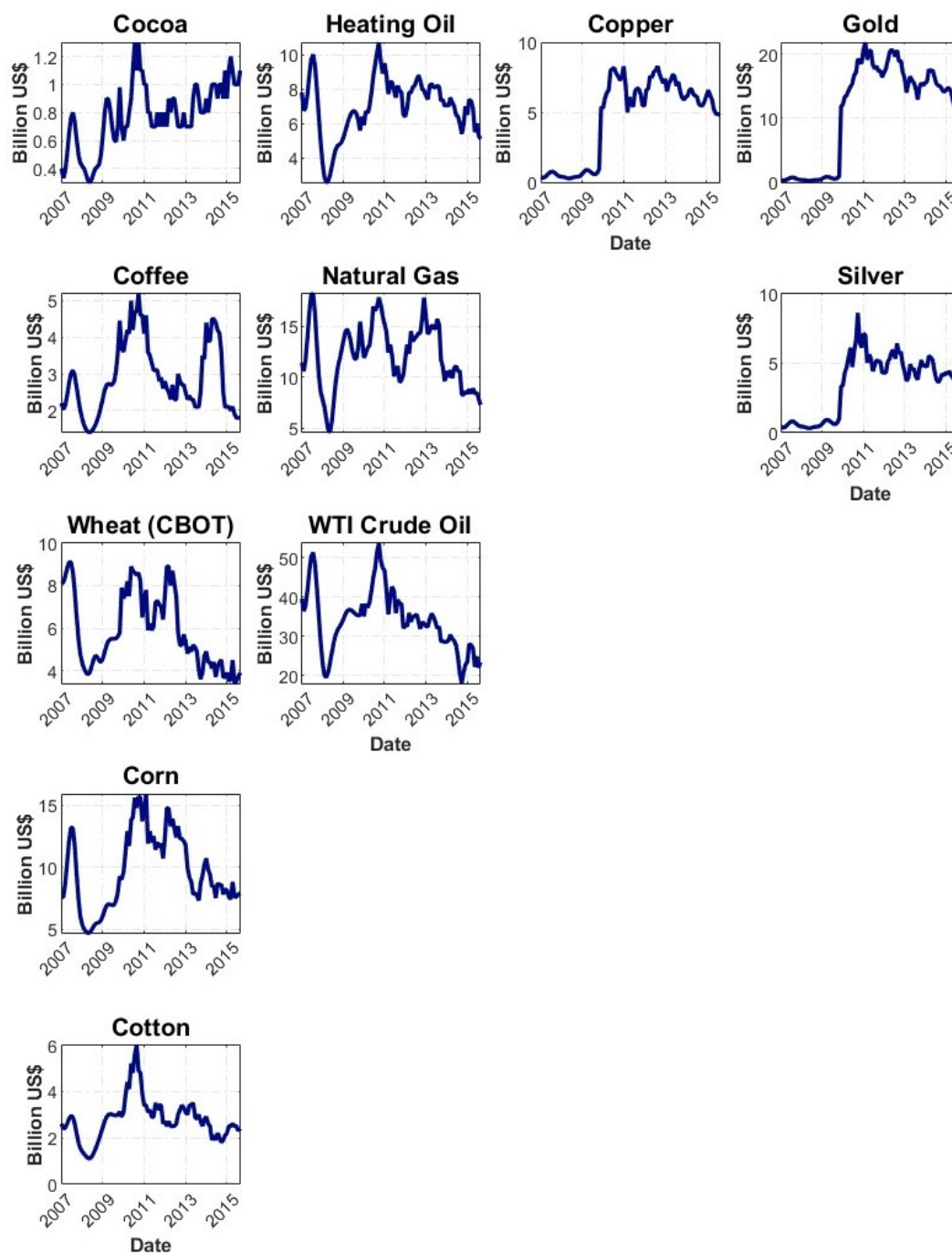


Figure A3: Index Investment Data.

Note. The figure illustrates the net long notional value of Index Investment Data provided by the Committee of Futures Trading Commission. The CFTC originally sent a special call to 43 entities to report the total notional value of index positions. The institutions include 16 swap dealers that were known to the CFTC staff to have large positions in commodity index investment, 14 commodity index funds and additionally included index funds, pension funds, mutual funds and hedge funds. The sample period starts in December 2007 and ends in November 2015.

Commodity futures	R^2 decomposition	Eq., cmdty co-movement	Index investment
Coffee	0.261	0.795	−0.598
Cocoa	0.722	0.704	−0.146
Corn	0.695	0.774	0.134
Cotton	0.537	0.161	−0.535
Corn	0.800	0.773	−0.161
Brent	0.635	−0.099	0
Gasoil	0.605	0.454	0
Heating oil	0.711	0.657	−0.030
Natural gas	0.325	−0.637	0.075
WTI	0.736	0.084	0.544
Aluminium	0.758	0.840	0
Copper	0.936	0.897	−0.496
Lead	0.603	0.638	0
Nickel	0.827	0.828	0
Zinc	0.346	0.639	0
Gold	0.522	0.680	−0.693
Palladium	0.850	0.652	0
Platinum	0.892	−0.216	0
Silver	0.724	0.396	−0.720

Table A1: Individual correlation between financialization time series.

Note. The table reports the Pearson correlation between the contribution of financial variables to the variance decomposition and three other proxies of financialization. The R^2 decomposition is the financialization index of Adams et al. (2020). The equity, commodity co-movement the correlation approach of Nguyen et al. (2020); Ordu et al. (2018); Tang and Xiong (2012) among others and the last row is the Index Investment Data (IID) provided by the CFTC and used by Irwin and Sanders (2012). The 0 correlations in the third column indicate that for the given commodity no CFTC IID data is available. The sample period starts in December 2007 and ends in November 2015.

A.3 Impulse-response functions to energy commodity uncertainty shock

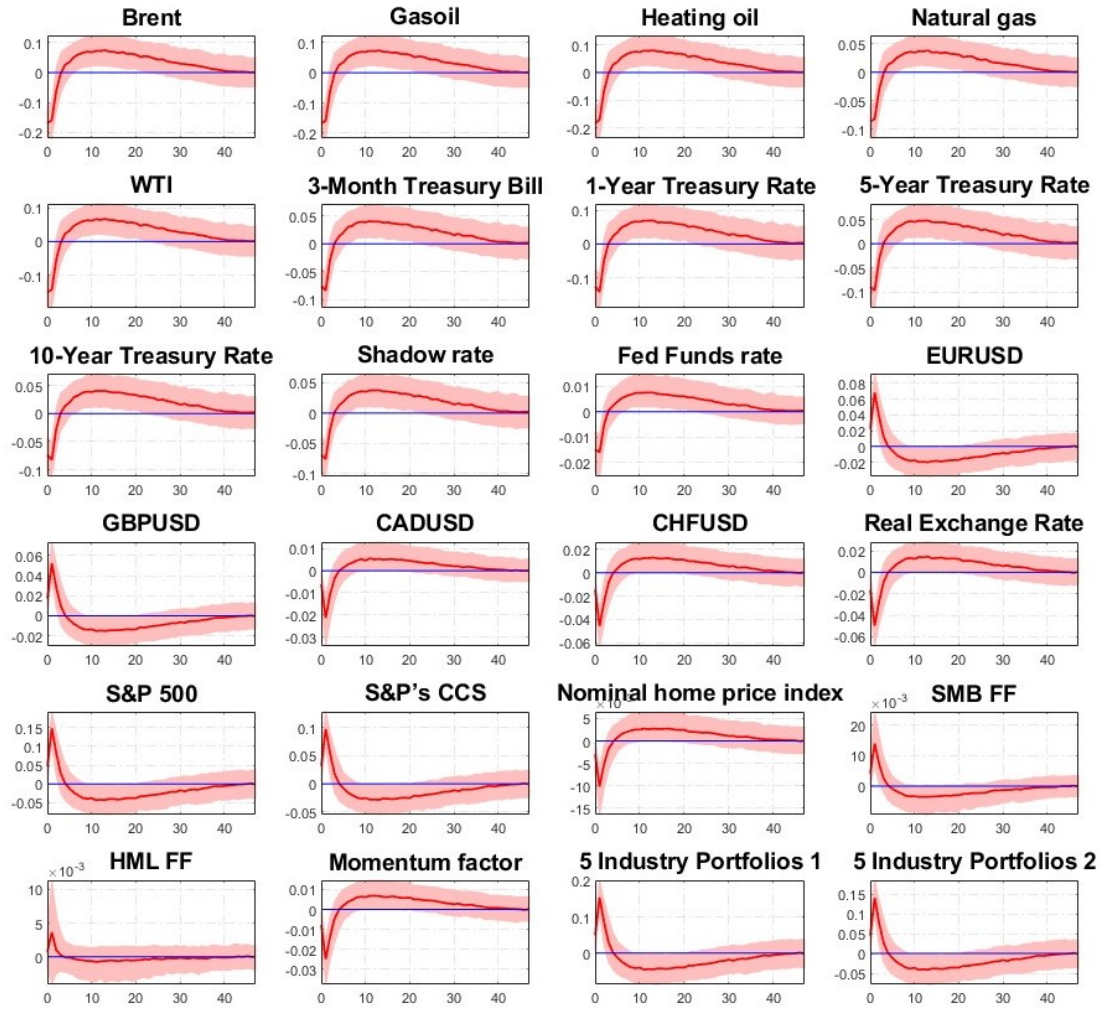


Figure A4: Impulse response functions of energy commodity futures and financial variables.

Note. The figure illustrates the 48 months ahead average impulse response function of the energy and financial variables to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.

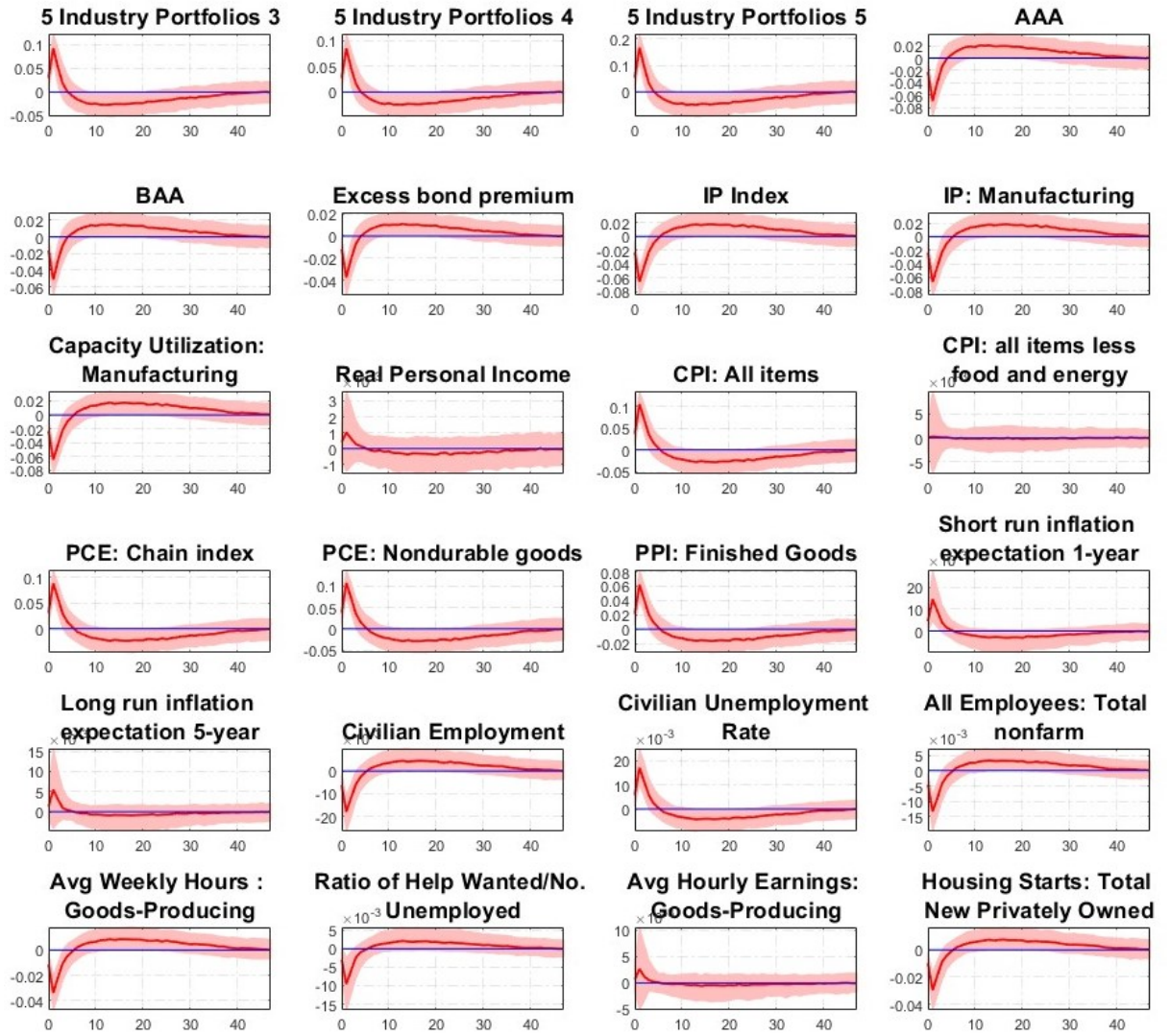


Figure A5: Impulse response functions of financial and macroeconomic variables.

Note. The figure illustrates the 48 months ahead average impulse response function of the macroeconomic and financial variables to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.

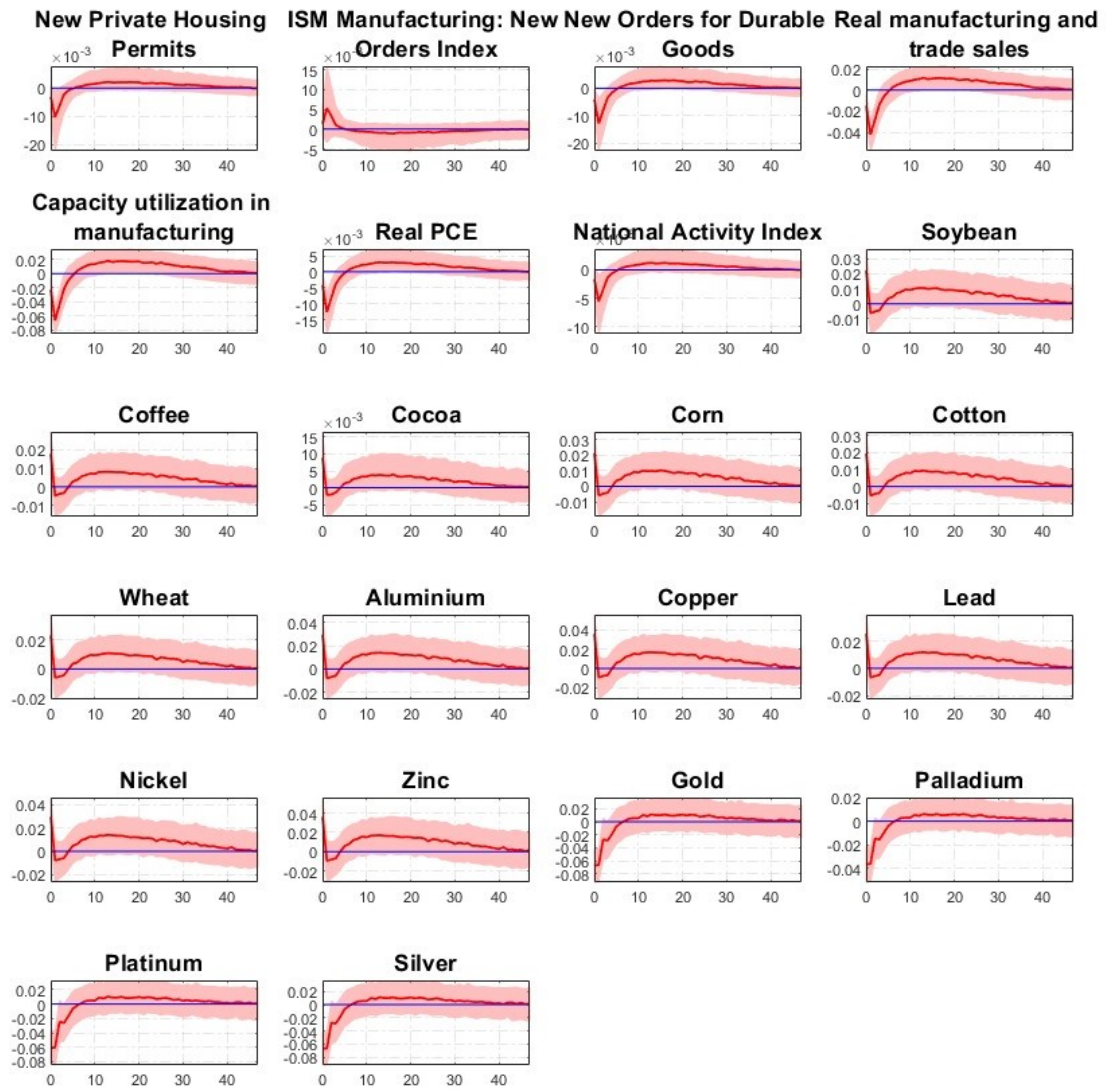


Figure A6: Impulse response functions of macroeconomic variables and commodity futures.

Note. The figure illustrates the 48 months ahead average impulse response function of the macroeconomic, commodity real economic activity and precious metals to a one standard deviation shock of the commodity uncertainty. The solid red line denotes the 68% confidence interval. The impulse-response function is calculated with Monte Carlo integration using the 1 000 posterior draws. The sample period starts in May 2001 and ends in December 2023.