Portfolio Investment: Are Commodities Useful?

by

Lei Yan and Philip Garcia

Portfolio Investment: Are Commodities Useful?

Lei Yan
and
Philip Garcia *

Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 21-22, 2014

*Lei Yan is a PhD student and Philip Garcia is the T.A. Hieronymus Distinguished Chair in Futures Markets, both in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. Correspondence may be sent to leiyan2@illinois.edu. We thank Anton Bekkerman, Jeffrey Dorfman, Randall Fortenbery, Scott Irwin, Berna Karali, Paul Peterson, Dwight Sanders, and participants at the NCCC-134 conference for helpful discussions and comments. Any remaining errors are our responsibility alone.

Portfolio Investment: Are Commodities Useful?

This paper investigates the usefulness of commodities in investors' portfolios within a mean-variance optimization framework. The analysis differs from previous research by considering multiple investment tools including individual commodity futures contracts, three generations of commodity indices and by controlling for estimation error in portfolio optimization process. Rather generally, the results demonstrate that including individual commodities or the first- and second-generation commodity indices do little to enhance portfolio performance. Similarly, when an initial portfolio is diversified, the risk-reducing ability of agricultural commodities is much weaker than identified by previous research. In contrast, including the third-generation indices substantially improves the portfolio's Sharpe ratio by generating higher returns and lower risk.

Keywords: commodity index investments, portfolio diversification, mean-variance optimization, estimation error

Introduction

Investments in commodities have grown rapidly in the past decade through commodity index funds. The total value of commodity index investments was about $210 billion by the end of 2012.[1] This large inflow of investment funds is mainly based on the perception that investors can obtain diversification benefits by including commodity futures since commodities show equity-like returns and low correlation with traditional assets (Gorton and Rouwenhorst, 2006). The question of interest is whether including commodity futures can improve a portfolio's performance. The literature provides mixed evidence. Early studies show that investors are better off in terms of reducing risk by including commodities in portfolios (e.g., Bodie and Rosansky, 1980; Fortenbery and Hauser, 1990; Ankrim and Hensel, 1993). However, recent research fails to identify consistent diversification benefits in an out-of-sample setting (Daskalaki and Skiadopoulos, 2011; You and Daigler, 2013). Where do commodities and particularly agricultural commodities fit into investors' portfolios, and does their usefulness provide insights into future non-traditional investors in these markets?

We analyze that most academic research on the usefulness of commodities in portfolios suffers from two shortcomings. First, most studies examine the role of commodities in a portfolio by employing either a few individual commodity futures contracts or widely used indices (e.g., S&P Goldman Sachs Commodity Index). Few investigate the roles that different commodity indices and multiple agricultural commodities can play in a portfolio. Miffre (2012) argues that returns to commodity indices can greatly depend on imbedded strategies, which suggests that the usefulness of commodities may vary by the type of index considered. In addition, Commodity Index Traders (CITs) have become large participants in twelve agricultural futures markets with an average percentage of positions rising from 7% to 34% for 2000-2009 (Aulerich et al., 2013). Despite the increasing exposure to agricultural commodity markets, formal assessments on whether including agricultural commodity futures benefits a

[1] See CFTC Index Investment Data. http://www.cftc.gov/MarketReports/IndexInvestmentData/index.htm

portfolio are scarce. Second, studies that apply portfolio theory to commodity futures often ignore estimation error. Estimation error occurs whenever sample moments are used to estimate population values. Standard portfolio optimization that replaces expected returns by sample estimates often leads to concentrated and unstable asset allocations (Kan and Zhou, 2009). For example, You and Daigler (2013) find that the ex-post optimal portfolios are quite unstable and suggest that reducing bias in the expected returns would make optimal portfolio models more useful. Failure to account for these shortcomings can influence our understanding of the role that commodities play particularly during periods of high volatility which have occurred recently.

The paper re-examines the usefulness of commodities in a portfolio and contributes the literature in two dimensions. First, we evaluate the impacts on investor portfolios of twelve agricultural commodities monitored by the CFTC, and three commodity indices (S&P Goldman Sachs Commodity Index (SPGSCI), Deutsche Bank Optimum Yield Commodity Index (DBOYCI), and Morningstar Long-Short Commodity Index (MSLSCI)), which reflect the classification scheme developed by Miffre (2012). The first two indices represent passive long-only exposures to commodities, while MSLSCI allows for both long and short positions and more aggressive momentum strategies.[2] The different proxies for commodities help us determine whether the usefulness of commodities varies by the type of investment tools. Second, to control for estimation error, we use Black and Litterman's (1992) procedure that conducts portfolio optimization with shrinkage estimates. Historical returns for equities and bonds are shrunk towards a prior, which is derived from reversing the Capital Asset Pricing Model (CAPM) with equal weights. For commodities the prior returns are assumed to be zero, consistent with findings in the literature (Sanders and Irwin, 2012) and our calculations that returns to commodities are not significantly different from zero.

The data consist of monthly returns of multiple asset types for 1991-2012. We construct benchmark portfolios using U.S. equities, U.S. bonds, global equities, and global bonds, and assess the effect of adding commodity indices and individual commodities. The mean-variance optimization with sample estimates and shrinkage estimates for expected returns is implemented in both in- and out-of-sample settings. The results suggest that including individual commodities does not significantly improve the portfolio performance as measured by the Sharpe ratio. Similarly, including the first- or second-generation commodity indices fails to increase the portfolio Sharpe ratio, but evidence does emerge that the third-generation commodity indices can either increase returns or decrease risk (or both) which leads to enhanced portfolio performance. The results also confirm previous findings that standard mean-variance optimization leads to concentrated and unstable asset allocations through time and controlling for estimation error produces more diversified and balanced portfolios. However, we find that commodities play a much smaller role when the remainder of portfolio is diversified. Our results are consistent over multiple robustness analyses.

[2] Added description of the indices is provided in the text below.

Related Literature

In this section, we first review studies on commodity futures returns with focus on the difference in performance between individual futures and commodity indices. Next, we discuss the correlation between equity and commodity returns, since low correlation is generally viewed as a necessary condition for commodity to be part of a portfolio. While return and correlation are two important factors that influence the commodities' investment value, they cannot assure portfolio benefits. Consequently, we discuss research that applies a portfolio framework to evaluate the usefulness of commodities.

Literature on commodity futures returns falls into two categories. The first line of research, dating back to Keynes (1930), seeks evidence on risk premium for individual commodity futures markets. For instance, Dusak (1973) and Fama and French (1987) find only limited evidence of a constant risk premium. However, Bessembinder and Chan (1992) and Bjornson and Carter (1997), when allowing for a time-varying risk premium, do find support for a risk premium in several commodity futures markets. Despite these mixed findings, the results identify a few factors such as hedging pressure that appear to be related to risk premium. The other line of literature examines returns to a portfolio of commodities or commodity indices. For example, in a direct assessment Gorton and Rouwenhorst (2006) document that an equally-weighted commodity portfolio can achieve "equity-like" returns. In contrast, Sanders and Irwin (2012) show that returns to individual commodity futures contracts do not differ from zero, question the source of value in an equally-weighted commodity portfolio, and suggest that the superiority of commodity indices may be largely due to their imbedded strategies. The notion that differences in strategies may influence commodity portfolio performance is informative for investors and analysts seeking to understand commodity market behavior. Miffre (2012) classifies commodity indices into three generations. First-generation indices provide passive long-only exposure to broad range of commodities and include widely used ones (e.g., SPGSCI). Second-generation indices also are long-only, but improve on the first generation indices by mitigating the impact of negative roll yields from rolling positions when the market is in contango. Third-generation indices differ from the former two by allowing both long and short positions based on selected strategies. Miffre (2012) further shows that third-generation commodity indices perform best for 2008-2012. This is consistent with recent research (Fuertes et al., 2010; Szymanowska et al., 2013) that momentum and term structure based strategies can work well in commodity futures markets.

In addition to returns, correlation also plays an important role in determining portfolio performance. Commodity futures have long been considered as an isolated asset class, whose returns are uncorrelated or even negatively correlated with returns of traditional assets (Gorton and Rouwenhorst, 2006). Recently, Büyükahin et al. (2010) investigate the time-varying correlation between equity and commodity returns and find the correlation generally remains low for 1991-2008 despite a temporary modest increase during the financial turmoil. However, low correlation only implies diversification opportunity not real diversification benefits. Moreover, recent evidence points to increasing correlation between commodities (Tang and Xiong, 2012) and between commodities and other financial assets (Silvennoinen and Thorp, 2013), which calls into question the diversification benefits of commodity futures.

The investment potential of commodities also has been investigated in a portfolio choice framework. Bodie and Rosansky (1980) and Georgiev (2001) show that blending commodity futures with common stocks can reduce risk without sacrificing return for the periods 1950-1976 and 1990-2001. They form portfolios by considering a range of proportions between assets instead of using an optimization procedure. Other researchers examine this issue within the Markowitz's (1952) mean-variance optimization framework. Fortenbery and Hauser (1990) find that the addition of agricultural futures contracts to a stock portfolio rarely increases portfolio return but does reduce portfolio risk. Ankrim and Hensel (1993) use SPGSCI to represent exposure to commodities and find similar improvement on return-risk profile. Satyanarayan and Varangis (1996) augment Ankrim and Hensel's (1993) work by considering global investors with broader asset classes and find that the inclusion of commodities shifts the efficient frontier upwards. While most prior research agrees on the diversification benefits of commodities, contradictory results also exist. Daskalaki and Skiadopoulos (2011) extend analysis to an out-of-sample setting and find that the in-sample benefits, if any, by including commodities do not persist out-of-sample.[3] You and Daigler (2013) also compare the in-sample and out-of-sample performance of portfolios with and without futures contracts using mean-variance optimization. They find that portfolios with futures outperform the traditional portfolio in-sample but out-of-sample gains in performance are negligible.

The identified research on the role of commodities in investors' portfolios is subject to two shortcomings. First, studies on asset allocation with commodities examines either a few individual commodities or a passive long-only index (e.g., SPGSCI), which fails to reflect the variety of commodity indices (Miffre, 2012). Also, few examine the usefulness of including the twelve agricultural commodity futures in portfolios in light of rapid growth of CIT positions in those markets (Aulerich et al., 2013). You and Daigler (2013) examine 39 futures contracts which include these agricultural commodity futures within a classical mean-variance optimization framework, but their results may be susceptible to estimation errors. Second, previous research ignores estimation error that can arise when the sample estimates of mean and variance of returns are treated as "true" values in optimization. You and Daigler (2013) is the only study that recognizes the potential impact of estimation error. They find that the optimal portfolio weights are highly volatile which is mainly driven by errors in expected return estimates. With imprecise mean and variance as inputs, the derived optimal portfolios tend to be concentrated and extremely sensitive to the mean estimate (e.g., Michaud, 1989). The problem is worse in a more realistic out-of-sample setting because in-sample distributions of returns often do not persist out-of-sample (Kan and Zhou, 2009). Failure to control estimation error can lead to misallocated portfolios and incorrect evaluation on the role of commodities. To address these issues, we consider alternate ways of investing in commodities including twelve agricultural commodity futures and three generations of commodity indices. The use of alternative indices allows us to test whether the usefulness of

[3] Daskalaki and Skiadopoulos (2011) use different approaches to evaluate the role of commodities under in-sample and out-of-sample settings. They conduct spanning tests in-sample and implement portfolio optimization out-of-sample.

commodities depends on specific investment instruments. Employing agricultural commodity futures enables us to explore whether they can provide diversification benefits to a portfolio given increasing CIT exposures. With regards to the estimation error, different methods have been proposed (Fabozzi et al., 2007). Here, we follow Black and Litterman's (1992) approach which allows us to incorporate the priors that many economists have about the zero expected returns to agricultural futures markets.

Data

The dataset consists of monthly returns of multiple types of assets for 1991-2012. To investigate the role of commodities in a portfolio, we construct a benchmark portfolio that includes four asset types - U.S. equities, U.S. bonds, global equities, and global bonds. The U.S. equity market is represented by the S&P 500 Index - a widely used investment benchmark. The Barclays US Aggregate Bond Index is chosen to reflect the performance of investment grade bonds as it covers treasury bonds, government agency bonds, mortgage-backed bonds, corporate bonds, and a small amount of foreign bonds traded in US. Global equity performance is measured by the MSCI World Index (ex U.S.), which includes securities from 23 countries and is often used as a common benchmark for global stock funds. The JPMorgan Global Aggregate Bond Index is considered as a representative of global bonds which tracks instruments from over 60 countries.

For commodities, we consider both commodity indices and individual commodity futures contracts. In particular, we use three commodity indices - S&P Goldman Sachs Commodity Index (SPGSCI), Deutsche Bank Optimum Yield Commodities Index (DBOYCI), and Morningstar Long-Short Commodity Index (MSLSCI), which belong to the three generations classified by Miffre (2012).[4] The first-generation indices are designed to provide broad exposure to commodities, among which SPGSCI is widely used as a benchmark. These indices hold liquid contracts with the shortest maturity, which can lead to poor performance when prices of distant futures contracts are higher than near-by contract prices. The second-generation indices differ in that they take into account the shape of futures price curve and also are willing to invest in distant contracts. DBOYCI, as its name suggests, choose contracts with the maximum implied roll yield.[5,6] The first two generations indices are both passive long-only exposures to commodities. In contrast, the third-generation indices can take long and short positions based on strategies they identify. For example, strategies based on momentum and term structure signal have been shown to work well (Fuertes et al., 2010; Szymanowska et al., 2013). MSLSCI uses a momentum rule to determine its long or short

[4]These three indices are chosen because they have historical data back to 1991, which is uncommon. We also use another set of indices for a robustness check.

[5]Implied roll yields are measured as price differences between near-by and distant futures contracts scaled by maturity differences. For a market in contango, negative roll yield arises by rolling from the near-by contract to the distant contract that has higher price. DBOYCI is designed to invest into a particular distant contract that generates the smallest negative roll yield. See DBOYCI index guide. https://index.db.com/dbiqweb2/data/guides/DBLCI-OY_v15.pdf

[6]Recently, researchers have begun to explore the sources of commodity index returns and cast doubt on the role of roll yields (Willenbrock, 2011; Sanders and Irwin, 2012). Here, we abstract away from those questions and simply use the indices as commonly done in the industry.

positions every month - taking long in the subsequent month if current price exceeds its 12-month moving average and taking short otherwise.[7] Note these three commodity indices are all broad and cover major commodity sectors including energy, precious metals, industrial metals, and agriculture. The data on equities, bonds, and commodity indices come from the Bloomberg database.[8] We also consider twelve agricultural commodities, including cocoa, coffee, cotton, sugar, feeder cattle, lean hogs, live cattle, corn, soybeans, soybean oil, Chicago wheat, and Kansas wheat. These twelve commodities are monitored by the CFTC in their weekly Supplemental Commitment of Traders reports and research has shown that the CIT positions in those markets have increased pronounced since 2004. The prices of individual commodity futures come from the Commodity Research Bureau.

Monthly returns are calculated as the percentage changes of asset prices,

$$R_t = ln(\frac{p_t}{p_{t-1}}) \times 100 \tag{1}$$

where p_t is the closing price on the last trading day in month t.[9] Several points require attention in excess return construction. First, total return indices are used for equities, bonds, and commodity indices since they are target indices tracked by many funds. Excess returns are obtained by subtracting total returns from equation 1 by risk-free rate, which is represented by the one-month Libor rate. Second, investing in futures, different from investing in equities or bonds, requires no principal except margins. Futures margins are typically much smaller than the contracts' nominal values; that is, futures investments involve substantial leverage. To draw a meaningful comparison between the performance of futures and other asset classes, a full-collateralized assumption is often made (Gorton and Rouwenhorst, 2006; Sanders and Irwin, 2012). Following this convention the excess returns to futures positions are just percentage changes in futures prices as shown in equation 1.[10] Third, commodity futures returns are calculated using near-by contracts and adjusted for roll dates (see appendix of Singleton, 2014).

Table 1 provides annualized mean and standard deviation of monthly excess returns for all assets for different periods. For 1991-2012, investing in U.S. equities and bonds achieved average returns about 5.7% and 3.23% per year, while allocating funds to global equities and bonds obtained 0.4% and 3.28% per year. The average returns for three commodity indices are 0.25%, 6.12%, and 5.28%. Returns for individual commodities vary considerably, ranging from 6.38% (soybeans) to -6.24% (cotton). Compared with equities, bonds tend

[7] See MSLSCI Fact Sheet.
http://corporate.morningstar.com/US/documents/Indexes/CommodityFactsheet.pdf

[8] Second- and third-generation commodity indices are all launched after 2005, but some provide historical data based on backward computation using the same strategies. Use of these returns to make comparisons here assumes that the relative effectiveness of the strategies has not changed.

[9] p_t denotes the settlement price for futures contract.

[10] Leverage only matters in calculating return for individual commodity futures, commodity indices are designed to be fully collateralized. Investing in a particular index is typically performed through an index fund, which ties to mimic the index performance by holding positions in both commodity futures and risk-free bonds markets.

to display much lower standard deviations and most commodities (except feeder cattle and live cattle) exhibit higher standard deviations as expected. The second- and third-generation commodity indices (DBOYCI and MSLSCI) show equity-like returns and moderate standard deviations. We also report test results for the hypothesis that the mean excess returns are equal to zero. Five markets show significantly positive returns at the 10% significance level including US equities, US bonds, global bonds, the third-generation commodity index, and live cattle futures. The strong performance of MSLSCI is probably related to its embedded strategies. In contrast, all individual commodity futures returns do not differ from zero except live cattle, which is consistent with findings by Erb and Harvey (2006) and Sanders and Irwin (2012) that the historical returns for individual commodities are almost zero.[11]

Table 1 also reports the mean and standard deviation of monthly excess returns for temporal subsamples. To reflect increases in commodity prices beginning in 2004 and their collapse following financial crisis in mid-2008, we consider three subsamples - 1991-2003, 2004-2007, and 2008-2012. Wide differences in asset performance can be identified across subsamples. For example, U.S. equities performed best before 2004; global equities did best for 2004-2007; U.S. and global bonds have become more valuable since 2008. For the commodity indices, DBOYCI provides consistently higher excess return and lower standard deviation than SPGSCI. Interestingly, MSLSCI shows moderate return or loss but consistently lower standard deviation compared to the other two indices. Seven of twelve individual commodities show higher excess returns in the commodity boom period 2004-2007. Three commodities (cotton, sugar, and soybeans) perform better after 2008 and the two cattle contracts show similar performance across subsamples. The standard deviations tend to increase for most of assets post 2008. The p-values are not reported for subsamples due to limited observations.

Table 2 provides correlations between asset returns for different periods. For the entire period, U.S. equities are highly correlated with global equities (0.78) but are weakly correlated with U.S. and global bonds (0.02 and 0.09). The correlation between U.S. equities and three generations of commodity indices are 0.2, 0.23, and -0.1. For the subsamples, we find that U.S. equities are nearly uncorrelated or negatively correlated with commodities before crisis, but the correlation increases considerably in 2008-2012. This is consistent with Büyükahin et al. (2010) and Silvennoinen and Thorp (2013), who find a recent increase in correlation between equity and commodity markets. Informatively, the third-generation commodity index MSLSCI maintains negative correlation with equities even in the post-crisis period, which is a reflection of its ability to take short positions in the downside market. The last column provides the average correlation between equities or bonds return and twelve agricultural commodity futures returns.[12] Despite an increase after 2008, average correlations between traditional assets and twelve individual commodity futures are generally low. U.S. bonds are highly correlated with global bonds and almost uncorrelated with commodities throughout the sample, and global equities and bonds seem to have higher correlations with commodities

[11] Erb and Harvey (2006) also find that the live cattle shows marginally significant positive returns in 1982-2004. In part, this may be related to the marginally significant time-varying risk premium found by Frank and Garcia (2009) in the live cattle market.

[12] Since we add only one commodity to the traditional portfolio at time, the pairwise correlations between individual commodities do not matter. So, we report average values to save space.

than U.S. equities and bonds in all periods.[13]

Methods

In this section, we first briefly review the classical mean-variance optimization framework. We then discuss the importance of estimation error and introduce Black and Litterman's (1992) approach to mitigate its effect. Finally, we describe in-sample and out-of-sample implementation of portfolio optimization.

Mean-variance optimization

Markowitz's (1952) mean-variance paradigm is by far the most common framework to study portfolio choice. Consider N risky assets with an $N \times 1$ random return vector $\tilde{r}$ and a risk-free asset with known return r_f. Define excess returns as $\tilde{r}^e = \tilde{r} - r_f \iota$, $\iota = [1, 1, ..., 1]'$, and denote the expected mean and variance-covariance matrix by $\mu = E[\tilde{r}^e]$ and $\Sigma = E[(\tilde{r}^e - \mu)(\tilde{r}^e - \mu)']$. Given a weight vector $w = [w_1, w_2, ..., w_N]'$ the portfolio's expected return and variance are $w'\mu$ and $w'\Sigma w$, respectively.

When there is a risk-free asset, the two-fund separation theorem implies that all investors should hold a combination of the risk-free asset and the same portfolio of risky assets - the tangent portfolio (Tobin, 1958). The tangent portfolio is achieved by maximizing the Sharpe ratio, which is defined as the ratio of excess return and standard deviation,

$$\max_{w} \frac{w'\mu}{\sqrt{w'\Sigma w}} \tag{2}$$

$$\text{s.t. } w'\iota = 1,\ \iota' = [1, 1, ..., 1]$$

Optimal portfolio weights can be solved as,

$$w^* = \frac{1}{\iota'\Sigma^{-1}\mu}\Sigma^{-1}\mu \tag{3}$$

In the analysis we also impose non-negative constraints on asset weights, consistent with the common practice that most investors only hold long asset positions.

Estimation error and shrinkage estimation

In practice, classical mean-variance optimization does not always perform well. To implement it, sample mean and covariance are used as inputs ($\hat{\mu}$ and $\hat{\Sigma}$) to solve for optimal weights based on equation (3). There is considerable research documenting the imprecision of sample estimates (e.g., Michaud, 1989; Kan and Zhou, 2009). The general conclusions are that sample estimates involve large estimation error and the derived optimal weights inherit those errors, resulting in unreliable portfolios. In addition to being imprecise, sample

[13]The reason why commodities are more correlated with global financial markets than U.S. markets is unclear and warrants further research.

estimates tend to produce concentrated allocations, which seems to contradict the notion of diversification.

The literature has suggested different ways to control estimation error.[14] Shrinkage estimation is attractive because of its ability to incorporate investors' own beliefs. Developed by James and Stein (1961) and applied to portfolio choice by (e.g., Jorion, 1986; Ledoit and Wolf, 2003), the idea is to shrink the sample mean towards a prior, [15]

$$\mu_s = (1 - \delta)\mu_0 + \delta\hat{\mu} \tag{4}$$

where μ_s, μ_0, and $\hat{\mu}$ denote the shrinkage estimate, the prior, and the sample estimate for expected excess return. The shrinkage factor $\delta \in [0, 1]$ reflects the relative precision between sample and prior estimates. When $\delta \to 0$ the shrinkage estimate converges to the prior and when $\delta \to 1$ it converges to the sample estimate, leading to the traditional mean-variance optimization. As illustrated in table 1, the sample estimate $\hat{\mu}$ varies considerably across subsamples, but the prior μ_0 by specification will be more stable. In this sense, the shrinkage estimates provide a more conservative framework in which investors underreact to both good and bad past performance in optimizing their portfolios. There is no fixed rule to select δ. The literature often assume that the prior is more precise than the sample estimate (e.g., Drobetz, 2001). Here, we use $\delta = 0.2$ and check the sensitivity of the results to alternate values in robustness analyses.

While appealing, the shrinkage estimation requires an informative prior. Black and Litterman (1992) lend an intuitive way of specifying the prior μ_0. Originally, Black and Litterman (1992) propose a framework to combine investor's subjective views and implied returns which derived from the CAPM model. To make correspondence, our prior returns refer to the implied returns and the sample returns are used to represent investor's subjective views. The use of sample returns as a proxy of investor's views reflects the fact that many investors make decisions based only on past prices. The implied returns are derived using the reverse optimization process described in Sharpe (1974). The intuition is that the market itself is in equilibrium and investors should only deviate from the CAPM market equilibrium returns if they have reliable information on future returns. Start with the CAPM model,

$$E(\tilde{r}_i) - r_f = \beta_i(E(\tilde{r}_m) - r_f), \qquad \beta_i = \frac{cov(\tilde{r}_i, \tilde{r}_m)}{\sigma_m^2} \tag{5}$$

where $E(\tilde{r}_i)$, $E(\tilde{r}_m)$, and r_f are the expected return on asset i, the expected return on market portfolio, and the risk-free rate, respectively. $cov(\tilde{r}_i, \tilde{r}_m)$ and σ_m^2 denote the covariance between returns of asset i and market portfolio and the variance of returns of market portfolio. Let $w_m = (w_{m,1}, ..., w_{m,N})'$ be market capitalization, then $\tilde{r}_m = \sum_{j=1}^{N} w_{m,j}\tilde{r}_j$. Substitute $\tilde{r}_m$

[14] Fabozzi et al. (2007) provide a review on recent developments in robust optimization and its applications in portfolio management.

[15] Shrinkage estimation can also be applied to the covariance matrix. We focus on the mean because the prior on covariance is more difficult to establish and because estimation errors in mean have a much larger influence (Best and Grauer, 1991).

in equation (5) and put it in matrix form,

$$\mu_0 \equiv E(\tilde{r}) - r^f \iota = \gamma \Sigma w_m, \qquad \gamma = \frac{E(r_m) - r_f}{\sigma_m^2},\ \iota = [1, 1, ..., 1]' \tag{6}$$

where $\mu_0 \equiv E(\tilde{r}) - r^f \iota$ is defined as the CAPM market equilibrium return vector, Σ is the covariance matrix, w_m is the weight vector measured by market capitalization, and γ can be explained as the coefficient of risk aversion (He and Litterman, 1999). In the same way options traders imply volatility from option prices using the Black-Scholes model, the implied returns are implied from market capitalization weights and covariance matrix.

To implement equation (6), we need to specify each term on the right-hand side: Σ, γ, and w_m. The sample covariance ($\hat{\Sigma}$) can be used as an estimate for Σ. Since $E(r_m)$ is unobservable, Black and Litterman (1992) suggest calibrating γ such that the resulting implied return μ_0 (or Sharpe ratio) for particular asset looks reasonable. In the original paper, they solve γ to make sure that the Sharpe ratio of U.S. equities is equal to 0.5. If the annualized standard deviation of U.S. equities ranges between 16% and 20%, fixing Sharpe ratio at 0.5 corresponds to a total return from 8% to 10% per year. Both these numbers are reasonable and widely accepted by industry and academia, and their procedure is used to calibrate γ. To develop a structure for w_m, Black and Litterman (1992) argue for the use of market capitalization, but the difficulty is to measure the scale of any asset market. For commodity futures, market capitalization is not a meaningful concept since the net position is always equal to zero. Moreover, it is unclear whether commodity futures can be priced using the CAPM model. To circumvent these issues we directly specify the prior returns of commodities to be zero, which is consistent with Sanders and Irwin (2012) and our own findings that the average returns to commodity futures are almost zero.[16] For the other assets (U.S. equities, U.S. bonds, global equities, and global bonds), we simply use equal weights ($w_m = [\frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}]'$) to derive their prior returns, which may capture real world constraints often placed by investors to ensure a more diversified portfolio (Brentani, 2004). Once we obtain the prior (μ_0) based on equation (6), the shrinkage estimates μ_s are then calculated as the weighted average between μ_0 and $\hat{\mu}$ from equation (4).

Implementation

The impact of including commodities in a portfolio is evaluated in three steps. First, we construct sample estimates for the mean ($\hat{\mu}$) and covariance ($\hat{\Sigma}$) of expected returns. To control for estimation error the shrinkage estimate for the mean (μ_0) is also established following procedure just described.

Second, we implement the mean-variance optimization for two sets of assets - benchmark and expanded. The benchmark consists of U.S. equities, U.S. bonds, global equities, and global bonds, and the expanded portfolio includes these assets as well as commodities. One commodity (either an index or an individual commodity futures contract) is introduced at

[16]This may be inappropriate for MSLSCI since it shows significantly positive returns. Shrinking towards zero provides a conservative assessment of the impact of MSLSCI on a portfolio.

a time. We explore the role of commodity under both in-sample and out-of-sample settings. Most of the literature falls into the in-sample category - deriving optimal portfolios and assessing their performance based on the same sample period. Out-of-sample analysis may be more meaningful to investors who are concerned more about future performance. We follow Daskalaki and Skiadopoulos (2011) in assessing out-of-sample portfolio performance with a monthly rebalancing scheme. At the end of each month, the previous K-months returns are used to generate the mean and covariance estimates, which are then used for mean-variance maximization to derive optimal weights. The portfolio's return in the next month is product of optimal weights and the month $K+1$ return vector. This process is repeated until the end of the sample is reached. By doing so, we obtain a series of monthly portfolio returns, which are used for evaluation. Note we use K observations for estimation and the portfolio evaluation is actually for 1996-2012 instead the whole period 1991-2012. Estimation window K is set at 60 months and alternative sizes of $K = 36,\ 48,\ 72$ are checked for robustness.[17]

Finally, we compare the performance between benchmark and expanded portfolios to assess if including commodities is beneficial. In particular, we calculate the portfolio's annualized excess return (Mean), annualized standard deviation (SD), and Sharpe ratio (Sharpe). If the expanded portfolio achieves a higher Sharpe ratio, we conclude that including commodities is beneficial to portfolios. Optimal weights are also examined to fully reveal the role of commodities in a portfolio.

Results

In-sample performance

Table 3 compares in-sample performance of the benchmark and expanded portfolios. Panel A is based on mean-variance optimization with sample estimates for expected returns. The benchmark portfolio achieves a Sharpe ratio 0.98 with annualized excess return and standard deviation 3.45% and 3.54%. The first two generations of commodity indices, represented by SPGSCI and DBOYCI, have limited influence on the portfolio, while including the third-generation commodity index MSLSCI improves return-risk profile, resulting in a substantial enhancement in Sharpe ratio. For many individual commodities (cocoa, coffee, cotton,lean hogs, corn, soybean oil, wheat CBOT, wheat KBOT) Sharpe ratios of expanded portfolios do not change from those of the benchmark because these commodities do not enter portfolio. The exceptions are feeder cattle and live cattle, which is partly explained by their marginally significant positive returns identified earlier. Since including twelve individual agricultural commodity futures fail to improve portfolio performance except for feeder cattle and live cattle, we only report those two in table 3 for conciseness.

Panel B shows the results from mean-variance maximization with shrinkage estimates. Compared with results in panel A, the means are similar but the standard deviations have increased substantially, leading to a drop in the Sharpe ratios. Including the third-generation

[17] In a forecasting context, one may be interested in the optimal selection of K in terms of trade-off between reliability and flexibility especially in the presence of structural breaks. Selecting optimal estimation window is not considered in the paper. Instead, we show the effect of different lengths on the results.

commodity index MSLSCI still improves the portfolio performance relative to the benchmark, but the effects of individual commodities on the portfolio are negligible.

These results can be further understood by examining the optimal portfolio weights. Table 4, panel A shows the optimal weights based on in-sample mean-variance maximization with sample and shrinkage estimates. With sample estimates, the optimal portfolios are dominated by U.S. bonds with a weight around 90%. Global equities and bonds are entirely excluded. This confirms the argument that classical mean-variance tends to generate concentrated asset allocations (e.g., Michaud, 1989). Of the three commodity indices, SPGSCI plays no role in the portfolio, DBOYCI takes 5.9%, while MSLSCI has the largest proportion 16.4%. On average, the individual commodities have a small part in the portfolio (1.7%). In contrast, the optimal weights based on optimization with shrinkage estimates are more balanced. The proportion of U.S. bonds declines to about 37%, and both U.S. equities and global bonds play larger roles in the portfolio. The increased share of equities raises both return and standard deviation of the whole portfolio which was seen in table 3. Weights on commodities are reduced, suggesting commodities are even less useful in more diversified portfolio. These small weights are consistent with findings in table 3 that individual commodities have limited impacts on the portfolios.

Out-of-sample performance

Table 5 shows the out-of-sample performance of benchmark and expanded portfolios using both sample and shrinkage estimates. Recall the estimates are generated using past 60-months observations and the portfolio is rebalanced monthly. To test whether the differences between Sharpe ratios from the benchmark and expanded portfolios are significant, we use the test proposed by Ledoit and Wolf (2008) which has demonstrated robust finite sample properties when returns are non-i.i.d..

Panel A shows the out-of-sample performance based on mean-variance optimization with sample estimates. The Sharpe ratio for the benchmark portfolio is 0.31, which does not differ statistically from those of the expanded portfolios except for the third-generation commodity index. Including MSLSCI achieves higher returns, lower risk, and a significantly higher Sharpe ratio (0.67). With regards to the individual commodities, ten of twelve reduce the portfolio's standard deviation, although they also decrease the returns. Again, we report expanded portfolios by feeder cattle and live cattle as they improve the Sharpe ratios though insignificantly. The standard deviation of benchmark portfolio is 6.78%, and the average standard deviation of portfolios expanded by those ten commodities is 6.56%, a 0.22% average reduction in portfolio risk. This risk reduction result is consistent with findings by early literature that commodities contribute to the portfolio by reducing the volatility (e.g., Fortenbery and Hauser, 1990).

Panel B shows results based on mean-variance optimization with shrinkage estimators. MSLSCI still improves the portfolio's Sharpe ratio (from 0.29 to 0.4), reflecting higher returns, lower risk, and negative correlation with traditional assets identified earlier. Compared with panel A, commodities appear to play a smaller role in portfolios. First, the third-generation commodity index MSLSCI raises the portfolio Sharpe ratio from 0.29 to 0.4, which is smaller in magnitude than the Sharpe ratio increase (from 0.31 to 0.67) when sample estimates are used. However, this Sharpe ratio increase becomes insignificant with p-value of 0.21. Second, the second-generation commodity index DBOYCI and two agricultural commodity futures (feeder cattle and live cattle) have no impact on the Sharpe ratio. Third, the risk-reducing ability of agricultural commodity futures is also weaker. Although ten of twelve expanded portfolios show some degree of reduction in standard deviation, the magnitude decreases from 0.22% to 0.05% when shrinkage estimates are used.

These findings can be further understood by inspecting the optimal portfolio weights. Table 4, panel B reports the out-of-sample optimal weights and their standard deviations. Each value represents the average optimal weights over time for a particular asset and values in the last column denote an average over twelve individual commodities for the period. With sample estimates, U.S. bonds has a weight close to 54%, much larger than weights estimated for the other assets. SPGSCI assumes only 3.1%, but the other two commodity indices are weighted heavily, reaching 12.8% and 22.7%. The individual commodities are on average 3.5%. With shrinkage estimates, the portfolios are more diversified. The weight of U.S. bonds is less than 34% and other assets contribute more in the portfolio. Note the weights of the commodity indices and composites are all reduced, suggesting weaker impacts. Standard deviations of optimal weights are shown in parentheses. The variability of optimal weights is uniformly smaller when shrinkage estimates are used, which again confirms that shrinkage estimation generate stable portfolios in an out-of-sample environment.

Robust analysis

We provide several robustness checks.[18]. First, we use different indices to represent the three generations of commodity indices. Second, we consider including all twelve agricultural commodity futures in the portfolio instead of just one at a time. Third, we examine alternative values of the shrinkage factor δ. Fourth, we examine how the usefulness of commodities changes over time by repeating analysis for different periods. Furthermore, we investigate the impact of futures leverage, different sizes of estimation window, and transaction costs.

Alternate commodity indices

We consider a different set of commodity indices - Dow Jones UBS Commodity Index (DJUBS), Merrill Lynch Commodity Index (MLCI), and CYD Long-Short Commodity Index (CYDLS), which enables examining whether the usefulness of commodity indices varies by the index type. DJUBS is a first generation index similar to SPGSCI but differs by being less highly concentrated on energy and by imposing upper bounds on individual commodity (15%) and sector (33%) to ensure diversification. MLCI is a second generation index similar to DBOYCI which reduces the negative effects of rolling contracts when the market is in

[18] Due to space limit, we only report a portion of the results. The complete set of findings are available from the authors. All robust tests are conducted in an out-of-sample setting since the in-sample portfolio is dominated by bonds and is less useful.

contango.[19] Similar to MLSCI, the CYDLS index also allows for long and short positions depending on the term structural signals.[20]

Table 6 shows the out-of-sample portfolio performance with sample and shrinkage estimates. For both the sample and shrinkage estimates, the DJUBS and MLCI indices fail to improve the Sharpe ratio, but including CYDLS does increase the Sharpe ratio though insignificantly from 0.31 to 0.58 with the sample estimates. Again, the benefit from Sharpe ratio increase disappears when shrinkage estimates are used.

Multiple commodities

In table 7, we include all twelve individual commodities in the portfolio, which enables us to check whether there exists any subset of commodities that jointly benefit the portfolio. Panel A shows the performance measures for benchmark and expanded portfolios. Based on mean-variance optimization with sample estimates, the expanded portfolio has a lower Shapre ratio (0.20) compared to the benchmark portfolio (0.31). Although including commodities can reduce standard deviation from 6.78% to 6.33%, it decreases the return even more from 2.13% to 1.27%. Similar results arise using the shrinkage estimates. Panel B provides the average optimal weights for each asset. With sample estimates, the U.S. bonds are 55% of the portfolio, followed by global bonds (21%), U.S. equities (19%), and global equities (5%). Once commodities are included, the weights of benchmark assets drop, and eight commodities enter the portfolio with a total weight about 21%. If the shrinkage estimates are used, the portfolios become more diversified. The total weight of commodities is reduced to 9%, indicating a less important role in the portfolio. On balance, including multiple commodities does not improve the portfolio performance.

Shrinkage factor

We examine the impact of shrinkage factor δ by changing it from 0.2 to 0.5, in which case the sample estimates and the prior are equally weighted. Since larger δ represents a more confident for the sample estimates, it is expected that the results using shrinkage estimates will look closer to those using sample estimates. The general conclusions are similar. All commodities fail to improve the portfolio's Sharpe ratio except MSLSCI. Including individual commodities seems to reduce the portfolio's standard deviation, but the magnitude is smaller with shrinkage estimates.

Subsamples

To investigate the temporal usefulness of commodities, we split the whole period into three subsamples, 1991-2003, 2004-2007, and 2008-2012. Period 2004-2007 represents the commodity price boom period, while 2008-2012 includes the market meltdown and subsequent period in which cross-market correlation is strengthened. Recall that the out-of-sample optimization needs 60-month observations for parameter estimation. We use returns for

[19]See Merrill Lynch commodity paper. http://www.ml.com/media/67354.pdf

[20]See CYD indices overview. http://www.cyd-research.com/en/indices/longshort_tr_index.php

1991-1995, 1999-2003, and 2003-2007 to start the optimization for each subsamples. Table 8, panel A reports the Sharpe ratios derived from optimization with sample estimates. As expected, the first- and second-generation commodity indices (SPGSCI and DBOYCI) increases portfolio Sharpe ratio only in the commodity boom period 2004-2007. In contrast, the third-generation commodity index MSLSCI improves the portfolio performance in the other periods. Including agricultural commodity futures barely enhances portfolio performance in any periods except feeder cattle and live cattle. When shrinkage estimates are used (table 8, panel B), benefits to commodities almost disappear except for the third-generation commodity index in 1991-2003.

Other robustness tests

In addition, we consider the impacts of full-collateralization, length of estimation window, and transaction costs. First, we relax the full-collateralized assumption since trading individual commodity futures does not require principal except a small margin. Following Egelkraut et al. (2005), we assume a futures margin as 10% of the contract value.[21] Given the leverage, the futures return is ten times as the usual return measured by percentage price changes. Using these levered returns we re-examine the role of twelve individual commodity futures and find similar results. The reason is that the leverage amplifies return and volatility simultaneously, leading to no difference in portfolio performance. We also consider alternative sizes of estimation window (K) in the out-of-sample optimization. The initial K is specified as 60 months. We check for the cases $K = 36,\ 48,\ 72$ and the results are rather robust.

To include transaction costs, we define the net-of-transaction-cost returns as,

$$r_{nc,t+1} = (1 + r_{c,t+1})[1 - c \times \sum_{j=1}^{N}(|w_{j,t+1} - w_{j,t}|)] \tag{7}$$

where $r_{c,t+1}$ and $r_{nc,t+1}$ are portfolio returns before and after the transaction costs in period $t+1$, c denotes the transaction cost vector, and $\sum_{j=1}^{N}(|w_{j,t+1} - w_{j,t}|)$ measures the amount of portfolio rebalanced. Since the transaction costs may differ between asset types, we follow Daskalaki and Skiadopoulos (2011) by setting c equal to 50 basis points (0.5%) per transaction for equities and bonds and 35 basis points (0.35%) for individual commodity futures contracts.[22] For commodity indices we set it at 100 basis point (1%) since the fee charged by most commodity index funds ranges from 0.75% to 1.5%. We obtain qualitatively similar conclusions when allowing for transaction costs.

[21] The actual margins are not available for all commodities. Analyzing several commodities including corn, soybeans, wheat, live cattle, and lean hogs, we find that the historical margins fluctuate about 5-10%. See historical margins from CME. http://www.cmegroup.com/clearing/risk-management/historical-margins.html

[22] Daskalaki and Skiadopoulos (2011) establish the transaction costs levels based on discussion with practitioners. Trading commodities may induce much lower transaction costs (Fuertes et al., 2010). To ensure the robustness, we have also considered lower cost values for trading commodities and the results remain similar.

Concluding remarks

This paper examines whether investors benefit by including commodities in their portfolios. We differ from previous literature in two aspects. First, we evaluate whether the usefulness of commodities varies by the type of investment tool. Specifically, we use the twelve agricultural commodities monitored by the CFTC, and three commodity indices, which allow for a broader range of commodities and represent the three generations of commodity indices developed by Miffre (2012). Second, we explicitly control for estimation error in the process of optimization, which has been ignored in prior research on this topic. Estimation error has been shown to induce concentrated and unstable portfolios, which may mislead the evaluation on the role of commodities. To reduce estimation error we shrink the sample mean estimates to a prior following Black and Litterman's (1992) approach. For commodities the prior returns are assumed to be zero, consistent with findings that returns to commodities do not differ from zero (Sanders and Irwin, 2012), while the prior returns of traditional components - U.S. and global equities and bonds - are derived from reversing the CAPM with equal weights.

Overall, the findings indicate that including individual agricultural commodity futures or the first- or second-generation commodity indices fail to significantly improve portfolio Sharpe ratios in both in-sample and out-of-sample analyses. In contrast, the third-generation commodity indices improve the portfolio performance due to their imbedding strategies related to momentum and term structure signals. Accounting for estimation error with shrinkage estimates, we find that the optimal asset allocations are more diversified and stable and that commodities play a smaller role in a more diversified portfolio. Specifically, while including commodities can reduce volatility in highly concentrated portfolios, their risk-reducing effects almost disappear in the more diversified portfolios considered here. In general, the results are robust to alternative commodity indices, degree of shrinkage, and including multiple commodities together. We also relax the full-collateralized assumption, take into account transaction costs, change sizes of estimation window, and perform the analysis over subsamples, and find similar results.

Our findings are largely consistent with recent studies on the role of commodities in investor's portfolios. Using out-of-sample utility maximization, Daskalaki and Skiadopoulos (2011) find no significant improvements in Sharpe ratios by including popular commodity indices (SPGSCI and DJUBS) and five individual commodities (cotton, crude oil, copper, gold, and live cattle) in a portfolio. Similarly, You and Daigler (2013) examine a number of commodity and financial futures in a mean-variance framework and find that the portfolio with futures contracts outperforms the traditional portfolio in-sample but improved performance does not continue in an out-of-sample context. Here, we identify limited improvement in Sharpe ratios for some commodities (e.g., feeder and live cattle) in the in-sample analysis, but no out-of-sample evidence of statistically significant portfolio improvement for both individual commodities and the first-generation indices emerges. We expand the literature by considering the second- and third-generation commodity indices and find that the third-generation commodity indices can improve portfolio Sharpe ratios significantly in most cases. This finding highlights the importance of commodity index categorization (Miffre, 2012) when studying portfolio investment by researchers as well as investors.

Our findings contrast to some degree with the view that including commodities in portfolios will be beneficial to investors by reducing volatility (e.g., Bodie and Rosansky, 1980; Fortenbery and Hauser, 1990; Ankrim and Hensel, 1993). While commodities can marginally reduce risk in concentrated portfolios, we show that in more balanced portfolios the risk reducing ability of commodities is negligible. Here, the more balanced portfolios resulted from using shrinkage estimates to reduce estimation error. You and Daigler's (2013) conjecture that the failure to account for estimation error, which can lead to concentrated portfolio allocations, might influence our understanding of the role commodities play in portfolios was perceptive. Our analysis demonstrates that in terms of risk mitigation their role is reduced.

These findings have practical implications in a broader investment context and implications for market behavior. Investors often place constraints on portfolio allocations to diversify their risk, and the notion that more balanced portfolios makes commodities less useful may influence allocations and the presence of some financial investors in commodity markets. Our findings seem to be consistent with recent actions by some large pension funds. For example, the California Public Employees Retirement System (CalPERS), the largest U.S. pension fund, reduced its commodity allocation from 1.4% to 0.6% in October 2012 and to 0.5% in early 2013. The same has been done by the California State Teachers' Retirement System and the Ohio Police & Fire Pension Fund.[23] On the market level, since the top ten largest exchanged-traded products (ETPs) on broad commodities track the first- or second-generation commodity indices according to the 2013 Q1 report by *ETF Securities*,[24] the growth of commodity index investments may slow given few benefits to portfolios. However, investors may shift investment tools. Since some of the third-generation commodity indices produce better performance, it is likely that investors may move towards third-generation indices and be selective in identifying those indices that provides most attractive performance. In a market context, this suggests that future non-traditional investors in commodity markets may be more informed and selective in their market activities.

[23] http://www.pionline.com/article/20130218/PRINT/302189978/investors-get-active-over-commodities

[24] http://www.etfsecurities.com/Documents/Global%20Commodity%20ETP%20Quarterly%20-%20Q1%202013%Europe.pdf

References

Ankrim, E. M. and Hensel, C. R. (1993). Commodities in asset allocation: A real-asset alternative to real estate? *Financial Analysts Journal*, 49(3):20–29.

Aulerich, N. M., Irwin, S. H., and Garcia, P. (2013). Bubbles, food prices, and speculation: evidence from the CFTCs daily large trader data files. *Working paper*.

Bessembinder, H. and Chan, K. (1992). Time-varying risk premia and forecastable returns in futures markets. *Journal of Financial Economics*, 32(2):149–168.

Best, M. and Grauer, R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational retults. *Review of Financial Studies*, 4(2):315–342.

Bjornson, B. and Carter, C. (1997). New evidence on agricultural commodity return performance under time-varying risk. *American Journal of Agricultural Economics*, 79(3):918–930.

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43.

Bodie, Z. and Rosansky, V. (1980). Risk and return in commodity futures. *Financial Analysts Journal*, 36(3):27–39.

Brentani, C. (2004). *Portfolio Management in Practice*. Essential Capital Markets. Elsevier Science.

Büyükahin, B., Haigh, M., and Robe, M. (2010). Commodities and equities: Ever a market of one? *Journal of Alternative Investments*, 12(3):76–95.

Daskalaki, C. and Skiadopoulos, G. (2011). Should investors include commodities in their portfolios after all? New evidence. *Journal of Banking & Finance*, 35(10):2606–2626.

Drobetz, W. (2001). How to avoid the pitfalls in portfolio optimization? Putting the Black-Litterman approach at work. *Financial Markets and Portfolio Management*, 15(1):59–75.

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

Egelkraut, T., Woodard, J., Garcia, P., and Pennings, J. M. E. (2005). Portfolio diversification with commodity futures: Properties of leveraged futures. *NCR-134 Conference paper*.

Erb, C. and Harvey, C. (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 62(2):69–98.

Fabozzi, F. J., Kolm, P. N., Pachamanova, D., and Focardi, S. M. (2007). Robust portfolio optimization. *Journal of Portfolio Management*, 33(3):40–48.

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

Fortenbery, T. and Hauser, R. (1990). Investment potentical of agricultural futures contracts. *American Journal of Agricultural Economics*, 72(3):721–728.

Frank, J. and Garcia, P. (2009). Time-varying risk premium: further evidence in agricultural futures markets. *Applied Economics*, 41(6):715–725.

Fuertes, A.-M., Miffre, J., and Rallis, G. (2010). Tactical allocation in commodity futures markets: Combining momentum and term structure signals. *Journal of Banking & Finance*, 34(10):2530–2548.

Georgiev, G. (2001). Benefits of commodity investment. *Journal of Alternative Investments*, 4(1):40–48.

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

He, G. and Litterman, R. (1999). The intuition behind Black-Litterman model portfolios. *Goldman Sachs Asset Management working paper.*

James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 361–379. University of California Press.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292.

Kan, R. and Zhou, G. (2009). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(03):621–656.

Keynes, J. M. (1930). *A Treatise on Money.* London: Macmillan.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603–621.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

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

Michaud, R. O. (1989). The Markowitz optimization enigma: Is the optimized optimal? *Financial Analysts Journal*, 45(1):31–42.

Miffre, J. (2012). Comparing first, second and third generation commodity indices. *EDHEC working paper.*

Sanders, D. R. and Irwin, S. H. (2012). A reappraisal of investing in commodity futures markets. *Applied Economic Perspectives and Policy*, 34(3):515–530.

Satyanarayan, S. and Varangis, P. (1996). Diversification benefits of commodity assets in global portfolios. *Journal of Investing*, 5(1):69–78.

Sharpe, W. F. (1974). Imputing expected security returns from porfolio composition. *Journal of Financial and Quantitative Analysis*, 9(3):463–472.

Silvennoinen, A. and Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, 24(0):42–65.

Singleton, K. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60(2):300–318.

Szymanowska, M., De Roon, F., Nijman, T., and den Goorbergh, R. V. (2013). An anatomy of commodity futures risk premia. *Journal of Finance, forthcoming*.

Tang, K. and 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.

Willenbrock, S. (2011). Diversification return, portfolio rebalancing, and the commodity return puzzle. *Financial Analysts Journal*, 67(4):42–49.

You, L. and Daigler, R. (2013). A Markowitz optimization of commodity futures portfolios. *Journal of Futures Markets*, 33(4):343–368.

Table 1: Summary statistics on asset excess returns

	1991 - 2012			1991 - 2003		2004 - 2007		2008 - 2012	
	Mean	*SD*	*p*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
U.S. equities	5.70	14.76	0.07	7.30	13.86	5.45	8.56	1.83	20.09
U.S. bonds	3.23	3.57	0.00	3.32	3.58	0.79	3.12	4.89	3.83
Global equities	0.40	16.93	0.91	-0.36	14.72	11.05	10.19	-6.05	24.67
Global bonds	3.28	5.92	0.01	3.28	5.89	1.65	4.96	4.57	6.70
SPGSCI	0.25	21.53	0.96	0.89	17.95	8.86	21.49	-8.17	28.90
DBOYCI	6.12	17.41	0.10	4.29	14.35	22.47	14.42	-2.08	24.79
MSLSCI	5.28	10.67	0.02	6.30	8.82	6.80	11.49	1.48	13.97
Cocoa	-2.38	28.88	0.70	-7.90	27.38	9.53	25.05	2.34	34.98
Coffee	-2.94	35.39	0.70	-5.10	38.85	6.46	26.10	-4.81	32.71
Cotton	-6.24	29.07	0.31	-6.01	25.84	-21.82	25.15	5.44	38.35
Sugar	4.41	32.27	0.52	4.79	29.06	-1.21	27.32	7.87	42.57
Feeder cattle	3.43	11.69	0.17	3.14	11.20	8.65	11.88	0.04	12.80
Lean hogs	-1.89	27.42	0.75	-2.66	29.30	2.98	24.58	-3.76	24.82
Live cattle	5.52	14.09	0.07	5.16	14.90	7.73	13.19	4.68	12.81
Corn	-4.31	25.48	0.43	-5.56	19.65	-7.10	29.60	1.09	34.18
Soybeans	6.38	23.90	0.21	5.15	18.93	3.57	29.90	11.71	29.75
Soybean oil	0.05	24.56	0.99	0.26	20.24	7.22	27.14	-6.12	31.82
Wheat CBOT	-3.13	28.19	0.60	-0.10	24.76	1.70	25.04	-14.67	37.51
Wheat KBOT	3.71	26.78	0.52	4.71	22.67	17.73	23.14	-9.89	37.16

Note: Excess returns are defined as the percentage changes of monthly prices in excess of risk-free rate. *Mean* and *SD* denote the average annualized excess returns and standard deviations in percentage formats. The p-value is for a two-tailed t-test that the mean equals to zero.

Table 2: Correlation between asset excess returns

		U.S. bonds	Global equities	Global bonds	SPGSCI	DBOYCI	MSLSCI	Comm-odities
U.S. equities	1991-2012	0.02	0.78	0.09	0.20	0.23	-0.10	0.14
	1991-2003	0.03	0.69	-0.01	-0.02	0.00	-0.14	0.06
	2004-2007	-0.14	0.70	-0.09	-0.24	-0.11	-0.24	0.08
	2008-2012	0.06	0.91	0.30	0.57	0.55	-0.03	0.29
U.S. bonds	1991-2012		0.06	0.68	-0.01	-0.02	-0.08	0.04
	1991-2003		0.02	0.66	0.04	-0.04	-0.02	0.02
	2004-2007		0.02	0.68	0.02	0.12	0.12	0.02
	2008-2012		0.15	0.74	-0.07	0.01	-0.28	0.10
Global equities	1991-2012			0.32	0.37	0.40	0.03	0.16
	1991-2003			0.23	0.18	0.17	0.01	0.04
	2004-2007			0.27	0.09	0.22	0.12	0.07
	2008-2012			0.48	0.63	0.62	0.01	0.33
Global bonds	1991-2012				0.14	0.16	0.05	0.07
	1991-2003				0.12	0.07	0.04	0.00
	2004-2007				0.20	0.30	0.23	0.04
	2008-2012				0.16	0.26	-0.02	0.20

Note: The last column reports averaged correlations across twelve individual commodity futures.

Table 3: Portfolio performance comparison: In-sample, 1991-2012

	Benchmark portfolio	Expanded portfolio				
		SPGSCI	DBOYCI	MSLSCI	Feeder cattle	Live cattle
Panel A: Mean-variance optimization with sample estimates						
Mean %	3.45	3.45	3.57	3.77	3.44	3.61
SD %	3.54	3.54	3.51	3.27	3.41	3.47
Sharpe	0.98	0.98	1.02	1.15	1.01	1.04
Panel B: Mean-variance optimization with shrinkage estimates						
Mean %	3.44	3.44	3.44	3.62	3.44	3.45
SD %	6.70	6.71	6.71	6.12	6.71	6.71
Sharpe	0.51	0.51	0.51	0.59	0.51	0.51

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds and is expanded by adding one commodity at a time. *Mean* and *SD* are based on the annualized portfolio excess returns. Sharpe ratio (*Sharpe*) is defined as the ratio between *Mean* and *SD*. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.

Table 4: Optimal portfolio weights: In-sample and out-of-sample

	Sample estimates					Shrinkage estimates				
	Benchmark	SPGSCI	DBOYCI	MSLSCI	Commodities	Benchmark	SPGSCI	DBOYCI	MSLSCI	Commodities
Panel A: In-sample										
US equities	0.090	0.090	0.070	0.082	0.087	0.254	0.254	0.254	0.244	0.254
US bonds	0.910	0.910	0.872	0.753	0.896	0.374	0.374	0.374	0.375	0.373
Global equities	0.000	0.000	0.000	0.000	0.000	0.149	0.150	0.150	0.127	0.150
Global bonds	0.000	0.000	0.000	0.000	0.000	0.222	0.222	0.222	0.186	0.223
Commodities		0.000	0.059	0.164	0.017		0.000	0.000	0.069	0.000
Panel B: Out-of-sample										
US equities	0.189	0.195	0.188	0.138	0.177	0.266	0.268	0.267	0.258	0.263
	(0.276)	(0.273)	(0.239)	(0.184)	(0.257)	(0.070)	(0.068)	(0.068)	(0.045)	(0.069)
US bonds	0.546	0.542	0.601	0.513	0.564	0.338	0.340	0.339	0.341	0.332
	(0.332)	(0.322)	(0.342)	(0.297)	(0.360)	(0.128)	(0.127)	(0.127)	(0.124)	(0.129)
Global equities	0.050	0.038	0.009	0.032	0.047	0.167	0.164	0.163	0.126	0.166
	(0.128)	(0.098)	(0.028)	(0.087)	(0.120)	(0.071)	(0.067)	(0.068)	(0.069)	(0.070)
Global bonds	0.215	0.194	0.075	0.090	0.178	0.230	0.226	0.227	0.184	0.231
	(0.199)	(0.192)	(0.131)	(0.136)	(0.204)	(0.080)	(0.079)	(0.079)	(0.078)	(0.081)
Commodities		0.031	0.128	0.227	0.035		0.003	0.004	0.091	0.009
		(0.041)	(0.108)	(0.114)	(0.044)		(0.009)	(0.010)	(0.057)	(0.013)

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds and is expanded by adding one commodity at a time. For twelve individual commodities, the weights are further averaged across commodities and reported in the last column. Standard deviations of weights over periods are reported in the parenthesis. The out-of-sample optimization is based on 60-months estimation window ($K = 60$) and monthly rebalancing, and the weights are reported as averages over periods. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.

Table 5: Portfolio performance comparison: Out-sample, 1991-2012

	Benchmark portfolio	Expanded portfolio				
		SPGSCI	DBOYCI	MSLSCI	Feeder cattle	Live cattle
Panel A: Mean-variance optimization with sample estimates						
Mean %	2.13	1.90	2.32	3.62	2.40	2.74
SD %	6.78	6.86	7.36	5.40	6.60	6.48
Sharpe	0.31	0.28	0.32	0.67	0.36	0.42
p-value		0.67	0.99	0.05	0.50	0.17
Panel B: Mean-variance optimization with shrinkage estimates						
Mean %	2.24	2.22	2.26	2.79	2.15	2.35
SD %	7.77	7.76	7.76	7.00	7.66	7.70
Sharpe	0.29	0.29	0.29	0.40	0.28	0.31
p-value		0.62	0.73	0.21	0.72	0.73

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds, and is expanded by adding one commodity at a time. $Mean$ and SD are based on the annualized portfolio excess returns. Sharpe ratio ($Sharpe$) is defined as the ratio between $Mean$ and SD. The p-values are computed based on Ledoit and Wolf (2008) with the null hypothesis that the Sharpe ratio obtained from benchmark portfolio is equal to that derived from expanded portfolios. The out-of-sample optimization is based on a 60-month estimation window ($K = 60$) and monthly rebalancing. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.

Table 6: Portfolio performance comparison: Alternate commodity indices, 1991-2012

	Sample estimates				Shrinkage estimates			
	Benchmark	DJUBS	MLCI	CYDLS	Benchmark	DJUBS	MLCI	CYDLS
Mean %	2.13	1.54	2.34	3.01	2.24	2.24	2.27	1.92
SD %	6.78	6.52	6.98	5.19	7.76	7.76	7.76	7.00
Sharpe	0.31	0.24	0.34	0.58	0.29	0.29	0.29	0.27
p-value		0.38	0.92	0.17		0.79	0.57	0.76

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds, and is expanded by including a different set of commodity indices - Dow Jones UBS Commodity index (DJUBS), Merrill Lynch Commodity index (MLCI), and CYD Long-Short Commodity index (CYDLS). $Mean$ and SD are based on the annualized portfolio excess returns. Sharpe ratio ($Sharpe$) is defined as the ratio between $Mean$ and SD. The p-values are computed based on Ledoit and Wolf (2008) with the null hypothesis that the Sharpe ratio obtained from benchmark portfolio is equal to that derived from expanded portfolios. The out-of-sample optimization is based on a 60-month estimation window ($K = 60$) and monthly rebalancing. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.

Table 7: Portfolio performance comparison: Multiple commodities, 1991-2012

	Sample estimates		Shrinkage estimates	
	Benchmark	Expanded	Benchmark	Expanded
Panel A: Portfolio performance				
Mean %	2.13	1.27	2.24	1.90
SD %	6.78	6.33	7.76	7.50
Sharpe	0.31	0.20	0.29	0.25
p-value		0.50		0.54
Panel B: Optimal weights %				
U.S. equities	0.19	0.11	0.27	0.24
U.S. bonds	0.55	0.47	0.34	0.28
Global equities	0.05	0.03	0.17	0.15
Global bonds	0.21	0.17	0.23	0.25
Commodities (total)		0.21		0.09
Cocoa		0.04		0.02
Coffee		0.01		0.01
Cotton		0.00		0.00
Sugar		0.02		0.01
Feeder cattle		0.05		0.02
Hogs		0.00		0.01
Live cattle		0.05		0.01
Corn		0.00		0.00
Soybeans		0.02		0.00
Soy oil		0.01		0.01
Wheat CBOT		0.00		0.00
Wheat KBOT		0.01		0.00

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds, and is expanded by including all twelve individual commodities. $Mean$ and SD are based on the annualized portfolio excess returns. Sharpe ratio ($Sharpe$) is defined as the ratio between $Mean$ and SD. The p-values are computed based on Ledoit and Wolf (2008) with the null hypothesis that the Sharpe ratio obtained from benchmark portfolio is equal to that derived from expanded portfolios. The out-of-sample optimization is based on a 60-month estimation window ($K = 60$) and monthly rebalancing. Optimal weights are averaged across periods. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.

Table 8: Sharpe ratio performance: Subsamples

	Benchmark portfolio	Expanded portfolio				
		SPGSCI	DBOYCI	MSLSCI	Feeder cattle	Live cattle
Panel A: Mean-variance optimization with sample estimates						
1991-2003	0.54	0.53	0.54	1.10	0.56	0.60
2004-2007	0.20	0.28	1.05	0.47	0.33	0.39
2008-2012	0.01	-0.13	-0.19	0.07	0.08	0.16
Panel B: Mean-variance optimization with shrinkage estimates						
1991-2003	0.30	0.30	0.31	0.61	0.31	0.33
2004-2007	0.88	0.86	0.88	0.93	0.86	0.91
2008-2012	0.12	0.12	0.11	0.05	0.11	0.13

Note: Benchmark portfolio consists of U.S. equities, U.S. bonds, global equities, and global bonds, and is expanded by one commodity at a time.The out-of-sample optimization is based on a 60-month estimation window ($K = 60$) and monthly rebalancing. The shrinkage factor δ is assumed to be 0.2 when shrinkage estimates are used.