JOHANNES KEPLER UNIVERSITY LINZ

DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Hedging with commodity futures and the end of normal Backwardation

by

Jochen GÜNTNER

Benjamin KARNER

Working Paper No. 2021

November 2020

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

Corresponding author: jochen.guentner@jku.at

Hedging with commodity futures and the end of normal backwardation

Jochen Güntner* Benjamin Karner

JKU Linz *JKU Linz*

November 2, 2020

Abstract

Using the S&P GSCI and its five component sub-indices, we show that considering each commodity separately yields nontrivial hedging gains in and out of sample. During 1999–2019, the maximum Sharpe ratio portfolio assigns positive weights to the GSCI Energy, Industrial and Precious Metals, whereas only precious metals enter the optimal portfolio after the financial crisis. In out-of-sample optimizations based on dynamic conditional correlations, a subset of commodity futures excluding the GSCI Agriculture and Livestock outperforms conventional stock-bond portfolios with and without the overall GSCI. We argue that the "normal backwardation" in commodity markets has broken down during our sample period.

Keywords: Commodity futures, Diversification, Hedging, Financial crisis, Normal backwardation

JEL classification: C58, G11, G17, Q02

*Corresponding author: Department of Economics, Johannes Kepler University Linz. Altenberger Straße 69, 4040 Linz, Austria. Telephone: (+43)(0)732 2468 7360. Email: jochen.guentner@jku.at.

1 Introduction

At least since Markowitz (1952), it is well known that the correlation between candidate assets is crucial for effective portfolio diversification. The lower the correlation between an asset and all other assets in a portfolio, the larger is the potential reduction in the portfolio's risk from including this asset. In this paper, we investigate the scope of commodity futures in portfolio diversification and optimization, which hinges on the cross-correlations between their returns and the returns on traditional assets, such as stocks and bonds. Given that the correlations of the returns on stocks, bonds, and commodity futures are likely time-varying, the weight of the latter in the optimal portfolios might also change over time.

Using Standard&Poor's global commodity investment benchmark (formerly the *Goldman Sachs Commodity Index*) GSCI and its five component sub-indices, we show that considering the underlying commodity futures separately yields nontrivial hedging gains over a conventional stock-bond portfolio both in and out of sample. Throughout our sample period from January 1999 to December 2019, the maximum Sharpe ratio portfolio assigns substantial weight to the GSCI Precious Metals and, to a lesser extent, the GSCI Energy and Industrial Metals, whereas only precious metals are included in the optimal portfolio after the financial crisis of 2007–2008. In out-of-sample portfolio optimizations based on a dynamic conditional correlations (DCC) model, a subset of commodity futures excluding the GSCI Agriculture and Livestock outperforms conventional stock-bond portfolios as well as a portfolio considering the overall GSCI. We argue that the theory of "normal backwardation" proposed by Keynes (1930) and Hicks (1939), which implies that there are more participants with hedging demand than speculators willing to supply these hedging services and thus positive risk premia for buyers in commodity futures markets, ceased to hold for all GSCI sub-indices either prior to or during the financial crisis. This is likely due to the increased "financialization" of commodity markets after 2000.

We are not the first to consider the hedging gains from investing in commodity futures. Early research by Bodie and Rosansky (1980), Bodie (1983), and Edwards and Park (1996), for example, has illustrated the suitability of commodity futures as an inflation hedge.[1]

[1]The inflation-hedge property is not surprising, given that commodity futures are a bet on future spot prices, which are directly related to components in the basket of commodities underlying the consumer or

Jensen et al. (2000) report dramatic improvements in the return and risk performance of a portfolio from including commodity futures during their sample period from January 1973 to December 1997.[2] Similarly, Gorton and Rouwenhorst (2006) find that stocks and commodity futures obtained similar average returns during 1959–2004, while futures exhibited slightly lower volatility. More recent studies argue that commodities have been increasingly integrated in global financial markets after 2000, as commodity futures indices and exchange traded funds (ETFs) tracking these indices reduced the cost of entry. This development is often referred to as the "financialization" of commodities and suspected to increase the correlation of asset returns across markets, thus making commodity futures less suitable for diversification (see, e.g., Domanski and Heath, 2007; Tang and Xiong, 2012; Silvennoinen and Thorp, 2013). In contrast, other studies report no change or a reduction in the correlation of commodity returns with stock and bond returns (see, e.g., Büyüksahin et al., 2009; Chong and Miffre, 2010).[3] In light of mixed evidence and given that all of the above studies use data from before the financial crisis of 2007–2008, the recent developments in correlations across financial markets and the hedging benefits from investing in commodity futures remain an open empirical question.

Besides this potential structural break around the financial crisis, futures contracts for different commodities are often considered as a homogeneous class of assets. When commodity futures differ in their risk and return properties, however, in particular in their respective correlations with stocks and bonds, this may lead to biased conclusions about the scope for diversification. Jaiswal and Uchil (2018), for example, find evidence in favor of the common wisdom that precious metals are a safe haven in times of financial-market turmoil and exhibit thus widely different return dynamics than energy commodities or industrial metals, the prices of which tend to be driven by (expectations about) global aggregate demand.

In order to illustrate the importance of distinguishing between different commodities, consider Figure 1, which depicts the development of the GSCI Industrial Metals, Precious

producer price index used to calculate rates of inflation (Gorton and Rouwenhorst, 2006).

[2]Considering commodity futures increases the portfolio's average annual return by six percentage points relative to a portfolio with stocks and bonds only, while keeping the standard deviation constant at 4%.

[3]Büyüksahin et al. (2009), for example, conclude that the correlation of commodities with stocks and bonds has not changed significantly during 1991–2007, despite increased interest in commodity futures as an investment vehicle.

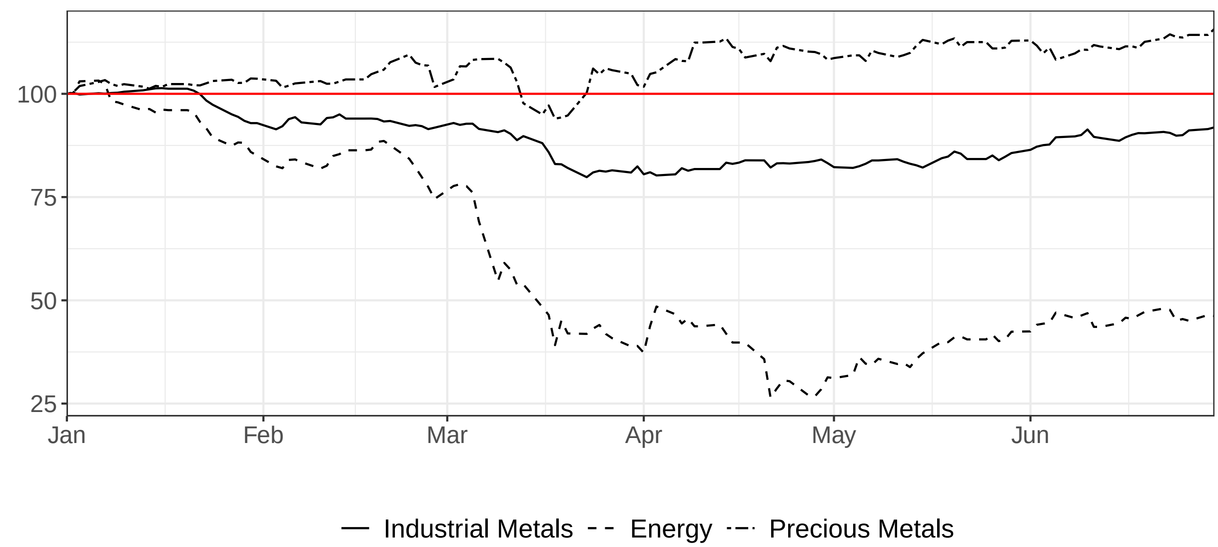

Figure 1: Index of cumulated percentage changes during the COVID-19 pandemic (normalized to 100 on January 1, 2020)

Metals, and Energy, respectively, where all three sub-indices are normalized to 100 in January 2020. During the ongoing COVID-19 pandemic, the cumulated returns on the three component sub-indices clearly diverge. While precious metals futures had realized cumulated returns of about 20% by the end of June, industrial metals and energy futures lost about 10% and more than 50%, respectively, as the global economy faced strong headwind.

In light of this heterogeneity across different commodities, we investigate the hedging benefits from considering the overall GSCI as well as each of its five component sub-indices — Energy, Industrial Metals, Precious Metals, Agriculture, and Livestock — separately for inclusion in a conventional stock-bond portfolio. Given that the GSCI is weighted according to world production volumes, it assigns a weight of 62% to Energy in 2019 (see Table 1), which is not necessarily optimal from a hedging perspective. Hence, portfolio optimizations considering the GSCI sub-indices separately may well imply very different optimal allocations across the different commodities.

The rest of this paper is structured as follows. Section 2 discusses the implications of Keynes' theory of normal backwardation for commodity futures markets. Section 3 presents the data and summary statistics for stocks and bonds as well as the overall GSCI and its five component sub-indices. In Section 4, we conduct in-sample portfolio optimizations based on different loss functions both for the entire sample period and splitting the latter into two subsamples, prior to and after the financial crisis of 2007–2008, to allow for a structural break due to the "financialization" of commodity markets. In Section 5, we conduct real-time (pseudo) out-of-sample portfolio optimizations using a dynamic conditional correlations (DCC) model to predict the correlations between candidate assets over the investment horizon. Section 6 discusses possible explanations for our in-sample and (pseudo) out-of-sample results and provides evidence that Keynes' theory of normal backwardation has broken down during our sample period. Section 7 concludes. Our robustness checks are deferred to the appendix.

2 Commodity Futures and "Normal Backwardation"

In commodity markets, the spot price for immediate delivery coexists with several prices for delivery at different points in time in the future. Accordingly, a commodity futures contract specifies the purchase or sale of a certain quantity of a commodity at a given price at a future point in time. The value of a futures contract at maturity equals zero, as the spot and the futures price coincide. Futures contracts are settled each day, whereby the original contract is replaced by a new one at the current futures price, while the buyer and the seller exchange the difference between the old and the new futures price via transfers from their margin accounts.[4]

Futures markets provide a hedging opportunity for commodity producers. If the time to market of a commodity amounts to six months, for example, the producer may sell forward her output immediately with delivery in six months at the 6-month futures price. This allows locking in the difference between the 6-month futures price and the commodity's production costs as a risk-free profit, while avoiding the exposure to fluctuations in the future spot price. According to Keynes (1930), producers of seasonal crops with a time to market of up to one year, which are exposed to uncertainty regarding weather and future market conditions, for example, are willing to pay a premium of 10 percent per annum in order to lock in their profits in advance of production.[5]

[4] Margin accounts represent liquidity deposited with the broker by both parties that are sufficient to absorb daily fluctuations in futures prices (Black, 1976).

[5] Keynes (1930) considers this to be a modest estimate of hedging premia, which may be substantially

Note that spot and futures markets may be in two relative states. If a commodity's spot price is above its futures price, the market is said to be "in backwardation".[6] If the futures price exceeds the commodity's spot price instead, the market is said to be "in contango". This opens up the possibility for substantially different returns on commodity futures and the corresponding spot contracts. Given that the fair futures price already account for the expectations of market participants, anticipated price changes such as seasonal fluctuations, for example, do not result in gains or losses on commodity futures.

Keynes' theory of "normal backwardation" originates from the observation that there are more sellers of futures contracts willing to accept a price below the expected future spot price in exchange for insurance against unexpected price drops. This hedging premium attracts investors and speculators, which are willing to absorb the supply of futures contracts. Due to an excess demand for hedging by producers, which suppresses current futures prices below expected future spot prices, commodity markets are assumed to be in backwardation most of the time — hence the adjunct "normal" (Keynes, 1930). Gorton and Rouwenhorst (2006) find empirical evidence for Keynes' theory during 1959–2004, when the returns on commodity futures significantly exceeded the returns on commodity spot prices. The question is whether this regularity still holds, despite the recent "financialization" of commodity markets described in the introduction.

Finally, it is important to distinguish between backwardation as a market state and the theory of normal backwardation. Against the intuition, a commodity market may be in a state of contango and still be consistent with Keynes' theory. As an example, suppose that the spot price of crude oil is expected to increase by 10 percent in the upcoming six months and that the 6-month futures price of crude oil is 5 percent above the current spot price. In this case the market is currently "in contango", as the futures price exceeds the *current* spot price, while normal backwardation continues to hold, as the *expected* future spot price exceeds the corresponding futures price.

higher in less organized markets.

[6] To be precise, "(normal) backwardation" describes a state, where the *expected* spot price at maturity exceeds the current futures price of a commodity. Given that the expected future spot price is unobservable, the current spot price is generally assumed to be the best predictor under the efficient market hypothesis.

3 Data

In this paper, we use weekly and monthly data for January 1999 through December 2019. This 21-year sample period allows for robust estimation and comprises similar sub-samples before and after the financial crisis of 2007–2008.

3.1 Global stocks and bonds

As a proxy for a diversified global stock portfolio, we use the Morgan Stanley Capital International (MSCI) *All Country World Index*, which includes more than 2,700 stocks listed in 23 developed and 24 emerging markets. The weight of each stock in the index is determined by its market capitalization.

As a proxy for a globally representative government bond index with varying duration, we include the Bloomberg Barclays *Global Treasury Total Return Index Value Unhedged*[7], which comprises local currency government debt of 37 investment grade countries with 24 different currencies, including both developed and emerging markets.

As a proxy for the risk-free rate of return, we use the yield on 3-month US treasury bills.

3.2 The S&P GSCI

Standard&Poor's global commodity investment benchmark (formerly the *Goldman Sachs Commodity Index*) GSCI represents a world production-weighted index of 24 exchange-traded futures contracts (as of January 2020). The weight of each commodity is based on its average worldwide production over the past five years. Although there is no general restriction on the total number of futures contracts, certain eligibility criteria ensure that only commodity futures with sufficient liquidity and investability are included in the GSCI.

[7]Source: Bloomberg Index Services Limited. BLOOMBERG® is a trademark and service of Bloomberg Finance L.P. and its affiliates (collectively "Bloomberg"). BARCLAYS® is a trademark and service mark of Barclays Bank Plc (collectively with its affiliates, "Barclays"), used under license. Bloomberg or Bloomberg's licensors, including Barclays, own all proprietary rights in the Bloomberg Barclays Indices. Neither Bloomberg nor Barclays approves or endorses this material, or guarantees the accuracy or completeness of any information herein, or makes any warranty, express or implied, as to the results to be obtained therefrom and, to the maximum extent allowed by law, neither shall have any liability or responsibility for injury or damages arising in connection therewith.

Table 1: Composition of Standard&Poor's global commodity investment benchmark GSCI

Subindex	Weight	Commodities Included
Energy	62.63%	Crude Oil (and supporting contracts) and Natural Gas
Petroleum	59.52%	Crude Oil (and supporting contracts)
Non-Energy	37.37%	All commodities not included in Energy Sub-Index
Industrial Metals	11.16%	Aluminum, Copper, Lead, Nickel, and Zinc
Precious Metals	4.14%	Gold and Silver
Agriculture	15.41%	Wheat (Chi. & Kan.), Corn, Soybeans, Coffee, Sugar, Cocoa, and Cotton
Grains	11.42%	Wheat (Chi. & Kan.), Corn, and Soybeans
Livestock	6.65%	Lean Hogs, Live Cattle, and Feeder Cattle

Note: Weights based on the average contract reference prices for the 2019 annual calculation period

The overall index comprises five sub-indices — Energy, Industrial Metals, Precious Metals, Agriculture, and Livestock. Table 1 reports the weight of each sub-index in the overall GSCI based on their average contract reference prices for 2019. It is important to note the dominant weight on crude oil futures, which is a direct consequence of the production-based weighting scheme. Also note that the weight on precious metals is only 4.14%, although gold is considered as a suitable asset for portfolio diversification and a safe haven in turbulent times. These observations already raise the question, whether the GSCI's weighting scheme is optimal from a financial investor's point of view, and indicates potential scope for improvement by including the GSCI sub-indices *separately* in a hypothetical global portfolio.

The purpose of the index is to replicate the development of a portfolio of commodities. Accordingly, the corresponding futures are not held until maturity, as this would result in the physical delivery of the commodities. At the beginning of each month, futures nearing their expiration dates are instead sold and substituted by futures with the next closest expiration date. These trades are executed in the 5-day roll period between the fifth and ninth business day of each month.

In what follows, GSCI consistently refers to the *S&P GSCI Total Return* (TR) index. Given that the corresponding futures contracts are fully collateralized, this commodity index represents the closest substitute for an investment in stocks or bonds.[8] The value of the GSCI

[8] In this case, "fully collateralized" means that the amount outstanding of a hypothetical payment at

TR on each business day equals the product of three components, i.e.

1. the value of the index on the preceding business day,

2. the sum of the *Contract Daily Return* (CDR) and the *Treasury Bill Return* (TBR) on the business day plus 1,

3. the TBR on each non-business day since the last business day plus 1.

Accordingly, the formal representation of the value on day d is given by

$$\text{GSCI TR}_d = \text{GSCI TR}_{d-1} \cdot (1 + CDR_d + TBR_d) \cdot (1 + TBR_d)^{days}, \tag{1}$$

where CDR_d denotes the ratio of the *Total Dollar Weight Obtained* ($TDWO$) on any given business day divided by the *Total Dollar Weight Invested* ($TDWI$) on the preceding business day minus 1, i.e.

$$CDR_d = \frac{TDWO_d}{TDWI_{d-1}} - 1, \tag{2}$$

where $TDWO_d$ denotes the sum of the dollar weights of all futures contracts obtained from an investment in the index on the preceding business day and $TDWI_{d-1}$ the sum of the dollar weights on the preceding business day. The TBR on each calendar day d is given by

$$TBR_d = \left[\frac{1}{1 - \frac{91}{360} \cdot r_{d-1}}\right]^{\frac{1}{91}} - 1, \tag{3}$$

where r_{d-1} denotes the interest rate on the treasury bill *as of the most recent business day.*

3.3 Descriptive Statistics

Each panel of Figure 2 plots an index of cumulated percentage changes in the overall GSCI and its component sub-indices, respectively, that is normalized to 100 in January 1999. Note that the scale of the vertical axis varies between panels. Panels (a) and (b) illustrate the similarity between the GSCI and its energy sub-index, reflecting the dominant weight of the later in the overall index. It is therefore important to note that the GSCI Energy peaks at

maturity is invested in US treasury bills.

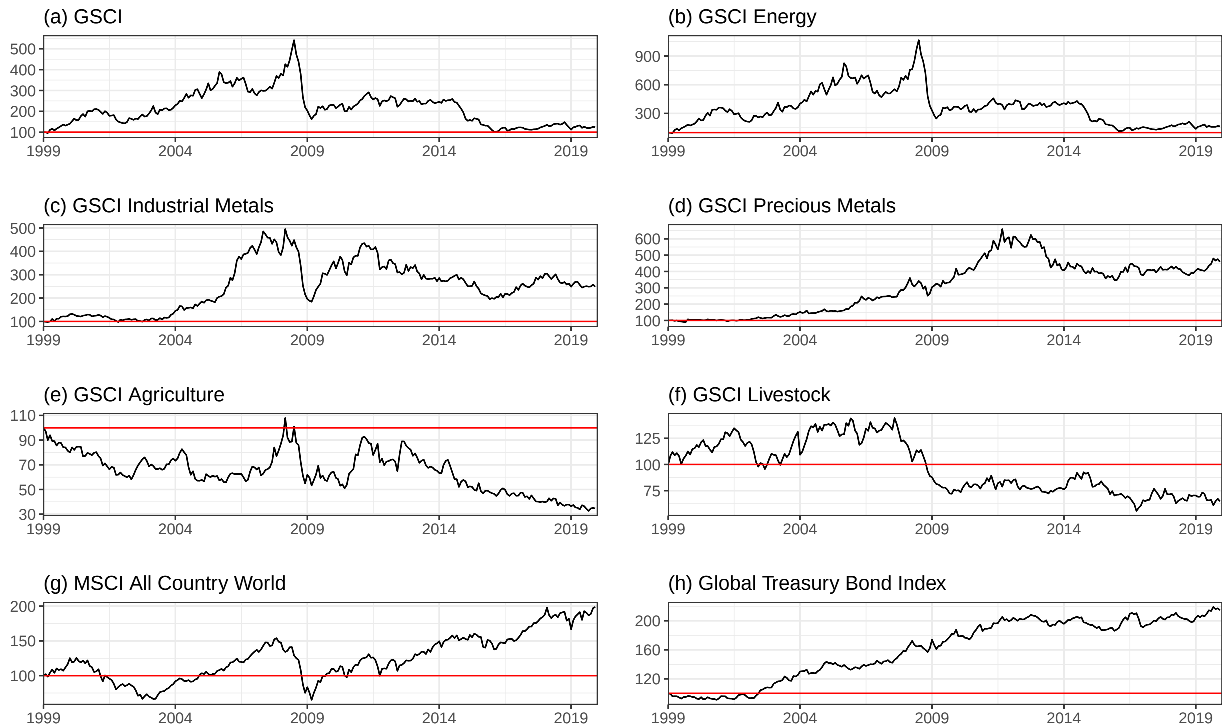

Figure 2: Index of cumulated percentage changes (normalized to 100 in January 1999)

an index value of 1,057% on the eve of the financial crisis of 2007–2008, while it virtually collapses during the crisis and performs poorly afterwards. Despite a slight recovery during 2009–2014, the price of crude oil futures contracts and thus the GSCI Energy declined again starting in 2014, due to an unexpected increase in global supply fueled by US shale oil producers and OPEC's reluctance to cut production as well as weaker-than-expected global economic growth and concerns about the Chinese economy on the demand side (see, e.g., Baumeister and Kilian, 2016; Ellwanger et al., 2017).

Similarly, the industrial and precious metals sub-indices realized substantial cumulative returns prior to the financial crisis, peaking at 495.2 and 659.6% of their 1999 values, respectively. At the same time, panels (c) and (d) illustrate the differences between the two commodity classes. While the prices of industrial metals futures surged prior to and suffered during the financial crisis, the prices of precious metals futures only started to pick up afterwards. The latter development is likely related to inflation concerns fueled by central banks' unconventional monetary policy actions and consistent with the widespread perception of

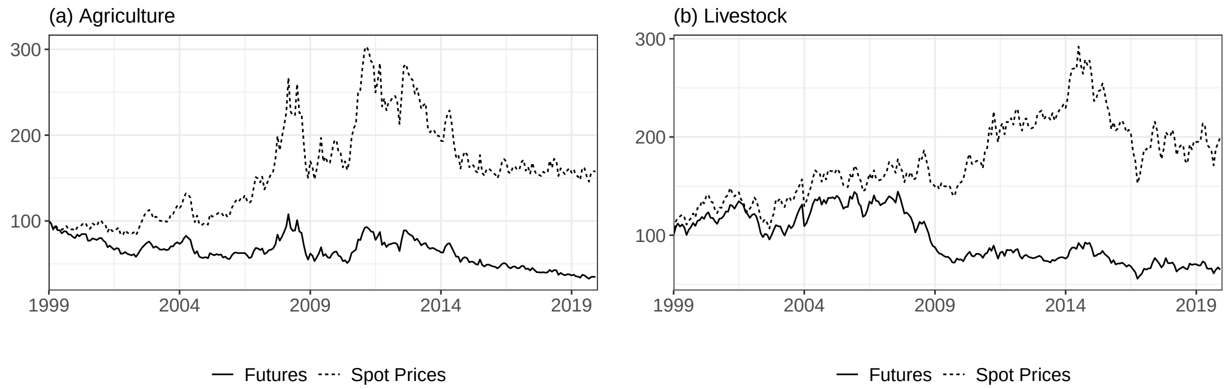

Figure 3: Spot versus futures price index for GSCI Agriculture and Livestock (normalized to 100 in January 1999)

precious metals, in particular gold, as an inflation hedge and a safe haven in times of financial market turmoil (Baur and McDermott, 2010).[9]

Regarding the cumulated changes in Agriculture and Livestock in panels (e) and (f), recall that the GSCI sub-indices are based on the prices of futures contracts rather than the prices of the underlying commodities. As outlined in the previous section, spot and futures prices may be disentangled for extended periods of time. While this will be discussed in greater detail in Section 6, Figure 3 illustrates that, during our sample period, the GSCI Agriculture and Livestock performed worse than the spot prices of the underlying commodities, where all are normalized to 100 in 1999. In any case, institutional investors as well as the providers of the GSCI index are bound to invest in commodities via futures, given that unprocessed orange juice and lean hog, for example, cannot be stored in large quantities over long periods and involve nontrivial inventory costs. Note also that a growing number of institutional investors refrains from investing in agriculture or livestock commodities for ethical reasons.

Panels (g) and (h) of Figure 2 plot the cumulated changes of the MSCI All Country World index and the Bloomberg Barclays Global Treasury Total Return Index Value Unhedged relative to their values in 1999. The MSCI World reflects the boom and bust of the dot-com bubble at the start of the sample period and the global stock market boom leading up to the financial crisis of 2007–2008. Due to their stable upward trend after 2009, stock investments

[9]There is an ongoing debate about the reasons for the decline in precious metals futures prices starting in 2013. Candidate explanations include dissolving inflation concerns, bullish stock markets, more careful monetary policy and forward guidance, as well as computerized trading.

nevertheless realized a decent average return during our sample period. Nevertheless, it was outperformed by the global bonds index over the past 21 years. With monetary policy rates close to the zero lower bound (ZLB) in many developed economies, investment-grade government bonds with longer maturities, which still yielded a decent nominal return, attracted large amounts of liquidity. It is important to note that, with policy rates remaining close to the ZLB, there is little room for similar price surges in the near future, consistent with a flattening of the index in panel (h) after 2013.

Table 2 reports the geometric mean and standard deviation of each asset's monthly return for 1999:1–2019:12 in percent. Consistent with the cumulated percentage changes in Figure 2, the GSCI Agriculture and Livestock realized negative returns on average over our sample period. Moreover, agricultural futures exhibited an exceptionally large volatility in returns. While futures contracts on precious metals and energy commodities yield the highest average returns, the standard deviation of the GSCI Energy exceeds that of the GSCI Precious Metals by about 80%. On average over our sample period, the overall GSCI performed worse than both the global stocks and bonds indices, also due to a negative average returns on two of its component sub-indices. In light of its competitive average return during our sample period, it seems in order to point out the low volatility of the global bonds index, which indicates that bonds are likely to carry a large weight in the optimal portfolios.

As pointed out beforehand, any hedging gains depend on the correlations between candidate assets. Table 3 therefore reports the contemporaneous correlations of monthly returns

Table 2: Summary statistics of monthly returns

Index	Geometric Mean (in %)	Standard Deviation (in %)
GSCI	0.085	6.650
GSCI Energy	0.201	9.146
GSCI Industrial Metals	0.362	6.099
GSCI Precious Metals	0.607	5.201
GSCI Agriculture	−0.420	6.038
GSCI Livestock	−0.172	4.281
MSCI Stocks	0.273	4.675
Global Bonds	0.303	1.904

Note: Geometric means and standard deviations of monthly returns for 1999:1–2019:12

Table 3: Contemporaneous correlations of monthly returns for 1999:1–2019:12

Index	GSCI	Energy	Industrial	Precious	Agriculture	Livestock	Stocks	Bonds
GSCI	1.000							
Energy	0.977	1.000						
Industrial	0.550	0.456	1.000					
Precious	0.288	0.219	0.349	1.000				
Agriculture	0.383	0.234	0.333	0.224	1.000			
Livestock	0.129	0.080	0.095	−0.064	−0.015	1.000		
Stocks	0.479	0.424	0.554	0.160	0.336	0.100	1.000	
Bonds	0.144	0.099	0.179	0.447	0.248	−0.164	0.194	1.000

on the GSCI and its component sub-indices with the global stocks and bonds indices. The lower the correlation of a given asset with all other candidate assets, the more suitable it is to reduce the standard deviation and thus the risk of the optimal portfolio. First, note that the large weight of the energy sub-index in the overall GSCI leads to an almost perfect correlation of 0.977, in line with the comovement of the corresponding time series in panels (a) and (b) of Figure 2. Moreover, the overall GSCI comoves strongly with its industrial metals sub-index and the global stocks index, where the contemporaneous correlation amounts to 0.550 and 0.479, respectively. Finally, the GSCI Industrial Metals correlates strongly with the global stocks index. Accordingly, these assets hardly qualify for portfolio diversification during our sample period.

In contrast, the global bonds index displays relatively low contemporaneous correlations with all asset classes but precious metals, which posses similar safe-haven properties and went through the financial crisis largely unscathed. Given that the contemporaneous correlation between bonds and livestock futures is negative and nontrivial, the two assets promise the largest hedging gains ex ante. Focusing only on the matrix of contemporaneous correlations, however, we ignores the negative average return on livestock futures during our sample period. While a negative correlation coefficient between two assets indicates possible hedging gains from including them in a portfolio, only two additional asset pairs display this property in Table 3, which both involve the GSCI Livestock as well as the GSCI Precious Metals and the GSCI Agriculture, respectively.

4 In-Sample Portfolio Optimization

As a starting point, we construct optimal in-sample portfolios that satisfy a selection of widely used objective functions. The aim is to show whether including (i) the overall GSCI and (ii) its five component sub-indices in the pool of candidate assets yields performance gains relative to the corresponding optimal portfolios with only stocks and bonds. In order to provide a comprehensive overview, we compare an equally weighted portfolio against a minimum risk efficient, a global minimum variance, and a maximum Sharpe ratio portfolio based on the respective return and standard deviation. Throughout, we exclude short positions, as investment trusts are generally not permitted to use short selling except for hedging purposes in order to neutralize long positions in the same asset.

4.1 Equally weighted portfolio

The equally weighted portfolio assigns constant and identical weights to all candidate assets. Accordingly, it represents a naïve benchmark rather than the outcome of an optimization procedure. When considering only the MSCI All Country World Index and the Bloomberg Barclays Global Treasury Total Return Index, the corresponding weights are one half each. When we also consider commodity futures, a naïve weight of one third is assigned to stocks and bonds as well as (i) the overall GSCI and (ii) its five component sub-indices, respectively. In the latter case, each component sub-index receives thus a weight of 1/15.

4.2 Minimum risk efficient portfolio

The minimum risk efficient portfolio minimizes the portfolio's variance subject to attaining a given target return — in our case the return of the equally weighted portfolio. When considering only stocks and bonds, this implies that the minimum risk efficient portfolio trivially replicates the equally weighted portfolio, which generally represents the unique only combination of stocks and bonds that yields exactly the target return. The constrained optimization problem for the minimum risk efficient portfolio can be formalized as

$$\min_{w_i} \quad \sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_i \sigma_j \rho_{ij} \quad \text{s.t.} \quad \sum_{i=1}^{n} w_i = 1, \quad 0 \leq w_i \leq 1, \quad r^* = \sum_{i=1}^{n} w_i r_i = r_{equal}, \tag{4}$$

where σ_p^2 denotes the portfolio's variance, w_i the weight of asset i, σ_i the standard deviation of asset i, ρ_{ij} the contemporaneous correlation between assets i and j, r^* the return of the minimum risk efficient portfolio, and r_{equal} the return of the equally weighted portfolio.[10]

4.3 Global minimum variance portfolio

The global minimum variance portfolio minimizes the portfolio variance without imposing a constraint on the portfolio's return. Accordingly, it is determined only by the standard deviations of and correlations between the candidate assets. The corresponding optimization problem resembles that of the minimum risk efficient portfolio, while we drop the constraint on the portfolio return, i.e.

$$\min_{w_i} \quad \sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_i \sigma_j \rho_{ij} \quad \text{s.t.} \quad \sum_{i=1}^{n} w_i = 1, \quad 0 \leq w_i \leq 1, \tag{5}$$

where all variables are as defined in (4).

4.4 Maximum Sharp ratio portfolio

Proposed by Sharpe (1994), the Sharpe ratio divides the excess return over the risk-free rate of return by the standard deviation of an asset i, i.e.

$$\text{Sharpe ratio}_i = \frac{r_i - r_f}{\sigma_i}, \tag{6}$$

where r_i and σ_i denotes the return and the standard deviation of the asset, respectively, and r_f the risk-free rate of return. The maximum Sharpe ratio portfolio maximizes the Sharpe ratio of the portfolio. Given that this occurs when the *capital allocation line* (CAL) is

[10]Note that the weights on assets in the portfolio sum to 1, i.e. the portfolio is fully invested, are bounded below by zero, which prevents short selling, and replicate the target return of the equally weighted portfolio.

tangential to the *efficient frontier*, i.e. the upward sloping part of the hyperbola of "efficient" portfolios, it is also known as the "tangency portfolio".

The CAL starts at the risk-free rate and its slope equals the ratio between the incremental increases in the portfolio's return and standard deviation, i.e. the Sharpe ratio. Accordingly, the Sharpe ratio is maximized, when the CAL is tangential to the efficient frontier:

$$\max_{w_i} \quad \text{Sharpe ratio}_p = \frac{r_p - r_f}{\sigma_p} \quad \text{s.t.} \quad \sum_{i=1}^{n} w_i = 1, \quad 0 \leq w_i \leq 1, \tag{7}$$

where all variables have been defined above. For the risk-free rate of return, r_f, we use the average monthly return on 3-month US treasury bills over the entire sample period.

4.5 Empirical results for 1999–2019

Table 4 reports the portfolio optimization results with only stocks and bonds. As mentioned above, the minimum risk efficient portfolio with two candidate assets trivially replicates the equally weighted portfolio and is therefore ignored.

The results show that the equally weighted portfolio attains the highest monthly return due to its comparatively high weight on stocks. At the same time, it features the highest volatility and thus the lowest Sharpe ratio in Table 4. The global minimum variance portfolio assigns a weight of 91.38% to bonds, which is similar to the optimal weight of 85.90% in the maximum Sharpe ratio portfolio. The results illustrate the gains from portfolio diversification, as the standard deviation of the global minimum variance and of the maximum Sharpe ratio portfolio is below that of bonds, although bond returns exhibit the lowest standard

Table 4: Optimal in-sample portfolios with stocks and bonds

Portfolio	Portfolio weight on		Monthly	Standard	Sharpe
	Bonds	Stocks	Return	Deviation	Ratio
Equally weighted	50.00	50.00	0.3539%	0.0269	0.0762
Global minimum variance	91.38	8.62	0.3280%	0.0186	0.0962
Maximum Sharpe ratio	85.90	14.10	0.3314%	0.0188	0.0971

Note: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2019.

deviation among the individual assets during our sample period. The results also confirm that bonds, especially those with higher maturity and thus longer duration, experienced a boom due to persistently low monetary policy rates.

Table 5 reports the optimal portfolios, when we consider the GSCI as a candidate asset. Note that we generally find no improvement over the portfolios with only stocks and bonds. During our sample period, the overall GSCI obtains zero weight in the minimum risk efficient and the maximum Sharpe ratio portfolio, while its optimal weight in the global minimum variance portfolio is a mere 1.5%. This seems at odds with results based on pre-2000 data, where Jensen et al. (2000), for example, find considerable improvements in an optimal portfolio's return and risk from considering commodity futures in addition to stocks and bonds. The seemingly contradictory findings in Table 5 suggest a structural break in commodity futures markets during our sample period.

The portfolio optimization in Table 6 replaces the GSCI by its five component sub-indices in order to qualify some of our previous findings. Although the overall GSCI turned out to be largely irrelevant above, its sub-indices receive nontrivial weights in the optimal portfolios. While the minimum risk efficient portfolio contains only 1.58% and 2.20% of the energy and industrial metals sub-index, respectively, precious metals account for 8.24% and outweigh thus stocks, which carry a mere 5.54%. The GSCI Agriculture virtually drops out of the portfolio due to its negative average return and comparatively high volatility, while the livestock sub-index accounts for 7.01%, despite attaining a slightly negative return on average over our sample period. Relative to the equally weighted portfolio, the standard

Table 5: Optimal in-sample portfolios with the GSCI

Portfolio	Portfolio weight on			Monthly	Standard	Sharpe
	Bonds	Stocks	GSCI	Return	Deviation	Ratio
Equally weighted	33.33	33.33	33.33	0.3396%	0.0344	0.0554
Minimum risk efficient	72.82	27.18	0.00	0.3396%	0.0205	0.0928
Global minimum variance	90.90	7.60	1.50	0.3272%	0.0186	0.0959
Maximum Sharpe ratio	85.90	14.10	0.00	0.3314%	0.0188	0.0971

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2019. The target return of the minimum risk efficient portfolio is the return of the equally weighted portfolio.

Table 6: Optimal in-sample portfolios with GSCI sub-indices

Portfolio	Portfolio weight on							Monthly	Standard	Sharpe
	Bonds	Stocks	Energy	Ind.	Prec.	Agri.	Live.	Return	Deviation	Ratio
EQW	33.33	33.33	6.67	6.67	6.67	6.67	6.67	0.3423%	0.0273	0.0707
MRE	75.42	5.54	1.58	2.20	8.24	0.00	7.01	0.3423%	0.0182	0.1062
GMV	76.47	3.91	0.00	0.00	0.00	0.32	19.30	0.2455%	0.0161	0.0599
MSR	52.35	6.06	4.00	4.42	33.17	0.00	0.00	0.4880%	0.0266	0.1273

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2019. EQW denotes the equally weighted, MRE the minimum risk efficient, GMV the global minimum variance, and MSR the maximum Sharpe ratio portfolio, respectively.

deviation of monthly returns in the minimum risk efficient portfolio decreases from 2.73% to 1.82%, and the Sharpe ratio increases thus from 0.071 to 0.106.

The performance of the global minimum variance portfolio illustrates the pitfalls of disregarding returns in the portfolio optimization. Its lower standard deviation of 1.61% relative to the minimum risk efficient portfolio comes at the cost of an average monthly return that is about 10 basis points lower, implying a lower Sharpe ratio, as well. Note that the optimal portfolio assigns nonzero weights only to commodity futures that correlate negatively with bond or stock returns and offer thus the largest hedging gains. As a consequence, the GSCI Agriculture and Livestock sub-indices carry positive weights notwithstanding their negative average returns, while the portfolio is dominated by the global bonds index.

The maximum Sharpe ratio portfolio seems more interesting from a diversification point of view. While the GSCI Energy and Industrial Metals carry minor weights of about 4%, the precious metals sub-index accounts for one third of the optimal portfolio. This mix of bonds, stocks, and commodity futures sub-indices yields an average monthly return of almost 50 basis points at a standard deviation below that of the equally weighted portfolio and maximizes thus the Sharpe ratio at 0.127. Recall that the maximum Sharpe ratio in both Tables 4 and 5 was 0.097. Due to their negative average returns, neither the GSCI Agriculture nor the GSCI Livestock are included in the maximum Sharpe ratio portfolio.

When we split our sample in the period prior to and the period after the financial crisis, i.e. January 1999 through December 2007 and July 2009 through December 2019, respectively, we find that the overall GSCI and its five component sub-indices allow for optimal portfolios with higher average returns and lower volatility mainly during the earlier sub-sample. While the minimum risk efficient and the global minimum variance portfolio assign weights of one sixth to the GSCI Livestock due to its hedging benefits, the maximum Sharpe ratio portfolio assigns weights between 19 and 29.4% to the GSCI Energy, Industrial Metals, and Precious Metals sub-indices mainly for return considerations.

In contrast, the overall GSCI virtually drops out of the optimal portfolios in the latter sub-sample, where only livestock futures carry positive weights in the minimum risk efficient and the global minimum variance portfolio, while precious metals futures retain a positive weight in the maximum Sharpe ratio portfolio. Detailed results for the period prior to and the period after the financial crisis are presented and discussed in Appendix A.1.

5 Out-of-Sample Portfolio Optimization

An important caveat of the in-sample portfolio optimization in Section 4 is that it determines the optimal constant allocations only *ex post*, when the realized returns and volatilities have been observed. The previous results are thus purely descriptive. In this section, we instead investigate the opportunity for hedging gains from considering commodity futures *ex ante*. For this purpose, we set aside part of our sample as a pseudo out-of-sample period and use an econometric model to predict the volatilities and correlations of asset returns in real time.

5.1 Econometric methodology

The weapon of choice for modeling time series in empirical finance are so-called autoregressive conditional heteroscedasticity (ARCH) models (Engle, 1982), which simultaneously estimate a model of the mean and of the conditional variance of the time series. The popular generalized ARCH (GARCH) extension by Bollerslev (1986) furthermore includes a moving-average component in the model of the conditional variance that may facilitate a more parsimonious representation of time-varying volatilities.

As mentioned earlier, the potential gains from portfolio diversification crucially depend on the correlations between candidate assets. Given that standard GARCH models predict the conditional mean and variance of *univariate* time series, they cannot be used to predict the correlations between multiple assets. We therefore draw on the dynamic conditional correlation (DCC) model proposed by Engle (2002) (see also Chong and Miffre, 2010; Sadorsky, 2014; Pouliasis and Papapostolou, 2018), i.e. a multivariate approach to modeling both the conditional volatilities and the conditional correlations for a set of candidate assets.

The DCC model is estimated in two steps. In the first step, a univariate GARCH model is specified and estimated for each asset. In the second step, the univariate GARCH models are used to estimate the DCC coefficients (see Engle, 2002). Before setting up the DCC model and discussing the empirical results, the univariate GARCH model is shortly explained below. Given that the DCC model requires a sufficient number of observations in the estimation window, the subsequent analysis is based on weekly rather than monthly data.

5.1.1 GARCH model

Although the GARCH model for the conditional variance may be estimated jointly with an autoregressive moving average (ARMA) model for the conditional mean of each time series, neither of the weekly or monthly return series exhibits autoregressive or moving average patterns. Consistent with the efficient market hypothesis, the return series resemble white noise, whereas the volatility of returns is clustered in time and may thus be predictable.

For each asset, we estimate a GARCH(1,1) model, where the error term is modeled as

$$\varepsilon_t = \nu_t \sqrt{h_t}, \tag{8}$$

where ν_t denotes a white-noise process with $E[\nu_t] = 0$ and $\sigma_\nu^2 = 1$, and

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \tag{9}$$

Given that ν_t represents a white-noise process, both the conditional and the unconditional mean of ε_t equals zero, whereas the conditional variance of ε_t is given by $E_{t-1}[\varepsilon_t^2] = h_t$.

While less parsimonious ARCH or GARCH processes are conceivable for the conditional volatilities, the p-values of the Ljung-Box Q-statistics in Table 7 suggest that a GARCH(1,1) model eliminates any serial correlation in the simple and squared standardized residuals for

Table 7: p-values of Ljung-Box Q-statistics for univariate GARCH(1,1) processes

	standardized residuals			squared standardized residuals		
Lags	1	4	12	1	4	12
GSCI	0.1166	0.1836	0.1628	0.7243	0.5225	0.8721
Energy	0.1134	0.2238	0.1254	0.9560	0.3329	0.5320
Industrial Metals	0.3096	0.8635	0.3919	0.8511	0.8095	0.1495
Precious Metals	0.5434	0.8629	0.7910	0.0471**	0.0200**	0.0190**
Agriculture	0.5470	0.5439	0.7527	0.5680	0.9803	0.9570
Livestock	0.3527	0.0409	0.4165	0.5570	0.8382	0.9183
Stocks	0.8735	0.9686	0.9971	0.3573	0.6469	0.1181
Bonds	0.6058	0.2801	0.3660	0.7114	0.4733	0.4297

Note: ** indicates statistical significance at the 5% level.

all assets, except the GSCI Precious Metals, for the first 1, 4, and 12 weekly lags.

5.1.2 DCC model

In the DCC model, the vector of returns for N assets, $r_t = (r_{1t}, r_{2t}, ..., r_{Nt})$, is modeled as

$$r_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = z_t H_t^{1/2}, \tag{10}$$

where μ_t denotes the vector of conditional means of the returns. With no ARMA processes present in the returns, we assume a random walk for each asset price, implying constant conditional means, i.e. $\mu_t = \mu$, as in Fleming et al. (2001). In (10), z_t denotes the vector of standardized residuals and H_t the conditional covariance matrix

$$H_t \equiv D_t P_t D_t,$$

where

$$\begin{aligned}
D_t &= diag\left(h_{1t}^{1/2}, h_{2t}^{1/2}, \ldots, h_{Nt}^{1/2}\right), \\
P_t &= \left[diag\left(Q_t\right)\right]^{-1/2} Q_t \left[diag\left(Q_t\right)\right]^{-1/2}, \\
Q_t &= \left(1 - a - b\right) P + a z_{t-1} z_{t-1}' + b Q_{t-1},
\end{aligned}$$

i.e. D_t denotes the $N \times N$ diagonal matrix of volatilities and P_t the $N \times N$ symmetric matrix of dynamic conditional correlations with $\rho_{ii,t} = 1$ for $i = 1, 2, ...N$ $\forall t$. Q_t is an $N \times N$ symmetric positive-definite matrix, while $a \geq 0$ and $b \geq 0$ denote the shock and persistence parameter of the DCC model, respectively. The off-diagonal elements of H_t are equal to $h_{it}h_{jt}\rho_{ij,t}$, $i \neq j$ (see Bollerslev, 1990; Engle, 2002).

Given that the first step of estimating the DCC model assumes GARCH(1,1) processes, the conditional variance of asset i in period t is given by

$$h_{it} = \omega_i + \alpha_1 \left(r_{it-1} - \mu_{it}\right)^2 + \beta_1 h_{it-1},$$

where $\omega_i > 0$ and $\alpha_1, \beta_1 \geq 0$ guarantees non-negative variances, while $\alpha_1 + \beta_1 < 1$ ensures that the variance processes are stationary.

5.2 Empirical results for 1999–2019

In this section, we evaluate the out-of-sample performance of the optimal portfolios with and without the overall GSCI and its five component sub-indices, respectively, while splitting our sample into an estimation and an evaluation period. The optimal weight of each candidate asset is determined from an expanding estimation window that increases in length, as the evaluation period moves forward in time.[11] As a shortcut to accounting for transactions costs, we assume that the optimal portfolios are rebalanced at an annual frequency.

Given that the DCC model requires many observations in order to estimate the variances of and covariances between asset returns, the exercise is based on weekly data. As an initial estimation period, we set aside the first four years of our sample, i.e. 1999–2002, to predict correlations and derive thus optimal portfolio weights for the initial evaluation period, 2003. In each round of the portfolio optimization, the estimation window expands, as more recent observations are included, while the length of the evaluation period is constant at one year. Having estimated the coefficients of the DCC model, we then use the 52-week-ahead forecasts of the variances and covariances of asset returns to derive the optimal portfolio weights.

[11] Alternatively, we could use a rolling estimation window with constant length, which replaces the most distant observations in the window by more recent ones, as we move forward in time. In Appendix C, we show that our main results are robust to using a rolling rather than an expanding estimation window.

As in Section 4, we compare the performance of conventional stock-bond portfolios with an extension that considers either the overall GSCI or its component sub-indices in order to investigate whether commodity futures enhance the performance of the optimal portfolios and whether distinguishing between commodities allows for additional risk or return gains. Before presenting the optimal out-of-sample portfolios, we report the estimation results for the DCC model.

5.2.1 Dynamic conditional correlations

Table 8 summarizes the coefficient estimates of the DCC model including stocks, bonds, and the overall GSCI. For each asset class, μ denotes the intercept of the return series, i.e. the only coefficient estimated for the model of the mean. Similarly, ω denotes the intercept, α the shock, and β the persistence parameter of the respective GARCH process. While α and β are highly statistically significant for stocks, bonds, and the GSCI, ω is insignificant at

Table 8: DCC coefficient estimates for stocks, bonds and the GSCI

		Coefficient	Std. error	t-statistic	p-value
GARCH (1,1)					
Stocks	μ	0.0022	0.0006	3.8203	0.0001
	ω	0.0000	0.0000	1.9401	0.0524
	α	0.1640	0.0558	2.9362	0.0033
	β	0.8021	0.0609	13.1697	0.0000
Bonds	μ	0.0006	0.0003	2.2362	0.0253
	ω	0.0000	0.0000	0.7266	0.4675
	α	0.0515	0.0193	2.6655	0.0077
	β	0.9332	0.0239	38.9988	0.0000
GSCI	μ	0.0008	0.0009	0.8662	0.3864
	ω	0.0000	0.0000	1.6334	0.1024
	α	0.0685	0.0191	3.5907	0.0003
	β	0.9080	0.0279	32.5694	0.0000
DCC (1,1)					
	a	0.0129	0.0080	1.6166	0.1060
	b	0.8180	0.0960	8.5164	0.0000

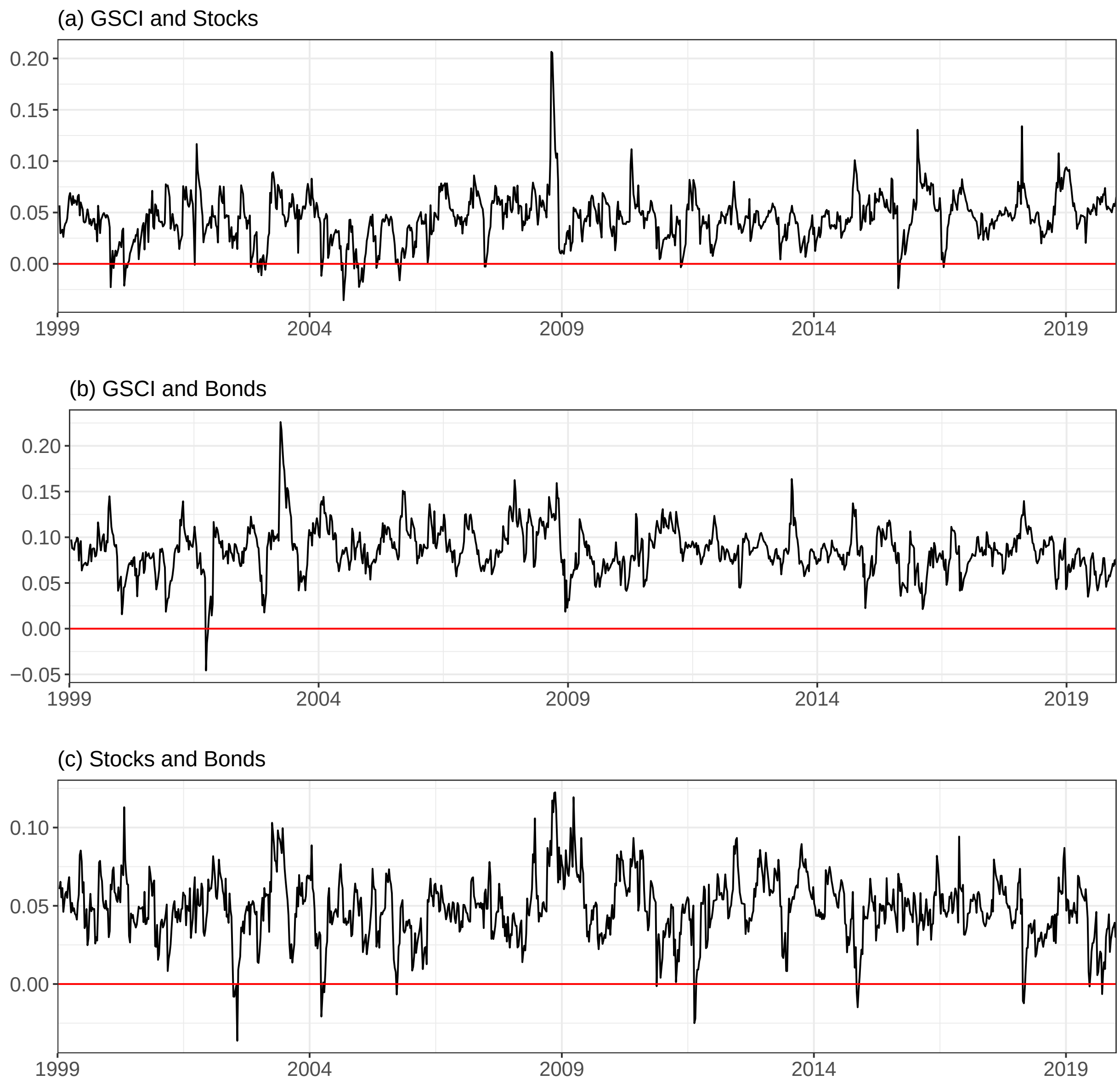

Figure 4: Dynamic conditional correlations of stocks, bonds, and the GSCI for 1999–2019

conventional levels, except for stocks.[12] The shock and persistence parameters of the DCC model, a and b, describe a non-linear relationship and are thus difficult to interpret.[13] Note that the parameters of the GARCH processes in the DCC model are the same as for the respective univariate processes.

[12]Given that $\alpha + \beta \leq 1$ by construction, without a non-zero intercept coefficient, the variance of an asset is weakly decreasing over time.

[13]With $a = 0$ and $b = 1$, the DCC model is identical to the constant conditional correlation (CCC) model.

Figure 4 plots the estimated dynamic conditional correlations of the model summarized in Table 8. Recalling that low or even negative correlations are most valuable for hedging during times of financial market turmoil, it is important to note the peak in panel (a), which suggests that the correlation between stocks and commodity futures exceeded 0.20 during the financial crisis of 2007–2008, eliminating thus any benefits from diversification. Panel (b) does not display a similar peak in the correlation between bonds and commodity futures. Instead, the latter declined and even became slightly negative during the dot-com bubble in 2002, when GSCI returns dropped, albeit not as strongly as stock returns.

Table 9 summarizes the coefficient estimates of the DCC model, when the overall GSCI is replaced by its five component sub-indices. Again, all estimated α and β coefficients as well as DCC parameters a and b are highly statistically significant. Figure 5 illustrates the corresponding dynamic conditional correlations. It is important to note that the correlation between stocks and the GSCI Energy sub-index increases during the financial crisis and remains high afterwards. Similarly, the correlations of all assets but bonds with the GSCI Livestock, which displays the lowest correlations on average over the sample period, increase during the financial crisis, when hedging benefits would have been most valuable. Finally, the positive correlation between the GSCI Precious Metals and bonds after the burst of the dot-com bubble and the US housing bubble in panel (o) illustrates the safe-haven properties of precious metals during turbulent times.

5.2.2 Maximum Sharpe ratio portfolio

Due to the so-called "separation property", a linear combination of the risk free asset and the maximum Sharpe ratio portfolio aligns with different investor preferences regardless of the degree of risk aversion, as the capital allocation line is weakly above the efficient frontier (Tobin, 1958). The only difference between a risk-averse and a risk-loving investor is that the former invests a larger share of the portfolio in the risk-free asset (see, e.g., Bodie et al., 2014). For this reason, we start by comparing the maximum Sharpe ratio portfolios while gradually increasing the set of candidate assets. Given the lack of serial correlation in the return series, we use the means over the 4-year estimation period to predict the returns

Table 9: DCC coefficient estimates for stocks, bonds, and GSCI sub-indices

		Coefficient	Std. error	t-statistic	p-value
GARCH (1,1)					
Stocks	μ	0.0022	0.0006	3.8225	0.0001
	ω	0.0000	0.0000	1.9506	0.0511
	α	0.1640	0.0554	2.9585	0.0031
	β	0.8021	0.0603	13.3007	0.0000
Bonds	μ	0.0006	0.0003	2.2392	0.0251
	ω	0.0000	0.0000	0.7237	0.4692
	α	0.0515	0.0194	2.6531	0.0080
	β	0.9332	0.0240	38.8062	0.0000
Energy	μ	0.0015	0.0012	1.2503	0.2112
	ω	0.0000	0.0000	1.3779	0.1682
	α	0.0764	0.0229	3.3310	0.0009
	β	0.8996	0.0349	25.8026	0.0000
Industrial	μ	0.0007	0.0007	0.9905	0.3220
	ω	0.0000	0.0000	0.4281	0.6685
	α	0.0878	0.0316	2.7773	0.0055
	β	0.9013	0.0254	35.4467	0.0000
Precious	μ	0.0010	0.0007	1.4207	0.1554
	ω	0.0000	0.0000	1.8832	0.0597
	α	0.0935	0.0319	2.9281	0.0034
	β	0.8738	0.0406	21.5032	0.0000
Agriculture	μ	-0.0011	0.0007	-1.5267	0.1268
	ω	0.0000	0.0000	3.0052	0.0027
	α	0.1272	0.0292	4.3501	0.0000
	β	0.8158	0.0368	22.1501	0.0000
Livestock	μ	0.0001	0.0006	0.2151	0.8297
	ω	0.0000	0.0000	6.6665	0.0000
	α	0.0468	0.0048	9.7694	0.0000
	β	0.9315	0.0061	153.0461	0.0000
DCC (1,1)					
	a	0.0181	0.0050	3.6032	0.0003
	b	0.9277	0.0310	29.8866	0.0000

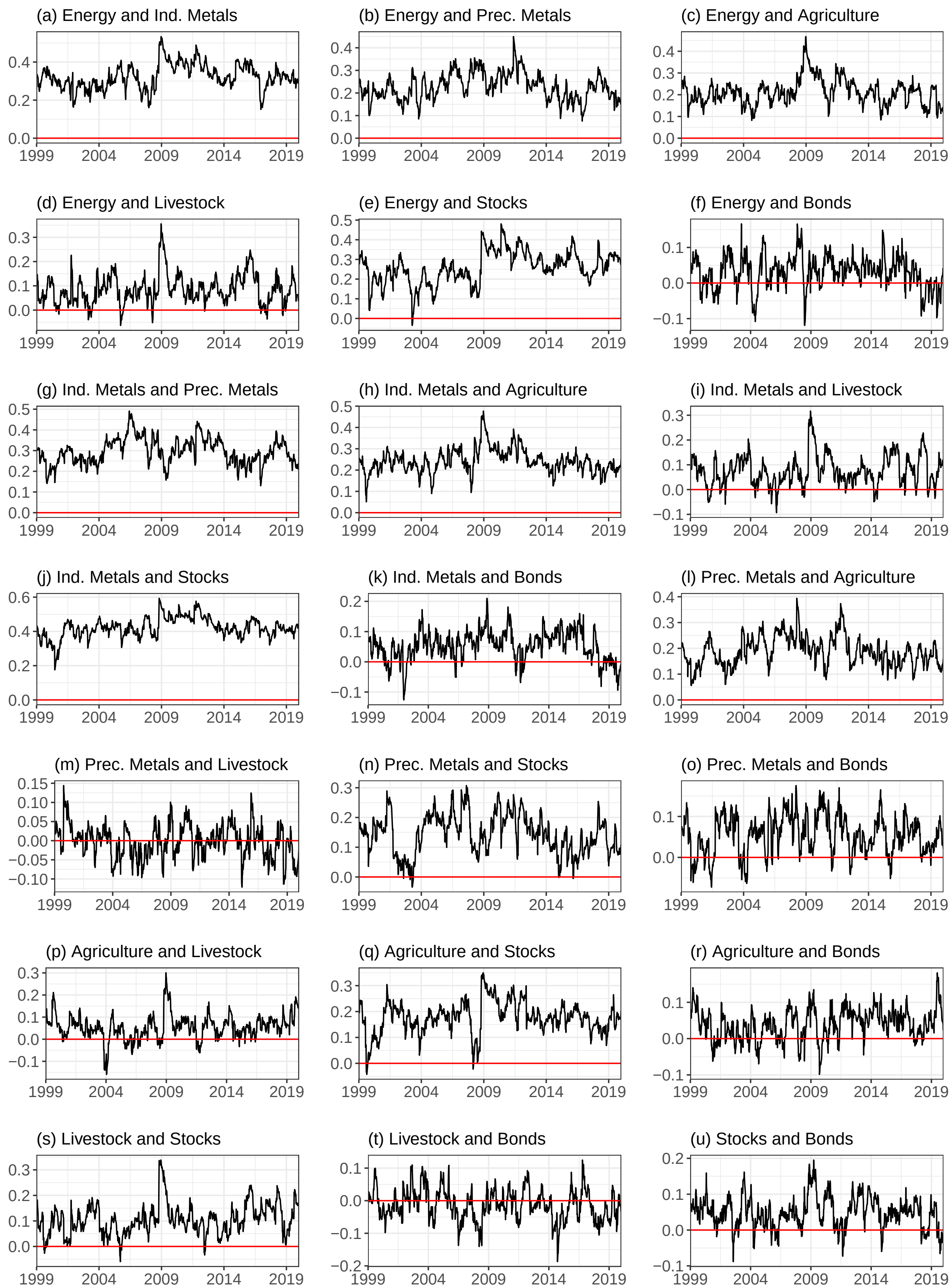

Figure 5: Dynamic conditional correlations of stocks, bonds, and GSCI sub-indices for 1999–2019

during the evaluation period.[14] As a proxy for the future risk-free rate, we use the average return on 3-month US treasury bills over the past 2 years. Due due to their stability over time, the results for 4 years are qualitatively identical.

Figure 6 illustrates the performance of three maximum Sharpe ratio portfolios with and without the overall GSCI and its five component sub-indices, respectively, for January 2003 through December 2019, where each portfolio is normalized to 100 in the first week of 2003. Recall that, in this exercise, the portfolios are rebalanced on an annually basis. The optimal out-of-sample weights for the maximum Sharpe ratio portfolios and their performances in each year are reported in Appendix B.

Prior to the financial crisis, the optimal portfolio with commodity futures outperforms the optimal portfolio with stocks and bonds only slightly. In the former, the overall GSCI carries an average weight of 23.2% during 2003–2008. This advance is lost during 2008–2009, when both the portfolio with and without the GSCI suffer similar losses. For 2009 through 2012, the two portfolios are fully invested in bonds and virtually moving together.

When considering the GSCI sub-indices separately, the relative gain over the conventional stock-bond portfolio is an order of magnitude larger. In 2008, the cumulated portfolio return peaks at 79.1%, as opposed to 43.2% for stocks and bonds and 48.9% for stocks, bonds, and the overall GSCI. Although none of the maximum Sharpe ratio portfolios passes the financial crisis without incurring substantial losses, the gap in cumulated returns widens rather than closes during 2009–2012, as the portfolios with and without the overall GSCI drop to around 80% of their initial values, as opposed to 125% when considering the sub-indices separately. Note that the latter portfolio assigns optimal weights between 27.6 and 43.5% to precious metals and a small weight to the GSCI Agriculture in 2011, while the remaining portfolio is invested in bonds. At the end of the sample period, the optimal portfolio including the GSCI sub-indices realizes a cumulated return of 98.9%, whereas the stock-bond portfolio with and without the overall GSCI realizes a cumulated return of 22.5% and 29.5%, respectively.

Figure 7 plots the annual Sharpe ratios of the optimal out-of-sample portfolios over time, where negative Sharpe ratios are set to zero due to their lack of interpretability. This is the

[14] Predicting returns based on a shorter window increases the volatility of the Sharpe ratios, in particular after the financial crisis of 2007–2008.

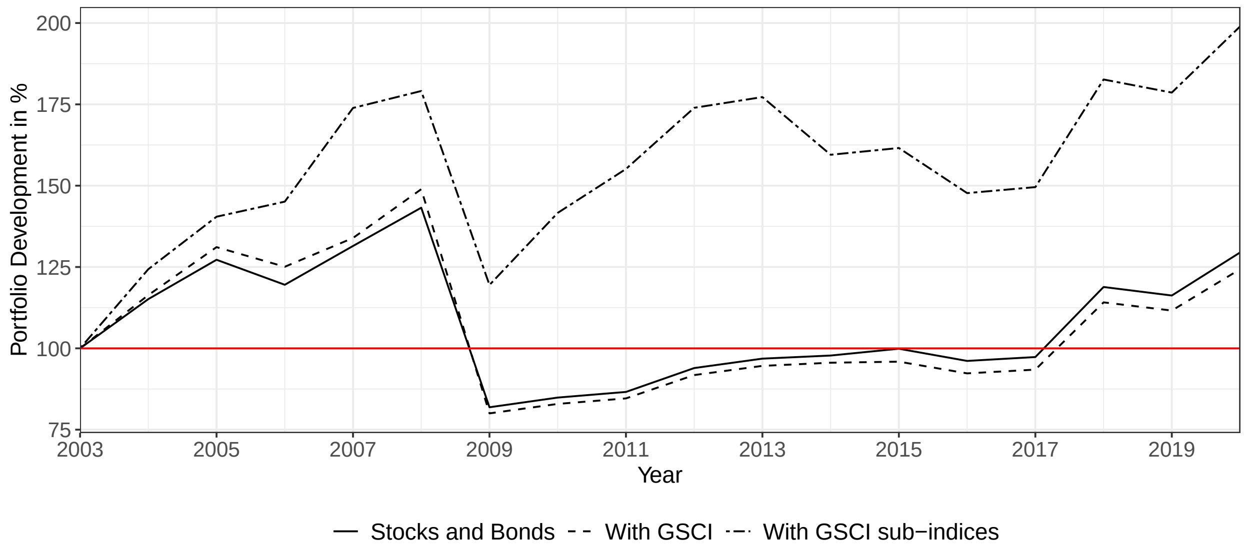

Figure 6: Performances of optimal portfolios for January 2003 through December 2019

case for all three portfolios in 2005, 2008, and 2015, for example.[15] A visual comparison of the Sharpe ratios reveals the superior return-to-risk ratios of the optimal portfolio including the GSCI sub-indices in 2009 and 2010, when it benefits from the relatively higher return and lower volatility of precious metals futures. Note that the Sharpe ratios in Figure 7 are determined ex-post, i.e. based on the observed portfolio returns and variances.

Note also that the Sharpe ratios of the optimal portfolios are identical in 2016, 2017, and 2019, when each of them is fully invested in stocks. This illustrates that the maximum Sharpe ratio portfolio is prone to putting high weight on individual asset classes that performed well in the recent past and incurs thus substantial risk. In what follows, we therefore complement our findings with an out-of-sample analysis based on an alternative investor objective.

5.2.3 Global Minimum variance portfolio

Historically, the global minimum variance portfolio has performed well in particular during periods of financial markets turmoil, such as the financial crisis of 2007–2008 (see, e.g., Clarke et al., 2011). Moreover, the minimum variance approach relies only on the out-of-sample predictions of variances and covariances, in line with the DCC model. By completely ignor-

[15] When the return of an asset falls short of the risk-free rate, the numerator in (6) becomes negative, and a larger variance in the denominator increases rather than decreases the Sharpe ratio, ceteris paribus.

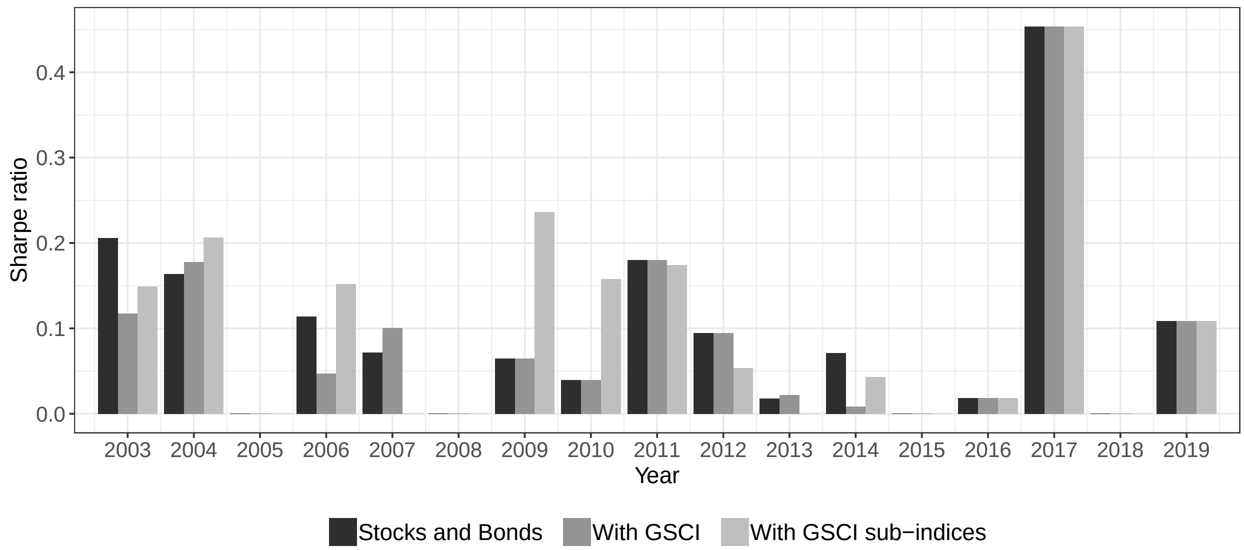

Figure 7: Sharpe ratios of optimal portfolios for January 2003 through December 2019

ing return considerations, however, the approach is prone to overweighing asset classes with inferior risk-return trade-offs or even negative average returns, such as the GSCI Agriculture and Livestock sub-indices. Given that many institutional investors refrain from investing in agricultural and livestock commodities for ethical reasons, we also compare the optimal portfolios from the previous section, which consider stocks and bonds as well as the overall GSCI and its five component sub-indices, with a global minimum variance portfolio that considers only a subset of the GSCI's components, while excluding the agriculture and livestock sub-indices from consideration.[16]

Figure 8 depicts the performance of the optimal minimum variance portfolios for January 2003 through December 2019, where each portfolio is normalized to 100 in the first week of 2003. In contrast to the maximum Sharpe ratio portfolios in Figure 6, all four portfolios pass the financial crisis relatively unscathed, illustrating the cardinal virtue of the minimum variance approach. By investing a larger share in lower-yielding yet safer asset classes, the approach avoids an excessively large exposure to unexpected turbulence in financial markets.

While the performance of the first three portfolios prior to the financial crisis is virtually

[16]Due to the negative average returns on the GSCI Agriculture and Livestock, a similar restriction on the set of candidate assets has no effect for the maximum Sharpe ratio approach, where they consistently carry very small weight in the optimal portfolio (see Table B.3).

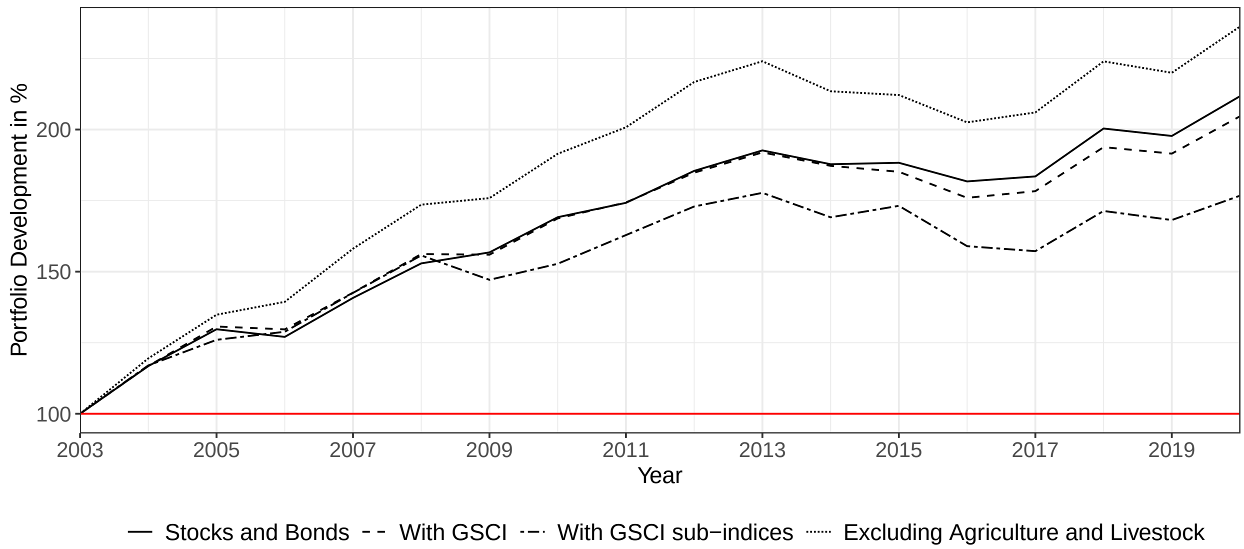

Figure 8: Performances of optimal portfolios for January 2003 through December 2019

identical, the portfolio considering the GSCI sub-indices separately falls short of the other two during 2007–2009 and fails to catch up over the rest of our sample period. In contrast, the stock-bond portfolios with and without the overall GSCI follow a very similar trend until the end of 2013, when the conventional portfolio starts to realize somewhat higher cumulative returns. It is important to note that all but the optimal portfolio with stocks, bonds, and the GSCI sub-indices clearly outperform the maximum Sharpe ratio portfolios in Figure 6 by the end of our sample period. The reason is that the global minimum variance portfolios don't suffer a major blow to their values during the financial crisis.

Note that the most successful portfolio, which excludes the GSCI Agriculture and Livestock from consideration, invests more than 50 percent into bonds in each year of the sample period (see Table B.8). Moreover, industrial and precious metals futures feature prominently, for instance in 2004, where they account for 16.1 and 20.9%, respectively. By the end of 2019, this portfolio realizes a cumulated return of 136.3% relative to its starting value in 2003, outperforming the most successful maximum Sharpe ratio portfolio, which considers all GSCI sub-indices separately, by a margin of 37.4 percentage points. The optimal stock-bond portfolios with and without the overall GSCI still attain cumulated returns of 104.7 and 111.7% by December 2019.

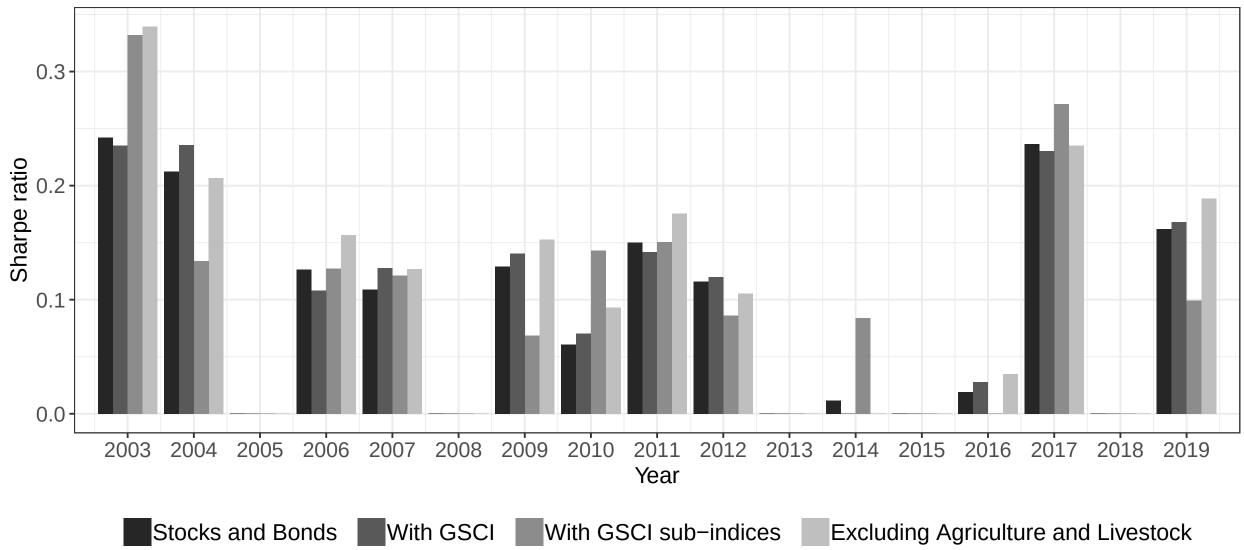

Figure 9: Sharpe ratios of optimal portfolios for January 2003 through December 2019

Note also that the GSCI Energy accounts for a maximum of 2.8% throughout our sample period, as opposed to a weight of 62.6% in the overall GSCI in 2019 (see Table 1). After 2011, it even drops out of the optimal portfolio including all GSCI sub-indices, illustrating yet again the importance of considering each component sub-index separately. The optimal out-of-sample weights for the global minimum variance portfolios and their performances in each year are reported in Appendix B.

Figure 9 plots the annual Sharpe ratios of the global minimum variance portfolios, where negative values are again set to zero. Note, in particular, that the portfolio excluding the GSCI Agriculture and Livestock attains the highest Sharpe ratio in six of the 12 years with non-zero values, including 2003, 2009, and 2019.

The (pseudo) out-of-sample exercise in this section warrants two conclusions. First, the overall GSCI helped enhance portfolio performances also *in real time* prior to the financial crisis of 2007–2008, while losing much of its suitability for diversification during the second half of the sample period. Second, investors may benefit from higher cumulated returns and more hedging opportunities when considering the GSCI's component sub-indices separately, in particular during periods of financial market turmoil.

6 The End of Normal Backwardation

In this section, we reassess the theory of normal backwardation described in Section 2. Consistent with Keynes' theory, earlier research by Bodie and Rosansky (1980) and Gorton and Rouwenhorst (2006) reports significant risk premia in a broad index based on an equally weighted cash-collateralized portfolio of commodity futures.[17]

When replicating our in-sample portfolio optimization for the period prior to and after the financial crisis of 2007–2008, we find that the overall GSCI and most of its component sub-indices drops out of the optimal portfolios in the later sub-sample (see Appendix A.2), suggesting a structural break in commodity futures markets after the crisis. However, average returns are not sufficient for contesting the theory of normal backwardation, as relatively lower returns on futures could simply reflect a downward trend in spot prices. To evaluate Keynes' theory, we therefore contrast the relative development of a commodity's futures price with that of the corresponding spot price during our sample period.

Panel (a) of Figure 10 illustrates the development of an index of spot and futures prices, respectively, for the overall GSCI during 1999–2019. Until October 2004, the relative price development is in line with normal backwardation, as the index of commodity futures prices increases faster than the spot price. Shortly before the financial crisis, however, the relative performance reverses, and there are considerable *rolling costs* rather than *rolling gains*, as postulated by Keynes.

The remaining panels of Figure 10 depict the relative price developments for the GSCI's component sub-indices. It is interesting to note that the energy futures in panel (b) perform better than the spot price prior to, whereas the situation reverses during the financial crisis. At the end of the sample period, the spot index outperforms the futures index by a margin of 331.8 relative to 65.7%, implying that the cumulated return on energy spot prices was about five times larger than on energy futures.[18] While one might argue that the unprecedented increase in the spot price is largely due to the all-time low in real oil prices in November

[17]Analyzing commodity futures during 1950–1976, Bodie and Rosansky (1980) find average returns comparable to those of common stocks. Similarly, Gorton and Rouwenhorst (2006) find competitive average returns during 1959–2004 and an annual risk premium on commodity futures of about 5%.

[18]Note that a spot-market investment in energy commodities requires buying and storing large amounts of oil and gas physically for an extended period, which is unlikely to be economically profitable due to non-trivial storing costs and availability in the short run.

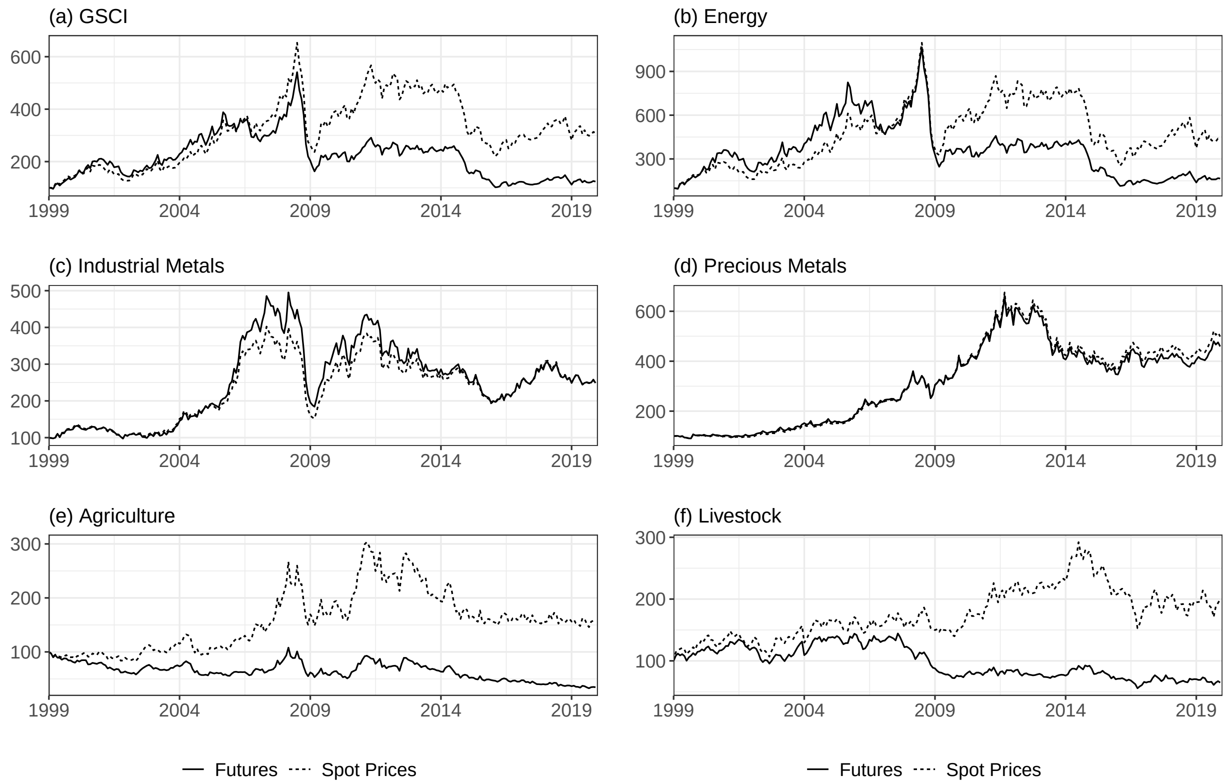

Figure 10: Spot versus futures price index for the GSCI and its five component sub-indices (normalized to 100 in January 1999)

1998, the divergence of the spot and futures price index in panel (b) over time is far from homogeneous and suggests a structural break around the financial crisis.

The spot and futures price indices for industrial metals in panel (c) display a very different relationship. While they move together at the beginning of our sample period, futures prices start outperforming the spot price prior to the financial crisis. After the crisis, the situation reverses and the returns on the spot price outperform the returns on futures prices. By the end of our sample period, the cumulated return on the GSCI Industrial Metals and on the spot price index amount to 146.54 and 248.6%, respectively. However, this similarity must be considered through the lens of a temporary excess return of 94.9 percentage points on the futures index in March 2008, which is used up completely by December 2019.

Similarly, the spot and futures price indices for precious metals in panel (d) are virtually indistinguishable up to their joint peak in September 2011 of 675.4 and 659.6, respectively. From this point onwards, the spot price index outperform the GSCI Precious Metals by about 40 percentage points until the end of our sample period.

Due to the importance of physical storage costs, the spot and futures price indices for agricultural and livestock commodities in panels (e) and (f) display a peculiar relationship.[19] From the beginning of our sample period, both commodities exhibit substantial rolling costs, which materialize as an increasing gap between the cumulated return on the spot and the futures price index. Consequently, an investment into the respective futures index in January 1999 realizes a loss in December 2019, whereas a spot index investment yields a gain. Note that the differences are sizeable. During our sample period, the spot price index for agricultural and livestock commodities increases to 157.9 and 199.7, whereas the corresponding futures price index decreases to 64.8 and 34.6%, respectively, of an initial investment.

Figure 10 illustrates substantial differences in the performances of spot and futures prices and suggests a structural break in commodity markets, except for agriculture and livestock, around the financial crisis of 2007–2008. The evidence in this paper qualifies thus the result in Gorton and Rouwenhorst (2006) that an investment in futures contracts outperforms an investment in commodities on the spot, consistent with the theory of normal backwardation. During our sample period, this was *not* the case for the GSCI Agriculture and Livestock, and is no longer the case for any other commodity in the GSCI in December 2019. As a result, commodity producers obtain rather than pay a premium for hedging their sales prices against unexpected fluctuations in the future spot price.

While this may seem counterintuitive, one candidate explanation is the relatively recent "financialization" of commodity markets, as institutional investors are searching for opportunities to diversify their portfolios. Another possible explanation for the persistent change in risk premia is that the effect of rolling activities on a futures price index increases in the volume of the index. Hirshleifer (1989, 1990), finally, argues that risk premia in futures markets depend on whether hedgers are short or long in the underlying asset. Building on this hedging-pressure hypothesis, Basu and Miffre (2013) propose longing commodity futures for which hedgers are short and speculators are long, while shorting commodity futures for which hedgers are long and speculators are short. The authors find that a long-short portfolio outperforms a long-only portfolio, consistent with hedging pressure.

[19]Note that some agricultural or livestock commodities are perishable and thus inherently unstorable.

7 Conclusion

Historically, commodity futures exhibited competitive returns and low correlations with traditional asset classes, such as stocks and bonds (see, e.g., Bodie and Rosansky, 1980; Gorton and Rouwenhorst, 2006). At the same time, an increasing "financialization" of commodities took place prior to and during the financial crisis of 2007–2008, possibly inducing fundamental changes in the hedging properties of commodity futures. In this paper, we therefore investigate whether the benefits recorded in the prior literature remain prevalent during 1999–2019.

In-sample portfolio optimizations reveal that Standard&Poor's global commodity investment benchmark (formerly the *Goldman Sachs Commodity Index*) GSCI obtains non-trivial weights in the optimal portfolios prior to the financial crisis, facilitating considerably higher Sharpe ratios. In contrast, the overall GSCI virtually drops out of the optimal portfolios after 2008, which is at odds with earlier findings.

We proceed to show that considering separately the GSCI's five component sub-indices considerably enhances the performance of the optimal minimum risk efficient and maximum Sharpe ratio portfolios on average over the sample period. Even after the financial crisis, precious metals futures feature prominently in the optimal maximum Sharpe ratio portfolio, corroborating the actual or perceived safe-haven properties of gold and silver. Similar benefits are not attainable by investing in the overall GSCI, where precious metals carry a weight of 4% in 2019, for example. This emphasizes the importance of considering each commodity sub-index as a separate candidate asset.

Our in-sample analysis is complemented by a (pseudo) out-of-sample portfolio optimization, where we use a dynamic conditional correlation (DCC) model to predict the correlations between candidate assets. Including the overall GSCI in this real-time portfolio optimization helps outperform the optimal portfolio with only stocks and bonds prior to 2008, whereas the optimal portfolio including the GSCI yields lower cumulated returns during the entire evaluation period, 2003–2019. The largest cumulated return for the maximum Sharpe ratio approach is again obtained, when considering the GSCI's component sub-indices separately. For the minimum variance approach, however, excluding the GSCI Agriculture and Livestock enhances the performance, as the benefit of low conditional correlations with other asset classes comes at the costs of negative returns on average over the sample period.

We argue that the observed changes in the hedging properties of commodity futures are due to the recent failure of Keynes' theory of normal backwardation. In line with this theory, rolling gains were a stylized fact prior to 2005, whereas all commodity sub-indices exhibit rolling costs after the financial crisis of 2007–2008, implying an inferior risk-return trade-off.

While the data and evidence in this paper does not warrant a conclusive explanation for the changes in commodity risk premia, it illustrates that the commodity futures included in the overall GSCI lost their suitability for portfolio diversification after the financial crisis or earlier due to considerable rolling costs. Relative to stock markets, for example, there remains a lack of knowledge about the return properties and hedging opportunities of commodity futures (see Rouwenhorst and Tang, 2012), warranting future research.

References

Basu, D. & Miffre, J. (2013). "Capturing the risk premium of commodity futures: The role of hedging pressure." *Journal of Banking & Finance* 37(7), 2652–2664.

Baumeister, C. & Kilian, L. (2016). "Understanding the decline in the price of oil since June 2014." *Journal of the Association of Environmental and Resource Economists* 3(1), 131–158.

Baur, D. G. & Dermott, T. K. (2010). "Is gold a safe haven? International evidence." *Journal of Banking & Finance* 34(8), 1886–1898.

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

Bodie, Z. (1983). "Commodity futures as a hedge against inflation." *The Journal of Portfolio Management* 9(3), 12–17.

Bodie, Z., Kane, A. & Marcus, A. J. (2014). *Investments* (10th ed.), McGraw-Hill Education.

Bodie, Z. & Rosansky, V. I. (1980). "Risk and return in futures markets." *Financial Analysts Journal* 36, 27–36.

Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 31(3), 307–327.

Bollerslev, T. (1990). "Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model." *The Review of Economics and Statistics* 72(3), 498–505.

Büyüksahin, B., Haigh, M. S. & Robe, M. A. (2009). "Commodities and equities: Ever a 'Market of One'?" *Journal of Alternative Investments* 12(3), 76–95.

Chong, J. & Miffre, J. (2010). "Conditional return correlations between commodity futures and traditional assets." *Journal of Alternative Investments* 12(3), 61–75.

Clarke, R., De Silva, H. & Thorley, S. (2011). "Minimum-variance portfolio composition." *Journal of Portfolio Management* 37(2), 31–45.

Domanski, D. & Heath, A. (2007). "Financial investors and commodity markets." *BIS Quarterly Review*, March 2017.

Edwards, F. R. & Park, J. M (1996). "Do managed futures make good investments?" *Journal of Futures Markets* 16(5), 475–517.

Ellwanger, R., Sawatzky, B. & Zmitrowitz, K. (2017). "Factors behind the 2014 oil price decline." *Bank of Canada Review*, Autumn 2017, 1–13.

Engle, R. F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation." *Econometrica* 50(4), 987–1007.

Engle, R. F. (2002). "Dynamic conditional correlation: A simple class of multivaritate generalized autoregressive conditional heteroskedasticity models." *Journal of Business & Economic Statistics* 20(3), 339–350.

Fleming, J., Kirby, C. & Ostdiek, B. (2001). "The economic value of volatility timing." *Journal of Finance* 56(1), 329–352.

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

Hicks, J. R. (1939). *Value and capital*. Oxford University Press, London, 1939.

Hirshleifer, D. (1989). "Determinants of hedging and risk premia in commodity futures markets." *Journal of Financial and Quantitative Analysis* 24(3), 313–331.

Hirshleifer, D. (1990). "Hedging pressure and futures price movements in a general equilibrium model." *Econometrica* 58(2), 411–428.

Jaiswal, R. & Uchil, R. (2018). "An analysis of diversification benefits of commodity futures using Markov regime-switching approach." *Afro-Asian Journal of Finance and Accounting* 8(1), 20–47.

Jensen, G. R., Johnson, R. R. & Mercer, J. M. (2000). "Efficient use of commodity futures in diversified portfolios." *Journal of Futures Markets* 20(5), 489–506.

Keynes, J. M. (1930). "A treaties on money: The applied theory of money, 2nd volume." *The Collected Works of J. M. Keynes* VI.

Markowitz, H. M. (1952). "Portfolio selection" *Journal of Finance* 7(1), 77–91.

Mandelbrot, B. (1963). "The variation of certain speculative prices." *Journal of Business* 36(4), 394–419.

Pouliasis, P. K. & Papapostolou, N. C. (2018). "Volatility and correlation timing: The role of commodities." *Journal of Futures Markets* 38(11), 1407–1439.

Rouwenhorst, K. G. & Tang, K. (2012). "Commodity investing." *Annual Review of Financial Economics* 4, 447–467.

Sadorsky, P. (2014). "Modeling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat." *Energy Economics* 43, 72–81.

Sharpe, W. F. (1994). "The Sharpe ratio." *Journal of Portfolio Management* 21(1), 49–58.

Silvennoinen, A. & Thorp, S. (2013). "Financialization, crisis and commodity correlation dynamics." *Journal of International Financial Markets, Institutions, and Money* 24, 42–65.

Tang, K. & Xiong, W. (2012). "Index investment and the financialization of commodities." *Financial Analysts Journal* 68(6), 54–74.

Tobin, J. (1958). "Liquidity preference as behavior towards risk." *Review of Economic Studies* 25(2), 65–86.

A In-Sample Portfolio Optimization

A.1 Empirical result for 1999–2007

Table A.1 reports the results of the in-sample portfolio optimization when considering stocks and bonds only prior to the financial crisis of 2007–2008. During the first sub-sample, the monthly returns of all three optimal portfolios are substantially higher than in Table 4 in the main text, while the Sharpe ratios are similar, due to a higher risk-free rate of return. Note that the average monthly rate of return on 3-month treasury bills dropped from 0.282% for 1999–2007 to 0.046% for 2009–2019.

Prior to the financial crisis, the global minimum variance portfolio is all but identical to the maximum Sharpe ratio portfolio, as both assign weights of 4/5 and 1/5 on bonds and stocks, respectively. The monthly return on either portfolio increases slightly relative to the equally weighted portfolio, while the standard deviation decreases from 2.21% to 1.76%, facilitating a substantially higher Sharpe ratio for either optimal portfolio.

Table A.2 illustrates the fundamentally different performances of commodity futures prior to and after the financial crisis of 2007–2008. While the overall GSCI dropped out of the minimum risk efficient and the maximum Sharpe ratio portfolio for the entire sample period, it now accounts for weights between 1/3 and 2/5, even entering the global minimum variance portfolio, albeit with a small weight. Across optimal portfolios, bonds continue to dominate with weights ranging from one half in the maximum Sharpe ratio to 4/5 in the global minimum variance portfolio, whereas the weight on stocks drops to between 6 and 18%.

Considering the overall GSCI generally increases the average monthly return and thus the

Table A.1: Optimal in-sample portfolios with stocks and bonds for 1999–2007

Portfolio	Portfolio weight on		Monthly	Standard	Sharpe
	Bonds	Stocks	Return	Deviation	Ratio
Equally weighted	50.00	50.00	0.4497%	0.0221	0.0759
Global minimum variance	80.74	19.26	0.4500%	0.0176	0.0956
Maximum Sharpe ratio	80.85	19.15	0.4500%	0.0176	0.0956

Note: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2007.

Table A.2: Optimal in-sample portfolios with the GSCI for 1999–2007

Portfolio	Portfolio weight on Bonds	Stocks	GSCI	Monthly Return	Standard Deviation	Sharpe Ratio
Equally weighted	33.33	33.33	33.33	0.7650%	0.0277	0.1746
Minimum risk efficient	58.25	8.45	33.30	0.7650%	0.0255	0.1897
Global minimum variance	78.49	18.19	3.32	0.4814%	0.0174	0.1143
Maximum Sharpe ratio	53.00	5.93	41.08	0.8386%	0.0292	0.1909

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2007. The target return of the minimum risk efficient portfolio is the return of the equally weighted portfolio.

Sharpe ratio, consistent with the finding in the existing literature that including commodity futures may increase a portfolio's performance and diversification. Note that, by including the GSCI among candidate assets, the maximum Sharpe ratio for 1999–2007 almost doubles from 0.096 in Table A.1 to 0.191 in Table A.2.

Table A.3 presents the optimal portfolios after replacing the GSCI by its five component sub-indices. The equally weighted portfolio now attains a lower average monthly return than in Table A.2, given that energy futures, which performed well prior to the financial crisis, carry a higher-than-equal weight in the overall GSCI. Both the minimum risk efficient and the global minimum variance portfolio assign weights of about 1/6 to the GSCI Livestock, while energy, industrial and precious metals futures enter only the minimum risk efficient portfolio with modest weights. This is due to their comparatively larger volatility.

When return considerations play a role, i.e. in the minimum risk efficient and the max-

Table A.3: Optimal in-sample portfolios with GSCI sub-indices for 1999–2007

Portfolio	Portfolio weight on Bonds	Stocks	Energy	Ind.	Prec.	Agri.	Live.	Monthly Return	Standard Deviation	Sharpe Ratio
EQW	33.33	33.33	6.67	6.67	6.67	6.67	6.67	0.6268%	0.0209	0.1648
MRE	57.53	8.33	5.74	7.14	5.23	0.00	16.02	0.6268%	0.0168	0.2057
GMV	61.47	14.47	1.88	0.00	0.00	4.82	17.37	0.4283%	0.0152	0.0963
MSR	19.62	0.00	19.17	29.44	23.72	0.00	8.05	1.2097%	0.0329	0.2819

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for January 1999 through December 2007. EQW denotes the equally weighted, MRE the minimum risk efficient, GMV the global minimum variance, and MSR the maximum Sharpe ratio portfolio, respectively.

imum Sharpe ratio portfolio, the GSCI Agriculture is not included. It is important to note that stocks are no longer included in the maximum Sharpe ratio portfolio either, while bonds no longer dominate the portfolio. Instead, industrial and precious metals futures now obtain optimal weights of 29.44 and 23.72%, respectively, while energy and livestock futures contribute another 19.17 and 8.05%, respectively. As a result, the monthly return of the maximum Sharpe ratio portfolio increases relative to Table A.2, from 0.839 to 1.210%, while the Sharpe ratio increases from 0.191 to 0.282.

A.2 Empirical results for 2009–2019

Table A.4 summarizes the optimal portfolios with bonds and stocks only after the financial crisis of 2007–2008. Note first that the Sharpe ratios across portfolios are substantially higher during the second half of our sample period, as the risk-free rate of return drops to 0.046%. Bonds again account for the lion's share in the global minimum variance portfolio, whereas stocks dominate by a small margin in the maximum Sharpe ratio portfolio.

Given that the overall GSCI had a positive weight prior to the financial crisis (see Table A.2), yet dropped from the optimal portfolios for the entire sample period (see Table 5), it is unlikely to feature prominently after the crisis. Indeed, the GSCI accounts for less than 3% of the minimum risk efficient and the global minimum variance portfolio and drops out of the maximum Sharpe ratio portfolio in Table A.5. Bonds account for about 92% in the former two portfolios, while the maximum Sharpe ratio portfolio is unaffected by the consideration of commodity futures. Finally, note that the Sharpe ratio of the global minimum variance portfolio with commodity futures is *lower* than for the corresponding portfolio in Table A.4.

Table A.4: Optimal in-sample portfolios with stocks and bonds for 2009–2019

Portfolio	Portfolio weight on		Monthly	Standard	Sharpe
	Bonds	Stocks	Return	Deviation	Ratio
Equally weighted	50.00	50.00	0.4540%	0.0236	0.1728
Global minimum variance	92.37	7.63	0.2337%	0.0164	0.1148
Maximum Sharpe ratio	46.42	53.58	0.4726%	0.0247	0.1730

Note: The risk-free rate equals the average monthly return on 3-month US treasury bills for July 2009 through December 2019.

Table A.5: Optimal in-sample portfolios with the GSCI for 2009–2019

Portfolio	Portfolio weight on Bonds	Stocks	GSCI	Monthly Return	Standard Deviation	Sharpe Ratio
Equally weighted	33.33	33.33	33.33	0.2070%	0.0306	0.0489
Minimum risk efficient	92.24	5.03	2.73	0.2070%	0.0164	0.0987
Global minimum variance	92.07	5.53	2.40	0.2112%	0.0164	0.1012
Maximum Sharpe ratio	46.42	53.58	0.00	0.4726%	0.0247	0.1730

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for July 2009 through December 2019. The target return of the minimum risk efficient portfolio is the return of the equally weighted portfolio.

Table A.6 illustrates the consequences of replacing the GSCI by its five component sub-indices. The minimum risk efficient portfolio now comprises only bonds, stocks and the GSCI Livestock. Relative to the equally weighted portfolio, its standard deviation drops from 2.48 to 1.65%, while the Sharpe ratio increases to 0.1523. The global minimum variance approach facilitates a further reduction of the portfolio's standard deviation at the cost of the lowest average monthly return in Table A.6, as the GSCI Agriculture and Livestock contribute negatively to the portfolio's return during 2009–2019. Interestingly, the maximum Sharpe ratio portfolio continues to assign the largest weight to stocks and a nontrivial weight to the GSCI Precious Metals, attaining an average monthly return of 0.511% and a Sharpe ratio of 0.175 even after the financial crisis.

Table A.6: Optimal in-sample portfolios with GSCI sub-indices for 2009–2019

Portfolio	Portfolio weight on Bonds	Stocks	Energy	Ind.	Prec.	Agri.	Live.	Monthly Return	Standard Deviation	Sharpe Ratio
EQW	33.33	33.33	6.67	6.67	6.67	6.67	6.67	0.2964%	0.0248	0.1013
MRE	68.62	23.49	0.00	0.00	0.00	0.00	7.89	0.2964%	0.0165	0.1523
GMV	79.44	1.50	0.00	0.00	0.00	0.98	18.08	0.1522%	0.0139	0.0769
MSR	33.48	57.15	0.00	0.00	9.36	0.00	0.00	0.5113%	0.0266	0.1751

Notes: The risk-free rate equals the average monthly return on 3-month US treasury bills for July 2009 through December 2019. EQW denotes the equally weighted, MRE the minimum risk efficient, GMV the global minimum variance, and MSR the maximum Sharpe ratio portfolio, respectively.

A.3 Discussion of sub-sample results

The previous analysis suggests that the properties of commodity futures and thus the hedging gains of including them in a global portfolio changed during the financial crisis of 2007–2008. Prior to the crisis, the GSCI and its five sub-indices facilitated considerable improvements in the portfolios' average returns and standard deviations. After the crisis, the GSCI virtually drops from all optimal portfolios, while the GSCI Precious Metals is the only sub-index included in the maximum Sharpe ratio portfolio in Table A.6.

Given that commodity futures accounted for a large share of the extraordinary returns prior to the financial crisis, average returns are considerably lower after 2009. Comparing the average monthly returns of the maximum Sharpe ratio portfolios in Tables A.3 and A.6, we find that the former is more than twice as large as the latter, while the Sharpe ratio decreases from 0.282 to 0.175. A candidate explanation is provided in Section 6 in the main text, where we consider changes in the rolling costs and returns of commodities in both spot and futures markets.

An important caveat of the in-sample portfolio optimizations is that they determine the optimal constant allocations only *ex post*, when the realized returns and volatilities have already been observed. The previous results are thus purely descriptive rather than suitable for improving a portfolio's performance *in real time*. Nevertheless, they demonstrate the possible hedging gains from including commodity futures in a conventional stock-bond portfolio as well as the possible value added of considering the GSCI sub-indices separately.

B Optimal Out-of-Sample Portfolio Weights

Table B.1: Optimal weights in maximum Sharpe ratio portfolio with stocks and bonds

Year	Stocks	Bonds
2003	0.00%	100.00%
2004	0.00%	100.00%
2005	0.00%	100.00%
2006	31.58%	68.42%
2007	66.21%	33.79%
2008	95.70%	4.30%
2009	0.00%	100.00%
2010	0.00%	100.00%
2011	0.00%	100.00%
2012	0.00%	100.00%
2013	23.38%	76.62%
2014	36.73%	63.27%
2015	48.81%	51.19%
2016	100.00%	0.00%
2017	100.00%	0.00%
2018	78.64%	21.36%
2019	100.00%	0.00%

Table B.2: Optimal weights in maximum Sharpe ratio portfolio with stocks, bonds, and the GSCI

Year	GSCI	Stocks	Bonds
2003	71.07%	0.00%	28.93%
2004	23.55%	0.00%	76.45%
2005	5.39%	0.00%	94.61%
2006	15.29%	30.87%	53.84%
2007	11.51%	62.42%	26.07%
2008	35.42%	64.58%	0.00%
2009	0.00%	0.00%	100.00%
2010	0.00%	0.00%	100.00%
2011	0.00%	0.00%	100.00%
2012	0.00%	0.00%	100.00%
2013	9.16%	22.22%	68.62%
2014	6.31%	35.34%	58.35%
2015	0.00%	48.82%	51.18%
2016	0.00%	100.00%	0.00%
2017	0.00%	100.00%	0.00%
2018	0.00%	78.64%	21.36%
2019	0.00%	100.00%	0.00%

Table B.3: Optimal weights in maximum Sharpe ratio portfolio with stocks, bonds, and GSCI sub-indices

Year	Energy	Industrial	Precious	Agriculture	Livestock	Stocks	Bonds
2003	59.58%	0.00%	21.18%	0.00%	2.51%	0.00%	16.73%
2004	12.59%	0.00%	29.31%	0.00%	2.50%	0.00%	55.60%
2005	3.71%	8.77%	18.58%	0.00%	1.82%	0.00%	67.12%
2006	7.21%	21.63%	15.43%	0.00%	3.69%	6.63%	45.43%
2007	0.53%	38.32%	7.55%	0.00%	3.68%	26.27%	23.65%
2008	5.90%	46.26%	32.24%	0.00%	0.00%	15.60%	0.00%
2009	0.00%	0.00%	43.51%	0.00%	0.00%	0.00%	56.49%
2010	0.00%	3.93%	27.60%	0.00%	0.00%	0.00%	68.47%
2011	0.00%	0.00%	32.98%	3.05%	0.00%	0.00%	63.97%
2012	0.00%	0.00%	28.81%	0.00%	0.00%	0.00%	71.19%
2013	0.00%	7.46%	22.08%	5.66%	0.00%	8.22%	56.58%
2014	3.35%	0.00%	2.94%	1.92%	5.93%	30.22%	55.63%
2015	0.00%	0.00%	0.00%	0.00%	24.11%	35.98%	39.91%
2016	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
2017	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
2018	0.00%	0.00%	0.00%	0.00%	0.00%	78.59%	21.41%
2019	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%

Table B.4: Out-of-sample performances of maximum Sharpe ratio portfolios

Year	Stocks and Bonds	With GSCI	With GSCI sub-indices
2003	100.00	100.00	100.00
2004	115.13	116.36	124.32
2005	127.23	131.12	140.47
2006	119.54	125.08	145.10
2007	131.46	133.92	173.89
2008	143.23	148.91	179.10
2009	81.89	80.01	119.48
2010	84.86	82.91	141.63
2011	86.59	84.61	155.13
2012	93.94	91.78	173.94
2013	96.83	94.61	177.24
2014	97.79	95.57	159.51
2015	99.87	95.90	161.58
2016	96.13	92.30	147.72
2017	97.32	93.45	149.56
2018	118.86	114.13	182.65
2019	116.23	111.61	178.62
2020	129.46	124.31	198.94

Table B.5: Optimal weights in global minimum variance portfolio with stocks and bonds

Year	Stocks	Bonds
2003	13.51%	86.49%
2004	20.63%	79.37%
2005	21.49%	78.51%
2006	38.72%	61.28%
2007	24.58%	75.42%
2008	17.16%	82.84%
2009	13.64%	86.36%
2010	12.69%	87.31%
2011	13.11%	86.89%
2012	10.15%	89.85%
2013	8.39%	91.61%
2014	11.18%	88.82%
2015	12.23%	87.77%
2016	12.70%	87.30%
2017	17.59%	82.41%
2018	14.15%	85.85%
2019	8.86%	91.14%

Table B.6: Optimal weights in global minimum variance portfolio with stocks, bonds, and the GSCI

Year	GSCI	Stocks	Bonds
2003	10.23%	12.31%	77.46%
2004	9.15%	19.01%	71.84%
2005	5.22%	20.99%	73.79%
2006	4.26%	38.29%	57.45%
2007	4.57%	24.31%	71.12%
2008	4.70%	16.87%	78.43%
2009	3.93%	13.32%	82.75%
2010	4.56%	12.29%	83.15%
2011	4.94%	12.64%	82.42%
2012	4.35%	9.86%	85.79%
2013	3.11%	8.23%	88.66%
2014	5.10%	10.83%	84.07%
2015	4.39%	11.89%	83.72%
2016	4.62%	12.30%	83.07%
2017	7.27%	16.48%	76.25%
2018	6.13%	13.42%	80.45%
2019	4.49%	8.45%	87.06%

Table B.7: Optimal weights in global minimum variance portfolio with stocks, bonds, and GSCI sub-indices

Year	Energy	Industrial	Precious	Agriculture	Livestock	Stocks	Bonds
2003	0.18%	10.90%	9.40%	9.87%	14.77%	5.03%	49.85%
2004	0.52%	10.26%	15.10%	13.30%	14.85%	6.47%	39.50%
2005	0.75%	5.26%	10.02%	9.45%	11.60%	10.34%	52.58%
2006	0.70%	2.46%	7.58%	6.83%	12.10%	25.05%	45.29%
2007	0.96%	1.79%	6.17%	7.82%	15.31%	14.34%	53.61%
2008	0.93%	1.43%	5.39%	7.19%	16.00%	9.41%	59.66%
2009	1.47%	0.00%	4.68%	0.75%	20.99%	7.68%	64.44%
2010	0.12%	0.33%	5.44%	6.29%	17.30%	5.94%	64.59%
2011	0.00%	0.00%	6.33%	4.67%	18.32%	6.23%	64.46%
2012	0.00%	0.00%	4.42%	5.24%	17.10%	4.73%	68.50%
2013	0.00%	0.35%	3.97%	3.97%	13.87%	3.90%	73.94%
2014	0.00%	1.99%	4.03%	4.59%	18.83%	4.54%	66.03%
2015	0.00%	2.17%	4.21%	4.60%	16.74%	5.60%	66.67%
2016	0.00%	0.19%	5.44%	4.77%	15.93%	6.53%	67.15%
2017	0.00%	0.00%	7.28%	6.16%	16.10%	9.30%	61.15%
2018	0.00%	1.01%	5.54%	4.78%	13.81%	7.36%	67.51%
2019	0.00%	1.48%	5.53%	4.35%	11.98%	3.93%	72.74%

Table B.8: Optimal weights in global minimum variance portfolio with stocks, bonds, and GSCI sub-indices excluding Agriculture and Livestock

Year	Energy	Industrial	Precious	Stocks	Bonds
2003	0.96%	15.55%	12.75%	6.70%	64.03%
2004	1.35%	16.11%	20.94%	9.48%	52.12%
2005	1.57%	7.61%	13.14%	13.43%	64.25%
2006	1.32%	3.77%	9.72%	30.62%	54.57%
2007	1.88%	3.22%	8.47%	18.81%	67.61%
2008	1.85%	2.34%	8.12%	12.81%	74.88%
2009	2.82%	0.00%	5.81%	11.15%	80.22%
2010	1.53%	1.65%	7.22%	9.35%	80.25%
2011	1.22%	0.44%	8.35%	10.04%	79.94%
2012	1.15%	0.00%	5.99%	8.11%	84.76%
2013	0.59%	1.12%	5.57%	5.38%	87.35%
2014	0.95%	3.44%	6.28%	6.62%	82.72%
2015	0.00%	2.79%	5.50%	10.41%	81.30%
2016	0.47%	0.60%	8.32%	10.29%	80.32%
2017	0.50%	0.71%	9.98%	14.19%	74.62%
2018	1.15%	2.62%	6.53%	10.64%	79.06%
2019	0.52%	3.58%	7.09%	5.47%	83.35%

Table B.9: Out-of-sample performances of global minimum variance portfolios

Year	Stocks and Bonds	With GSCI	With GSCI sub-indices	Excl. Agriculture and Livestock
2003	100.00	100.00	100.00	100.00
2004	116.77	116.80	117.00	119.48
2005	129.73	130.70	126.01	134.85
2006	127.04	129.69	128.85	139.38
2007	140.76	142.60	142.53	158.07
2008	152.91	156.20	155.71	173.57
2009	156.78	155.97	147.13	175.87
2010	169.16	168.73	152.77	191.47
2011	174.21	174.36	162.86	200.79
2012	185.45	184.82	172.92	216.73
2013	192.66	191.95	177.73	224.03
2014	187.79	187.23	169.12	213.47
2015	188.30	185.13	173.16	212.14
2016	181.75	175.98	158.96	202.53
2017	183.51	178.33	157.21	206.02
2018	200.36	193.81	171.35	224.00
2019	197.75	191.53	168.16	219.98
2020	211.74	204.68	176.73	236.26

C Robustness Checks

In order to corroborate our results, we conduct three robustness checks. First, we replicate the analysis based on continuous rather than discrete returns, which tend to be closer to zero. Although the monthly returns are lower, as well, our qualitative results are not affected, except that bonds obtain an even higher weight in the standard portfolio optimization exercise, as the average continuous return on bonds exceeds that on stocks during our sample period. Importantly, our main conclusion remains that the overall GSCI lost its suitability for hedging during our sample period, while the results for the out-of-sample portfolio optimizations are hardly affected. Based on the maximum Sharpe ratio approach, the optimal portfolio with only stocks and bonds continues to outperform slightly the portfolio considering the overall GSCI, whereas the portfolio considering the GSCI component sub-indices separately exhibits by far the best performance.

As a second robustness check, we replicate the out-of-sample portfolio optimization using a rolling rather than an expanding estimation window, where the length of the estimation period used to predict the conditional correlations remains constant. As the evaluation period is shifted forward in time, the most distant are replaced with more recent observations. As before, the first estimation window is 1999–2002, and the first evaluation period is 2003. After predicting the conditional correlations for 2003, however, 1999 is dropped from the estimation window in favor of 2003 in order to predict the conditional correlations for 2004. Again, our qualitative insights are not affected. Based on the maximum Sharpe ratio approach, the optimal portfolio with stocks and bonds outperforms the optimal portfolio considering the overall GSCI, while the optimal portfolio considering the GSCI component sub-indices separately performs best.

Finally, we use a Laplace (a.k.a. double exponential) rather than a normal distribution to model the standardized residuals of the DCC model, as financial market returns often exhibit a leptokurtic distribution with positive excess kurtosis and fat tails (see Mandelbrot, 1963).[20] All results are qualitatively robust to this variation, as well. The results for these and further robustness checks are available from the authors upon request.

[20]The Laplace distribution has an excess kurtosis of 3, while its mean, median, and mode are unchanged relative to the normal distribution.