Returns to Investing in Commodity Futures: Separating the Wheat from the Chaff

Scott H. Irwin, Dwight R. Sanders, Aaron Smith, and Scott Main*

January 14, 2019

Abstract: *Commodity futures investment grew rapidly after its popularity exploded—along with commodity prices—in the mid-2000s. Numerous individuals and institutions embraced commodity investments for their purported diversification benefits and equity-like returns; yet, real-time performance was disappointing. We investigate the reasons for the puzzling returns of commodity futures investments in the last decade. Our analysis shows that the disappointing commodity returns were not driven mechanically by contango or negative "roll yields." Six decades of daily return data suggest that the expected excess return in individual commodity futures markets is near zero before expenses. This implies that net losses to commodity futures investment will be equal to order execution and operating costs estimated at 3-4% per year. Finally, it is likely that rapid increases in commodity prices during 2004-2008 skewed investor return expectations upward much like it did in the early 1970's.*

Key words: backwardation, commodity, contango, futures prices, investments, storable

JEL categories: D84, G12, G13, G14, Q13, Q41

*Scott H. Irwin is the Lawrence J. Norton Chair of Agricultural Marketing, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. Dwight R. Sanders is a Professor, Department of Agribusiness Economics, Southern Illinois University-Carbondale. Aaron Smith is a Professor in the Department of Agricultural and Resource Economics, University of California-Davis. Scott Main is a former graduate student in the Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. Correspondence can be directed to Scott Irwin. Postal address: 344 Mumford Hall, 1301 W. Gregory Dr. University of Illinois at Urbana-Champaign, Urbana, IL 61801. Phone: (217)-333-6087, Email: sirwin@illinois.edu.

"...the question is whether commodity ETFs provide reasonably good correlation to the commodities (spot?) market with minimal "net" erosion from contango (after taking into account the positive effects of occasional backwardation). And, if not, is there anything better for retail investors to gain exposure to this investment class without betting the farm through a dedicated commodities trading house. If there's nothing better, then are the deficiencies in the commodity ETFs so chronic that one should avoid them entirely, and simply stay out of the commodities space? I can't find anyone who seems to know the answers, so thank you for any thoughts you might offer."

—email message sent to the authors by a bewildered commodity investor

1. Introduction

Investment in commodity futures is now regularly included in lists of alternative investments that should be considered in discussions about the portfolio mix for investors. These investments include commodity index funds, exchange traded funds (ETFs), and exchange traded notes (ETNs), and, all of which either track broad commodity indices or individual commodities. Investors find commodity ETFs especially appealing because of the ease of access. Individuals with standard stock brokerage accounts can gain exposure to commodity futures markets (e.g., WTI crude oil futures) without opening a separate futures brokerage account. Moreover, the structure of ETFs removes investor concerns about executing trades in the underlying futures market, margin calls, and trading through delivery periods.

While each type of commodity futures investment is managed in a somewhat different manner, all such investments provide exposure to commodity prices through the purchase of futures contracts in a specific market or a basket of markets. As a particular futures contract nears expiration (e.g., June contract), it is sold and a more distant expiration month is simultaneously bought to replace it (e.g., July contract). This mechanical "rolling" of futures contracts allows the investment manager to maintain a constant exposure to a particular commodity futures market or group of markets. The vast majority of investments provide passive, long-only exposure to commodity futures price movements.

As shown in Figure 1, commodity-linked investments boomed from the early 2000s through 2010-2012, growing from less than $100 billion to a peak of almost $450 billion.[1] The publication of several influential academic studies helped fuel the boom by claiming that commodity

[1] There is no definitive measure of total investment in commodity futures markets. The series from Barclays in Figure 1 is the longest and most comprehensive, but it may contain some non-futures commodity investments, particularly in metals. The best measure is the *Index Investment Data* (IID) series from the U.S. Commodity

futures investment generated "equity-like" returns and provided portfolio diversification benefits (e.g., Gorton and Rouwenhorst 2006; Gorton, Hayashi, and Rouwenhorst 2007). A flurry of research studies in the private sector around the same time corroborated and reinforced the findings in the academic studies. These newer studies built upon earlier work that reported evidence of attractive returns to commodity futures investment (e.g., Greer 1978, 2000; Bodie and Rosansky 1980; Fama and French 1987; Ankrim and Hensel 1993).

As quickly as investment in commodity futures boomed, it went bust even faster. Figure 1 reveals that after peaking during 2010-2012 investment in commodity futures plummeted by more than half during 2013-2015, to a low of $160 billion. While investment in commodities recovered, it remained well below peak levels at the end of the third quarter of 2018. Not surprisingly, there were numerous articles in the financial press chronicling the demise of commodity futures investments during the bust years. For example, a *Wall Street Journal* article (Dugan 2013) reported, "Pension funds and other institutions are retreating from popular investments linked to commodities after finding they did little to protect their portfolios against inflation risk and the unpredictable returns of stocks." An article in the *Washington Post* (Samuelson 2015) entitled, "Behind the Commodities Bust," began this way: "First was the dot-com bubble, then the housing bubble. Now comes the commodities bubble."

Investors must have been sorely disappointed with the performance of commodity futures investments or they would not have been so quick to divest. The experience of the California Public Employees' Retirement System (Calpers), the nation's largest pension fund, is instructive. In October 2012, Calpers pulled out of 55% of its holdings in commodity indexes after losing about 8% annually over the previous five years (Dugan, 2013). Calpers was a pioneer in pushing commodity index investing in the pension industry. A similar but more recent example is the Teucrium Corn Fund (CORN) ETF. As shown in Figure 2, an investor in the Teucrium ETF who bought in January 2015 at a price of $24.98 per share watched the price sink to $18.19 by April 2018 for a loss of 27%. Meanwhile, the nearby futures price for corn rose from $3.70 to $4.01 per bushel. This was not an isolated example. Table 1 shows data on returns for 23 futures-based ETFs and ETNs that have 5-year track records. These funds had assets of $3.5 billion as of May 7, 2018. Only three of the 23 produced a positive return over the previous five years. Worse, eight of the 23 funds have lost over half of their value over that time period, with an average total loss of 37% over the previous 5 years and 16% over the previous 3 years. The puzzling aspect of this performance is that it has occurred over a period

Futures Trading Commission (CFTC), but this series does not begin until December 2007 and the CFTC ceased collecting the data as of October 2015.

of generally flat or even upward price movements for many commodities, which could lead even an experienced investor to ask, “What happened to my money?”

One explanation for the disappointing commodity returns is the process of “financialization,” which is simply a label for the wave of large-scale institutional investment in commodity futures markets that started in the mid-2000s. Theoretical models (e.g., Acharya, Lochstoer, and Ramadorai 2013; Hamilton and Wu 2015; Basak and Pavlova 2016) show how buying pressure associated with financialization can exert downward pressure on risk premiums, or equivalently, upward pressure on commodity futures prices before expiration. Main et al. (2018) test this prediction and find that the average unconditional return for 19 individual commodity futures markets is approximately the same before and after financialization. Therefore, despite the logical appeal of theoretical models, average returns in commodity futures market appear to have been largely unaffected by the process of financialization.

Disappointing returns to long-only commodity futures investments have also been blamed on the “carry” structure of futures prices. The two basic types of carry are contango and backwardation, where contango (backwardation) occurs when futures contracts further from expiration on a given date have a higher (lower) price than those contracts closer to expiration. The process of rolling long positions from lower-priced nearby to higher-priced deferred contracts in a contango term structure is said to create a negative “roll yield.” (e.g., Moskowitz, Ooi, and Pedersen 2012). It is widely accepted in the investment industry that roll yield is a key part of the return generating process for commodity investments. For example, a *Wall Street Journal* article (Shumsky 2014) explains, “Most commodities ETFs get their exposure by buying futures contracts, and over time they typically shift, or roll, their positions from nearby to later-dated contracts. Sometimes, they have to pay more for the new contracts, which eats into returns.” Several researchers have dissented from this conventional wisdom about roll yield and commodity futures returns (e.g., Sanders and Irwin 2012; Bhardwaj, Gorton, and Rouwenhorst 2015; Bessembinder et al. 2016; Bessembinder 2018).

This discussion suggests there is a major puzzle regarding returns to commodity futures investments. On one hand, both academic and private-sector research offers the prospect of commodity futures investments earning “equity-like” returns. On the other hand, actual investor experience with commodity futures investments, at least in the last decade, has clearly disappointed. The purpose of this article is to analyze the reasons for the puzzling returns of commodity futures investments in the last decade. In particular, are lackluster returns the norm or anomalous? Our main focus is on long-only returns at the individual market level in order to: i) reflect the passive, long-only exposure desired by the vast majority of commodity futures investors; and ii) investigate the fundamentals of commodity futures returns without the complications introduced by portfolio strategies that are inherently active in nature rather than passive.

In the first part of the article we use the theory of storable commodity markets to show that the disappointing returns of the last decade were not driven by contango or negative roll yields, as is widely purported. Next, we show how contango markets may have actually inflated investor return expectations due to misperceptions about spot versus futures investing. Then, we use six decades of daily return data for 19 storable commodity futures markets to document that the expected return in individual commodity futures markets is near zero before expenses. This implies that net losses to commodity futures investments will be equal to order execution and operating costs estimated at 3-4% per year. This mirrors the actual experience of investors in commodity futures during the last decade. It is likely that rapid increases in commodity prices during 2004-2008 skewed investor return expectations upward much like it did in the early 1970's, only to be disappointed by subsequent returns nearer the longer-run average of zero before expenses. Finally, there is limited evidence that the slope of the futures term structure (roll returns) provides a reliable signal about expected returns for individual commodity futures markets.

2. Basic Theory of Commodity Futures Returns

We begin our analysis commodity futures investment with a graphical presentation of the basic theory of returns in storable commodity futures markets. Figure 3 shows hypothetical pricing curves for a commodity futures market with a spot (cash) price P_t equal to $100 on January 1. The figure shows contango (Panel A) and backwardation (Panel B) term structures for commodity futures contracts that expire on December 31 of the same year. The cost-of-carry, C, in the contango chart is $10 per annum. The components of the carry include forgone interest, physical storage costs, and convenience yield. The latter is the implicit benefit to owners of physical commodity inventories for immediate use and it enters C as a negative value. The three possible spot prices, P_T, on December 31 are $105, $110, and $115 depending on the risk premium (points A, B, and D on the chart in Panel A).

The middle line in Panel A of Figure 3 represents the case with no risk premium. This means that a trader could buy the commodity on January 1, pay the storage fee, and be willing to sell on December 31 for $100 + $10 = $110. Thus, the price on January 1 of a futures contract that expires on December 31 ($F_{t,T}$) is also $110. In equivalent terms, the current futures price equals the expected value of the December 31 spot price ($F_{t,T} = E(P_T) = \$110$). In the course of a year the spot price will rise to the futures price of $110 to compensate owners of the physical commodity for storage costs. With no risk premium, the net return to a holder of the spot commodity over January 1 to December 31 is zero ($110 - $100 - $10 = 0) after accounting for the cost C paid to physically store the spot commodity. Likewise, the return to a futures investor is zero because the futures contract purchased on January 1 at \$110 receives \$110 at the expiration of the contract on December 31.[2] No-arbitrage conditions force convergence of spot and futures prices at contract expiration.

If a positive risk premium π is introduced then the returns to both spot and futures are altered. As demonstrated by the upper line in Panel A of Figure 3, spot prices appreciate above the level dictated by cost-of-carry and holders of the spot commodity earn a positive return after paying the cost associated with storing the physical commodity. In this specific example, a spot investor purchasing the commodity on January 1 at \$100 earns a return of \$115 - \$100 - \$10 = \$5 to compensate for the risk of owning the commodity. The futures investor realizes the same return by purchasing the futures contract at \$110 on January 1 and then selling the contract at \$115 on December 31. The futures investor is able to purchase the contract at \$110 on January 1 because that is the risk-free price of the commodity on December 31. The risk premium is paid to the futures investor in the form of a downward-biased futures price ($F_{t,T} = \$100 < E(P_T) = \$115$). Since no-arbitrage conditions force convergence of spot and futures prices at contract expiration, the same risk premium is earned by spot and futures investors. In sum, the spot investor earns a profit over and above the cost-of-carry and the futures investor realizes the same return by purchasing the futures contract at a price less than the expected spot price.

The upper line in Panel A of Figure 3 represents the classic Keynesian theory of normal backwardation in a storable commodity futures market. Since there is a long for every short in the commodity futures market, someone must earn a negative return if the commodity futures investor earns a positive return. The Keynesian theory of normal backwardation assumes the risk premium is compensation paid by hedgers (shorts) to speculators/investors (longs) for bearing price risk.

The lower line in the Panel A of Figure 3 represents the case of a negative risk premium. Here, spot price appreciation is below the level dictated by cost-of-carry and the holder of the spot commodity will earn a negative return (\$105 - \$100 - \$10 = -\$5) after paying the cost associated with storing the physical commodity. Because the current futures price is biased upward compared to the expected spot price ($F_{t,T} = \$110 > E(P_T) = \$105$), a futures investor earns the same negative return as the spot holder. A negative risk premium reverses the logic of the

[2] The examples considered here do not have a stochastic (error) component. This will be considered in the formal mathematical presentation of the cost-of-carry model in the next section.

Keynesian theory of normal backwardation. That is, hedgers (shorts) earn the risk premium paid by speculators/investors.

Panel B in Figure 3 represents a futures market with a backwardation term structure.[3] The three possible spot prices P_T on December 31 are \$95, \$90, and \$85 depending on the risk premium (points A, B, and D on the chart in Panel B). The cost-of-carry C is assumed to be -\$10 per year, representing a market situation where the convenience yield (negative value) is larger than the other components of storage costs. This situation is not typical in most storable commodity futures markets and tends to occur when the supply/demand balance is "tight" and inventories are relatively low. Under such market conditions, inventories will only be held by those firms for which running out of inventory would be very costly. They pay a relatively high price to buy the commodity and then reap the \$10 convenience yield through maintaining operations when others may not. Once again, the middle line in Panel B represents the case of no risk premium, so a spot investor purchases at \$100 on January 1 and sells at \$90 on December 31. Somewhat counterintuitively, the holder of the spot commodity earns a net return of zero (\$90 - \$100 + \$10), reflecting the offset of warehouse, insurance, and interest costs by convenience yield.

The return to a futures investor with no risk premium is also zero because the purchase price on January 1 ($F_{t,T} = E(P_T) = \$90$) is the same as the selling price at expiration on December 31 (\$90). The introduction of a positive risk premium, represented again by the upper line in Panel B, benefits the spot holder who purchases at \$100 and sells at \$95. This offsets part of the depreciation in spot prices as determined by the (negative) cost-of-carry, so the net return is \$100 - \$95 + \$10 = \$5. A futures investor purchasing on January 1 at \$90 can expect the futures price to rise to the expected spot price at \$95, and thereby, earns the same return as the spot holder. Results are simply reversed for the case of a negative risk premium with a backwardated term structure (the lower line in Panel B).

The graphical analysis based on Figure 3 suggests two important results. First, the carry structure of the futures market, whether contango or backwardation, has no direct impact on the long return to holding a commodity futures contract in terms of cash flows to investors. Second, the return earned by a long futures investor is solely determined by the risk premium embedded in the futures price prior to expiration. If the risk premium is positive (negative) then the long investor will earn a positive (negative) average profit.

[3] "Backwardation term structure" and "Keynesian theory of normal backwardation" are distinct concepts that are often confused. The former refers to the difference between spot and futures prices on a particular trading date. The latter refers to the expected change in the price of a particular futures contract between trading dates.

3. The Cost-of-Carry Model and Commodity Futures Returns

The "cost-of-carry" model (e.g., Tomek 1997; Pindyck 2001) is a well-developed theoretical model of pricing storable commodities. We use the cost-of-carry model to formally analyze the drivers of returns to long-only commodity investments. The model is the same as that found in Bessembinder et al. (2016) except we assume a constant storage cost in order to simplify the exposition. To be consistent with the graphical analysis in the previous section, let P_t represent the spot price at date t, $F_{t,T}$ represent the futures price at time t for delivery on date T, and C is the per period cost-of-carry which includes foregone interest, physical storage costs, and the convenience yield associated with having stocks on hand. The cost-of-carry C normally is dominated by interest and physical storage costs, in which case the futures market is in a normal carry or contango ($C > 0$). Other times, the convenience of having stocks on-hand dominates, in which case the futures market is inverted ($C < 0$). In all cases, the cost-of-carrying inventory is revealed by the term structure of the futures market.[4]

The no-arbitrage cost-of-carry relationship depicted graphically in Figure 3 can be expressed in mathematical terms as follows,

(1) $$F_{t,T} = P_t e^{C(T-t)}.$$

The return on the spot commodity net of storage costs is,

(2) $$U_{t+1} = ln\left[\frac{P_{t+1}}{P_t e^C}\right].$$

Bessembinder et al. (2016) call U_{t+1} the *ex post* premium. It has two components: (i) the *ex ante* risk premium (π), which is the return that holders of the commodity expect to earn as compensation for risk, and (ii) the *ex post* price shock (ε_{t+1}), which includes unforeseen supply and demand shocks. To make this clear, we write $U_{t+1} = \pi + \varepsilon_{t+1}$.

Market forces imply that *ex post* price shocks ε_{t+1} should average zero, and therefore the *ex post* premium is determined, on average, by the risk premium.[5] For example, if traders expect

[4] There is an important exception to this result. Garcia, Irwin, and Smith (2015) show that the futures term structure provides a downward-biased measure of the cost-of-carrying inventory when the market price of physical storage exceeds the maximum storage rate allowed by the delivery terms of the futures market. This situation actually occurred frequently over 2005-2010 for grain futures markets, and Garcia, Irwin, and Smith show how this explains the much-discussed episodes of non-convergence that plagued the markets during this time period.

[5] Formally, this statement applies to the mean of the exponential rather than the level of U_{t+1}. In a rational expectations equilibrium, the expected price next period equals the current price plus the carrying cost and the

demand for the commodity to increase in the future, then they will hold some of the commodity off the market to store in anticipation of higher future prices. This action will cause current prices to rise and eliminates any excess returns from storage.

Equations (1) and (2) can be used to express the continuously compounded returns to holding spot and futures positions,

$$ln\left[\frac{P_{t+1}}{P_t}\right] = \pi + \varepsilon_{t+1} + C \tag{3}$$

$$ln\left[\frac{F_{t+1,T}}{F_{t,T}}\right] = \pi + \varepsilon_{t+1} \tag{4}$$

where, as above, $U_{t+1} = \pi + \varepsilon_{t+1}$. The gross return to the holder of the cash or spot commodity in (3) is the sum of the *ex post* premium U_{t+1} and the market-implied cost-of-carrying the inventory C. From (4), the return to a long futures position is the *ex post* premium U_{t+1}. Note, the difference in returns to the spot (3) and the futures (4) is the cost of carry C. The spot market prices will appreciate by the cost of carry; whereas, that cost C is incorporated into the futures price $F_{t+1,T}$.

Equations (3) and (4) yield three important predictions regarding returns in storable commodity markets. First, if there is a risk premium ($\pi \neq 0$), it appears in both the spot and futures returns. Different risk premiums in spot and futures prices would be a violation of no-arbitrage conditions. Second, if there is no risk premium ($\pi = 0$), then cash prices will change by exactly the carrying costs (C) and futures returns are determined entirely by *ex post* supply and demand price shocks. Carrying costs are incorporated in the period t futures price $F_{t,T}$ so they are not a component of the return from period t to $t + 1$. Third, long futures returns in (4) are not directly determined by the carry term structure of the futures markets or the level of C. In other words, the long futures return is not a function of whether the futures market is in contango ($C > 0$) or backwardation ($C < 0$).

4. Decomposing Commodity Futures Returns

A popular approach to characterizing long-only commodity futures returns is to decompose returns into the sum of spot and "roll" returns (e.g., Moskowitz, Ooi, and Pedersen 2012). Equations (3) and (4) can be easily combined to demonstrate this common decomposition.

---

ex ante risk premium, i.e., $E(P_{t+1}) = P_t e^{\pi+C}$, which implies $E(e^{U_{t+1}}) = e^{\pi}$ and, from Jensen's inequality, $E(U_{t+1}) < \pi$.

That is, solving (3) for $\pi = ln[P_{t+1}/P_t] - \varepsilon_{t+1} - C$ and substituting into (4) we obtain the relationship that futures returns equal spot returns minus the carry,

$$(5) \qquad ln\left[\frac{F_{t+1,T}}{F_{t,T}}\right] = ln\left[\frac{P_{t+1}}{P_t}\right] - C$$

In (5), the carry C adds to the futures returns when the market is inverted (negative C) and detracts from returns when the market is in contango (positive C). Note that C is defined exactly the same as the roll return in the standard decomposition (e.g., Moskowitz, Ooi, and Pedersen 2012) except the sign is the opposite in our case. Therefore, by simply reversing the sign on C in (5) we obtain the identity commonly used to decompose the return to a long futures position,

$$(6) \qquad \text{Futures Return} = \text{Spot Return} + \text{Roll Return}.$$

Note that equation (6) represents the same accounting identity as equation (5) in the sense that the futures returns must equal the spot return plus the carry (or roll return) in the no-arbitrage cost-of-carry model.

The standard decomposition given by equation (6) is frequently used to explain high and low returns to long-only commodity futures investments based on the carry structure of prices in futures markets. Examples in the financial press abound. For example, a recent *Bloomberg.com* article (Nussbaum and Javier 2016) stated, "Part of the problem is how fund managers try to mimic price changes. Rather than buy raw materials that have to be stored, they use futures contracts. However, when those expire—sometimes every month—returns suffer if contracts are replaced at higher cost. That occurs when markets are in contango, meaning that commodities for immediate delivery are cheaper than in the future, as they are now for everything from corn to crude."

The problem with the conventional wisdom regarding roll returns as a driver of future returns is that no such link is implied by the standard cost-of-carry model. The crucial observation is that no causal direction among the components is implied by whichever representation of futures returns, equation (5) or (6), is employed. In the rational theory of storage (e.g., Williams and Wright 1991), the futures price, spot price, and price of storage (carry) are determined simultaneously. That is, a market in contango (negative roll return or positive C) or backwardation (positive roll return or negative C) is determined by the supply and demand for storage in that particular commodity market in conjunction with the simultaneous determination of spot and futures prices. This means that the variable on the left-hand side of the standard decomposition can be stated with equal theoretical validity as, (7a) Spot Return = Futures Return - Roll Return,

(7b) Roll Return = Futures Return - Spot Return.

There simply is no theoretical basis for asserting a particular direction of causality among the components of (5) or (6) in the standard cost-of-carry model. Moreover, if a contango market caused a negative futures return in the sense used by most investment practitioners, it would imply a gross violation of market efficiency in futures markets, in that an extremely simple and highly profitable trading rule could be easily implemented for a given market (Sanders and Irwin 2012). In sum, these results conclusively demonstrate that the popular discussion around roll yield as a source of commodity futures returns is misguided. Bessembinder (2018, p. 51) provides a definitive statement, "In fact, however, the roll yield as an actual cash gain or loss to a futures investor is a myth."

5. Roll Returns and the Performance Gap

The discussion in the previous section suggests there is considerable confusion about the nature of returns to long-only commodity futures investments. In particular, the belief that roll returns drive futures returns can lead to the perception of a "performance gap" in commodity futures investments. We demonstrate how the perception of a performance gap can develop by comparing (implied) spot prices and the value of the three largest single-commodity ETFs—U.S. Oil Fund (USO), Powershares DB Gold Fund (DGL), and U.S. Natural Gas Fund (UNG). Our procedures are similar to those found in Bessembinder et al. (2016). The USO series begins on April 10, 2006, the DGL series begins on January 5, 2007, and the UNG series begins on April 18, 2007. All three ETF series end on June 30, 2017. We compute implied spot price prices using a two-step procedure. First, the daily cost-of-carry is estimated as $C_t = [1/(Z-T)] \cdot ln[F_{t,Z}/F_{t,T}]$, where $F_{t,T}$ is the price of the nearby futures contract with delivery date T, $F_{t,Z}$ is the price of the next-to-expire futures contract with delivery date Z, and $Z > T$. Second, the daily implied spot price is estimated by discounting the nearby futures price as $P_t = F_{t,T}/e^{C_t(T-t)}$. Implied spot prices are estimated in order to provide a consistent measurement of spot prices across all three commodities.[6]

Figures 4, 5, and 6 plot the path of the three ETFs over their lifetime compared to the corresponding (implied) spot price. We also plot the path of a simulated ETF implied by holding a long position in nearby futures and rolling to the next contract 21 days before first delivery

[6] Because crude oil, gold, and natural gas futures have contract expirations every month, the implied spot prices are very close to the nearby futures throughout the sample, i.e., the results that follow would be very similar if we used $F_{t,T}$ in place of the implied spot price.

day.[7] We also subtract a daily fee equal to the reported expense ratios for each of the three ETFs, i.e., annualized 0.76% for USO, 0.77% for DGL, and 1.7% for UNG.[8] The USO ETF was launched April 10, 2006 at a price of $68.02 per share, roughly equivalent to the crude oil implied spot price on the same date of $68.31/bbl. By June 2014, the crude oil implied spot price had risen to over $100/bbl. but the USO price had fallen to $38. In the latter part of 2014, the futures market went into backwardation, both funds fell sharply, and the gap closed somewhat. Figure 4 shows that the simulated ETF closely matches fluctuations in the USO fund. In particular, the simulated and actual funds did not rebound with the spot price in 2009 because the price of storage was very high in that period.

Perhaps the most vivid example of the performance gap is found in gold, where storage costs are very stable (essentially just interest costs) at 1.3% per year. Although both the spot and futures show a positive average return over this sample, the difference between the implied spot price and the futures return is a very consistent 1.8% per year. The difference manifests itself as a performance gap that accumulates in a stable fashion, as shown in Figure 5 for a simulated gold ETF that almost exactly matches the DGL.

Natural gas prices decreased significantly during the sample period, as shown in Figure 6. High storage costs further compounded the misery for investors in UNG. Figure 6 shows that, as the Henry Hub spot price declined from $7.66/MMBtu in April 2007 to $2.89 in December 2014, the UNG fund and the simulated ETF dropped to $0.28 and $0.30, respectively. In contrast to gold, the simulated ETF returns do not match the UNG returns exactly. We also saw in Figure 4 that simulated crude oil returns did not match the USO return exactly. These differences likely reflect differences in the trading strategies used by the USO fund and those employed in our estimation procedures. Trading strategies can matter because the implied price of storage can differ between nearby and more distant futures contracts, especially for natural gas, which has a seasonal storage pattern.

There is a straightforward explanation of the ETF underperformance when compared to spot commodity prices, i.e., the performance gap in Figures 4, 5, and 6. Equation (5) in the section on the cost-of-carry model shows that the expected difference between spot and futures returns

[7] The simulated ETF prices could also be computed by subtracting carrying costs from the return on the implied spot price.

[8] Source http://finance.yahoo.com. These expense ratios are in line with ETFs on commodity indexes. We collected the annualized expense ratio of four ETFs over 2008-2012. The average expense ratios for these funds were: iShares S&P GSCI Commodity-Indexed Trust (GSG, expense ratio = 0.75), iPath DJ-UBS Commodity Index Trust, Exchange Traded Note (DJP, expense ratio = 0.75), GreenHaven Continuous Commodity Index (GCC, expense ratio = 0.85) and GS Connect S&P GSCI Enhanced Commodity Trust, Exchange Traded (GSC, sxpense ratio = 1.25).

equals the carrying cost. Thus, if carrying costs are positive ($C > 0$), spot prices will appreciate faster than futures prices to cover the cost of storage. As a result, an apparent performance "gap" will develop between the price of the spot commodity and the value of a commodity futures investment. If carrying costs were to switch sign ($C < 0$) then this apparent performance gap will narrow or possibly even reverse in sign. However, the narrowing only occurs because the spot price declines to reflect the convenience yield. The difference between spot price returns and futures market returns is simply a function of the cost of storage, or roll returns, which drives a wedge between the two price series. Indeed, the price series would coincide in the absence of storage costs ($C = 0$).

In sum, comparing long-only commodity futures investment performance to a spot price benchmark for a market with a contango term structure will always be frustrating to investors as the spot price performance represents an unattainable strategy of buying-and-holding the spot commodity without paying storage. Conversely, comparing long-only commodity futures investment performance to a spot price benchmark in a backwardated term structure will always show the futures investment outperforming the spot. A trading strategy in futures does not entail holding the physical commodity and its return includes no compensation for storage costs as it is already built into the term structure of futures prices. Since contango is the norm in storable commodity markets, these examples highlight that a "performance gap" between spot prices and the value of commodity investments is also the norm. To the degree that commodity investors base return expectations on spot prices, this will lead to disappointment about "poor" returns when the actual problem is the use of a wrong performance benchmark.

6. Historical Commodity Futures Returns

It is certainly reasonable for investors to ask why they made or lost money in a long-only commodity futures investment, but this is a function of risk premiums rather than roll returns (at least directly). As Bhardwaj, Gorton, and Rouwenhorst (2015, p.2) state, "The source of value to an investor in commodity futures is the risk premium received for bearing future spot price risk." To that end, we estimate historic average returns for 19 storable commodity futures markets over July 1959 through June 2017. The markets, sample starting dates, and number of daily observations are listed in Table 2. The commodities include New York Mercantile Exchange (NYMEX) energy and metals markets, Intercontinental Exchange (ICE) softs markets, Chicago Board of Trade (CBOT) grain markets, and the Kansas City Board of Trade (KCBT) wheat market. We do not include any livestock markets in our sample because these are non-storable commodities and the cost-of-carry model does not strictly apply. Over 10,000 daily observations are available in 14 of the 19 markets and the remaining five markets have at least 6,700 daily observations. Therefore, the sample is broad in terms of the number and types of commodities and deep in terms of the number of observations available for each market. This should provide a comprehensive perspective on the long-only returns available to investors in storable commodity futures markets.

Daily (annualized) futures returns are computed as the log difference in nearby futures prices. In all computations, returns are for the same nearby futures contracts and contracts are rolled to the next contract 21 days before first delivery day. The daily averages are annualized by multiplying by 250. Finally, the futures return that we compute is technically considered to be an "excess" return because it does not add the interest (Treasury Bill) return associated with a fully collateralized long-only commodity investment.[9]

The full sample estimates in Table 3 show that average futures returns are negative in 12 of the 19 markets. Three markets show evidence of statistically significant negative returns—natural gas, corn, and rough rice. Only the crude oil and soybean complexes show consistently positive average returns. The only statistically significant and positive return is in copper. The average futures return across the 19 markets is -0.3% per year and the return is statistically indistinguishable from zero. Note that average returns are reported in annualized form, which implies that daily average returns are very small. Consider the largest annualized positive return of 7.1% for copper. This translates into slightly less than three basis points per day. Most of the average returns for individual markets are less than two basis points per day. When the small size of the average daily futures returns is combined with the high variability of futures returns, it is not surprising there is so little evidence of statistically significant futures returns. With daily standard deviations for the 19 markets ranging from 1.2 to 2.9%, it is natural to be concerned about low statistical power in such "noisy" return series (e.g., Summers 1986). This issue is less of a concern when one considers the fact that point estimates for average returns are negative for almost two-thirds of the markets. In addition, the point estimate of average returns across the 19 markets is negative. Overall, the full sample results in Table 3 provide strong evidence that realized excess returns for individual futures markets are near zero.

An important fact to keep in mind is that the "paper" returns reported in Table 3 do not account for the costs of long-only commodity investment. Bessembinder et al. (2016) estimate that the order execution (liquidity) cost of roll trades for the USO ETF to be about 3% per year. Earlier in Table 1, we reported that the direct operational expenses for long-only ETFs

[9] The addition of interest returns to excess futures returns to create a fully collateralized futures return will create a return series that may have quite different statistical properties than excess returns. The addition of interest returns will not only change the mean of the return series but also the volatility. For this reason, means, standard deviations, or statistical test results for excess futures returns should not be compared to those derived from fully collateralized returns.

average 0.75% per year. We can combine this information to estimate total expenses for long-only commodity futures in the range of 3-4% per year. Total expenses of this size would substantially increase net losses in the 12 markets with negative gross returns and turn positive gross returns into net negative returns in 4 of the 7 markets with positive gross returns. Average returns are large enough to remain positive with expenses of 3-4% per year in only three markets—RBOB gasoline, copper, and soybean oil.

Further perspective is provided by examining the pattern of commodity returns by decade. Returns for each decade since the 1960's are presented in Table 3. We present estimates only for complete decades. For example, a market that started trading in 1978 does not show any data in the 1970's but shows a complete return history for the 1980's. The data are arranged in this manner to keep the markets within each decade consistent and complete. The decade averages in Table 3 reveal that positive futures returns have really only marked two decades, the 1970's (13.6% with 11 of 12 markets positive) and the 2000's (1.3% with 12 of 19 markets positive). The positive futures returns in the 1970's were mostly driven by the rapid price adjustments in the early part of the decade. Sanders and Irwin (2012, p. 524) note that "…the three tumultuous years from 1972-1974, when the commodity markets underwent dramatic structural shifts, accounts for 96% of the decade's returns and 68% of all returns during 1961-2010." It is not surprising that returns of this magnitude were accompanied by an increased focus on commodities by investors. For example, Laby and Thomas (1975, p. 287) observe that there was a "switch of speculative funds away from traditional asset placements and towards commodity futures contracts" following the spike in prices during 1972-1974.

In a similar fashion, the upheaval in commodity prices from 2004-2010—which was accompanied by popular themes such as "peak oil" and "food for fuel"—piqued the interest of investors in commodity-related instruments. Investment firms met this demand with index funds, ETFs, and other vehicles that allowed investors convenient access to those markets. However, much like what happened in the 1970's, it appears that the performance of those investments was overhyped and investors anchored their expectations too much on the most recent high returns. Not surprisingly, subsequent years were characterized by poor performance, with the 1980's averaging -7.1% and the 2010's averaging -6.6%. Undoubtedly, commodity investors in both periods were likely disappointed—and perhaps puzzled—by the lackluster returns they earned. In this context, it is interesting to note there is well-documented precedent of investment returns in the commodity space being overhyped (Irwin 1994; Bhardwaj, Gorton, and Rouwenhorst 2014).

The historical averages by decade provide a very useful perspective on the pattern of expected returns in commodity futures markets since the 1960s. However, these comparisons are limited by the fact that most of the markets have different starting years, which means that all but the most recent decades contain different sets of markets. In order to standardize the results across markets we conduct a bootstrap simulation of 10-year average returns where: i) 2,500 daily returns are randomly sampled from the full sample of observed daily returns for each market; ii) the average return over the 2,500 randomly drawn returns is computed; iii) the simulation is repeated 1,000 times and the proportion of 10-year average returns greater than zero are counted. The results of the bootstrap simulation reported in Figure 7 are striking. Only 7 of the 19 markets have more than a 50% probability of positive average returns over a 10-year horizon and only 3 markets have more than a 70% chance of positive average returns. Equally interesting is the fact that 8 of the 19 markets have a 40% or smaller probability of positive average returns over a 10-year horizon. These simulation results highlight the limited potential for earning positive returns in long-only commodity futures over a 10-year investment horizon.

Finally, the results reported in this section are consistent with the findings in a number of previous studies that the average unconditional return to individual commodity futures markets is approximately equal to zero. For example, Kolb (1992) examines 29 commodity futures markets and finds evidence that 18 had positive returns and 11 had negative returns. He concludes that most markets exhibit no risk premium. Gorton and Rouwenhorst (2006) analyze 36 individual markets and find that 18 had positive returns and 18 had negative returns, with none of the individual markets having statistically significant positive returns. Erb and Harvey (2006) find that the average return to 12 commodity futures markets from 1982-2004 is -1.71% with no individual market producing statistically positive returns. Sanders and Irwin (2012) examine 20 commodity futures markets and report that, outside of the 1970s, the number of markets with negative returns roughly equals the number with positive returns and the average return is relatively close to zero.

In terms of the cost-of-carry model, these findings imply that the *ex post* premium ($U_{t+1} = \pi + \varepsilon_{t+1}$) for commodity futures markets consists solely of unforeseen supply and demand price shocks (ε_{t+1}) because the *ex ante* risk premium (π) for long-only positions is zero. Borrowing from Kolb (1992), the implication is that "normal backwardation is not normal." This conclusion is not as surprising when one considers the fact that commodity futures contracts are not included in the market portfolio because there is a short position for every long position (Black 1976). In other words, futures markets are simply side bets on commodity prices, not capital assets that earn a structural risk premium. An alternative explanation is that the supply of speculative services in commodity futures markets is perfectly elastic (e.g., Telser 1958; Dusak 1973; Hartzmark 1987).

7. Risk Premiums and Roll Returns

While the average unconditional return for individual commodity futures markets is approximately equal to zero, this does not necessarily imply that the conditional return is also equal to zero. There is a long strand of the literature that reports evidence of time-varying risk premiums (e.g., Cootner 1960; Carter, Rausser, and Schmitz 1983; Bjornson and Carter, 1997; Moskowitz, Ooi, and Pedersen 2012; Szymankowska et al. 2014), where expected returns in commodity futures markets change in relation to the position of hedgers or financial market conditions. It is crucial to recognize that any commodity futures investment based on conditional returns requires some type of underlying dynamic trading strategy. That is, earning conditional returns may require investors to follow a trading strategy that goes both long and short in commodity futures markets. A less dynamic strategy entails investors reducing their long exposure to a commodity based on some type of signal, but not taking outright short positions. Regardless, following any type of dynamic strategy is quite different from the passive, long-only strategy we have analyzed up to this point.

A strategy that has attracted considerable attention in recent years is one of buying commodities with low roll returns ("low basis/backwardation") and selling commodities with high roll returns ("high basis/contango"). In this case, conditioning on the level of roll returns serves as a signal for the magnitude of risk premiums. Several studies show that this strategy produces positive and statistically significant returns (e.g., Gorton and Rouwenhorst 2006; Erb and Harvey 2006; Bhardwaj, Gorton, and Rouwenhorst 2015). From a conceptual standpoint, this strategy implies that the risk premium π in the cost-of-carry model (Figure 3) is negative for markets in contango and positive for markets in backwardation. The theoretical model in Gorton, Hayashi, and Rouwenhorst (2013) provides rigorous justification for this type of strategy. Both storage costs and the risk premium depend on inventories and price volatility in their model. Under certain conditions, a market in contango will tend to have less volatile prices, which implies that risk-averse investors accept a lower risk premium than when the market is in backwardation. It is important to emphasize that saying a market in contango is a *signal* that the realized return in a subsequent period will be negative is quite different from saying that rolling commodity futures positions in a contango market *creates* a contemporaneous negative realized futures return.

We investigate whether roll returns serve as a signal for risk premiums in the 19 commodity futures markets included in our study. If such relationships exist, then it would imply that, even though there is no mechanical connection between the futures term structure (roll returns) and commodity futures returns, traders can profit by using roll returns as a trading signal. We use a two-step procedure to decompose futures returns into spot and roll returns. The first step involves computing daily returns as follows:

(8a) Non-Roll Days: $$ln\left[\frac{F_{t+1,T}}{F_{t,T}}\right] = ln\left[\frac{F_{t+1,T}}{F_{t,T}}\right] + 0.$$

(8b) Roll Days: $$\underbrace{ln\left[\frac{F_{t+1,Z}}{F_{t,Z}}\right]}_{\substack{\text{Futures}\\ \text{Return}}} = \underbrace{ln\left[\frac{F_{t+1,Z}}{F_{t,T}}\right]}_{\substack{\text{Spot}\\ \text{Return}}} + \underbrace{ln\left[\frac{F_{t,T}}{F_{t,Z}}\right]}_{\substack{\text{Roll}\\ \text{Return}}},$$

Equation (8a) is used to compute futures and spot returns on non-roll days. Note that futures and spot returns are the same on non-roll days because we account for all the storage costs on roll days. Equation (8b) is used to compute returns on roll days, or days when the contract in the nearby position changes between day t and $t+1$. On those days, the trader rolls out of the old nearby contract that expires on date T and into the new nearby contract that expires on date Z, with $Z > T$.

We assume that a trader holding the spot commodity pays for storage on the roll day. As shown on the right-hand side of (8b), the roll return, $ln[F_{t,T}/F_{t,Z}]$, is the log difference between the nearby contract and next-to-expire contract on day t. This is a standard way to compute storage costs for a commodity on a given date based on the term structure of commodity futures prices. Note that this formulation results in negative roll returns when a market has a contango term structure and positive roll returns when a market has a backwardation term structure. Since the nearby contract changes on roll days, the spot return shown on the right-hand side of (8b) includes the roll return. Specifically, the spot return on roll days is the log difference between the price of the nearby contract at $t+1$, $F_{t+1,Z}$, and the nearby contract at t, $F_{t,T}$. This, of course, means the spot return on roll days is computed across different futures contracts. The futures return on roll days, $ln[F_{t+1,Z}/F_{t,Z}]$, is computed using the "new" nearby futures contract at $t+1$.

The second step in the procedure is to compute an average of the daily returns generated by equations (8a) and (8b) for futures, spot, and roll returns. In all three cases, a simple average is computed across all non-roll and roll days. The averaging process is straightforward except perhaps in the case of roll returns, which are zero except on roll days. By averaging across zero roll returns on non-roll days and non-zero roll returns on roll days, the roll return (storage cost) is allocated across all days in the sample. Equivalently, storage costs are paid on a single day—the roll day—and then averaged across all days.[10]

The decomposition of full sample futures returns by market is reported in Table 4. The futures returns are the same as the full sample returns reported in Table 3. As shown in Equation (6), the futures return for WTI crude oil (1.8%) equals the spot return (1.9%) plus the roll return (-0.1%). The spot return is positive for all markets but it is not statistically significant in any of the 19 markets. Across all markets, the spot return averages a statistically insignificant 2.9%. The roll return is consistently negative (16 of the 19 markets) and the average roll return is -3.3% and statistically significant, indicating that contango is the normal market structure. Compared to futures and spot returns, the relatively low variability of the carry (roll return) leads to higher statistical precision around those estimates.

A close examination of Table 4 shows that three markets have positive roll returns and positive futures returns and 12 markets have negative roll returns and negative futures returns. Hence, roll returns and futures returns have the same sign in 15 of the 19 markets. Figure 8 plots the average roll and futures returns, excluding natural gas since it distorts the scales. There is a strong positive correlation in average roll and futures returns across markets (correlation = 0.96), consistent with the evidence in previous studies that commodities with higher roll returns have lower futures returns and *vice versa* (e.g., Gorton and Rouwenhorst 2006; Erb and Harvey 2006; Gorton, Hayashi, and Rouwenhorst 2013; Bhardwaj, Gorton, and Rouwenhorst 2015). While there is undoubtedly a high cross-sectional correlation between average futures and roll returns, "connecting the dots" between the cross-sectional averages obscures the inherent variability in the time-series average for each market. Figure 8 include 95% error bars for average futures returns, which illustrates the very substantial time-series variability obscured by only focusing on cross-sectional averages.

The fundamental question is whether a time-series relationship exists between roll returns and futures returns that commodity futures investors can use to reliably increase returns over that of a passive long-only strategy. To test this formally, futures returns for each of the 19 markets in this study are regressed on an indicator (dummy) variable that equals zero if the market is in contango and equals one if the market is in backwardation. This is a difference-in-means test. Specifically, if the coefficient on the dummy variable is statistically different from zero,

[10] This method of computing futures, spot, and roll returns is generalizable to any holding period. Specifically, the method is equally applicable to monthly returns. In contracts which are rolled each month (like crude oil), equations (8a) and (8b) will correctly calculate each component. In contracts which are rolled at irregular intervals (like grains), (8a) is used for non-roll months and (8b) is used for roll months. Then, similar to daily data, averaging the returns essentially prorates the roll return across all months.

then the roll returns (backwardation or contango) provides information about subsequent returns in the time series. Table 5 shows that 14 of the 19 commodities exhibit lower returns in contango periods than in backwardation periods over the entire sample period. The coefficients indicate mean differences are statistically significant for seven markets, with the estimates providing evidence of lower returns in 6 of 7 markets during contango and larger returns in one market (rough rice). To further investigate the reliability of these signals, we split the sample at the end of 1988. In the first half of the sample, five of the 19 markets have statistically lower returns during contango. In the second half of the sample the relationship seems to fade. That is, only two of the 19 markets have statistically significant differences between returns in contango markets and returns in backwardated markets. One of those has lower returns in contango (crude oil) and one has higher (rough rice).

In spite of the unevenness of the statistical significance of the results in Table 5, some of the point estimates are quite large. For instance, across the entire sample the annualized return to crude oil during backwardation markets is 23.6% larger than during contango. However, even that large of a return difference only generates a t-statistic of 1.98 across the entire sample. The variability is further illustrated in heating oil where in the second half of the sample a trader only holding long positions during backwardation would have a return 22.3% greater than those held in contango markets, but it would essentially be indistinguishable from luck (t-statistic = 1.64).

The combination of large coefficient estimates and low statistical significance arises because time-series returns at the individual market level are highly volatile. This volatility means that, even if the differences in returns are stable at the values displayed in Table 5, a trader could hold long positions for years in a given commodity futures market and still not realize a benefit from only buying in backwardated market conditions. These results do not provide for high confidence in investment strategies based on changing contango and backwardation. Moreover, the value of this signal varies widely across markets and is much less prevalent in recent years when eight of the 19 markets actually show higher returns during contango. An investor would have to be fortunate enough to pick the right market over the right time period to realize any benefit from roll return signals. It is certainly debatable whether investing in commodity futures based on roll returns is a reliable strategy or one that most investors in commodities would be willing to follow.

In sum, picking up risk premiums based on the level of roll returns is highly uncertain and requires following a dynamic trading strategy that has an unclear foundation. In addition, the cost of implementing and managing dynamic strategies undoubtedly is higher than the 3-4% per year estimated earlier for passive long-only strategies. This only raises the performance bar further. We agree with Erb and Harvey (2006, p. 91): “There is, of course, reason to doubt how broad based the demand for commodity TAA [tactical asset allocation] might be. Many, if not most, investors interested in investing in commodities are interested only in a long-only exposure to commodity futures. A TAA approach is unacceptable to these investors because they want to know that they will always have a well-defined long exposure to the commodities market. Tactical strategies that allocate among commodities, or go long or short commodity futures, will naturally leave these investors wondering about what sort of portfolio exposure they happen to have at any point in time." The bottom-line is that the return to a strategy based on roll returns is not one that can be earned by passive investing, which to date has been strongly preferred by investors.

8. Summary and Conclusions

The returns to long-only commodity futures investments have generally disappointed since the explosion in their popularity during the mid-2000s. The puzzling aspect of the poor performance is that it often occurred when the overall trend in commodity prices was flat or even upward. The purpose of this article is to analyze the reasons for the lackluster returns of commodity futures investments in the last decade.

The "carry," or term structure, of futures prices is often blamed for disappointing returns to commodity futures investments. We use the cost-of-carry model for storable commodity prices to show that the carry in a commodity futures market has no direct effect on the returns to long futures positions. As Bessembinder (2018, p. 51) states, "…the roll yield as an actual cash gain or loss to a futures investor is a myth." We then demonstrate how this common misconception about carry and commodity futures returns can create an apparent "performance gap" in the minds of investors. Using data for the three largest single commodity ETFs, we show that the apparent performance gap between spot price levels and long-only commodity ETF investments is due to an "apples and oranges" comparison problem. In particular, comparing long-only commodity futures investment performance to a spot price benchmark for a market with a contango term structure will always be frustrating to investors as the spot price performance represents an unattainable strategy of buying-and-holding the spot commodity without paying storage.

We next investigate daily returns for 19 storable commodity futures markets over July 1959 through June 2017 in order to provide historic evidence about expected returns. Unconditional futures returns are very close to zero on average, with 12 of the 19 markets showing a negative return. Daily average returns are very small, at most two to three basis points per day. The average annual futures return is -0.3% per year across the 19 markets and the return is statistically indistinguishable from zero. Overall, the results strongly suggest that the unconditional return in commodity futures markets is near zero before expenses. We concur with Erb and

Harvey (2006, p. 94) that the average commodity futures contract does not have an "equity-like" return. Since order execution and operating costs for commodity futures investments are estimated to be 3-4% per year, the results imply that the expected net loss for investment in long-only commodity futures is around 3-4% per year. This mirrors the actual experience of investors in commodity futures during the last decade. Hence, the attraction of passive, long-only commodity investments is rather elusive.

So, why the rush to commodities starting in the mid-2000s? The pattern of historical commodity returns is instructive. Positive commodity futures returns have really only marked two decades—the 1970's and the 2000's. The positive futures returns in the 1970's were mostly driven by the rapid price adjustments in the early part of the decade and this led to an increased focus on commodities by investors. In a similar fashion, the upheaval in commodity prices from 2004-2008—which was accompanied by popular themes such as "peak oil" and "food for fuel"—piqued the interest of investors in commodity-related instruments. Investment firms met this demand with ETFs and other vehicles that allowed investors convenient access to those markets. However, much like what happened in the 1970's, the performance of those investments was overhyped and investors anchored their expectations too much on recent peak returns. There is certainly precedent for investment returns in the commodity space being overhyped (Irwin 1994; Bhardwaj, Gorton, and Rouwenhorst 2014).

Lastly, we investigate whether the slope of the futures term structure (roll returns) provides a reliable signal about expected returns in storable commodity markets. The roll return is consistently negative (16 of the 19 markets) and the average roll return is -3.3% per year and statistically significant, indicating that contango is the normal market structure. We find some evidence of a relationship between the slope of the term structure of commodity futures markets and subsequent returns. While the difference in returns between contango markets and backwardated markets can be relatively large, the differences are highly variable with somewhat limited statistical significance. Moreover, the strength of this finding varies widely across markets and fades in the second half of the sample period, leaving investors with little confidence in the value of this signal. In addition, following the signal requires that investors go both long and short in commodity futures markets, something quite different from the passive, long-only exposure desired by the vast majority of commodity investors to date.

We purposely did not examine returns to portfolios of commodity futures in order to focus on the fundamentals of returns at the individual market level. Due to a "rebalancing bonus" or "diversification return," it is possible for portfolios of commodity futures positions to generate a positive return despite unconditional returns of individual commodity futures being near zero. However, the diversification return depends entirely on the strategy used to rebalance the portfolio (Willenbrock 2011), and hence, portfolios of commodity futures represent a form of active rather than passive investment. It is an open question whether investors will find commodity investment attractive if the case depends solely on a portfolio diversification return. Of course, there is always the possibility that sophisticated dynamic strategies can be used to earn positive returns in commodity futures markets (e.g., Boons and Prado 2018), but this is the realm of active rather than passive investment.

9. References

Acharya V.V., Lochstoer L.A., Ramadorai, T. 2013. Limits to Arbitrage and Hedging: Evidence from Commodity Markets. *Journal of Financial Economics* 109 2:441–465.

Ankrim E.M., Hensel C.R. 1993. Commodities in Asset Allocation: A Real-Asset Alternative to Real Estate. *Financial Analysts Journal* 49 3:20-29.

Basak S., Pavlova A. 2016. A Model of Financialization of Commodities. *Journal of Finance* 71 4: 1511–1556.

Bessembinder H. The "Roll Yield" Myth. *Financial Analysts Journal* 74 2: 41-53.

Bessembinder H., Carrion A., Tuttle L., Venkataraman K. 2016. Liquidity, Resiliency, and Market Quality around Predictable Trades: Theory and Evidence. *Journal of Financial Economics* 121 1:142-166.

Bhardwaj G., Gorton G.B., and Rouwenhorst K.G. 2014. "Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors." *Review of Financial Studies* 27 11:3099-3132.

Bhardwaj G., Gorton G.B., and Rouwenhorst K.G. 2015. Facts and Fantasies about Commodity Futures Ten Years Later. Working Paper No. 21243, National Bureau of Economic Research.

Bjornson B., Carter C.A. 1997. New Evidence on Agricultural Commodity Return Performance under Time-Varying Risk. *American Journal of Agricultural Economics* 79 3:918-30.

Black F. 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics* 3 1-2:167-179.

Bodie Z., Rosansky V.I. 1980. Risk and Return in Commodity Futures. *Financial Analysts Journal* 36 3:27-39.

Boons M., Prado M.P. 2018. Basis-Momentum. *Journal of Finance* (forthcoming).

Carter C.A., Rausser G.C., Schmitz A. 1983. Efficient Asset Portfolios and the Theory of Normal Backwardation. *Journal of Political Economy* 91 2:319-331.

Cootner P.H. 1960. Returns to Speculators: Telser vs. Keynes. *Journal of Political Economy* 68 4:396-404

Dugan I.J. 2013. Pension Funds Cut Back on Commodity Indexes. *Wall Street Journal*, February 5. (https://www.wsj.com/articles/SB10001424127887324761004578286172166045096)

Dusak K. 1973. Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums. *Journal of Political Economy* 81 6:1387-1406.

Erb C.B., Harvey C.R. 2006. The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal* 62 2:69-97.

Fama E.F., French K.R. 1987. Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *Journal of Business* 60 1:55-73.

Garcia P., Irwin S.H., Smith A. 2015. Futures Market Failure? *American Journal of Agricultural Economics* 96 1:40-64.

Gorton G., Rouwenhorst K.G. 2006. Facts and Fantasies about Commodity Futures. *Financial Analysts Journal* 62 2:47-68.

Gorton G.B., Hayashi F., Rouwenhorst K.G. 2007. The Fundamentals of Commodity Futures Returns. Working Paper No. 13249, National Bureau of Economic Research.

Gorton G.B., Hayashi F., Rouwenhorst K.G. 2013. The Fundamentals of Commodity Futures Returns. *Review of Finance* 17 1:35-105.

Greer R.J. 1978. Conservative Commodities: A Key Inflation Hedge. *Journal of Portfolio Management* 4 4:26-29.

Greer R.J. 2000. The Nature of Commodity Index Returns. *Journal of Alternative Investments* 3 1:45-52.

Hamilton J.D., Wu J.C. 2015. Effects of Index-Fund Investing on Commodity Futures Prices. *International Economic Review* 56 1:187-205.

Hartzmark M.L. 1987. Returns to Individual Traders of Futures: Aggregate Results. *Journal of Political Economy* 95 6:1292-1306.

Irwin S.H. 1994. Further Evidence on the Usefulness of CTA Performance Information in Public Commodity Pool Prospectuses and a Proposal for Reform. In *Advances in Futures and Options Research*, Don M. Chance and Robert R. Trippi, editors, Greenwich, Connecticut: JAI Press, pp. 251-265.

Kolb R.W. 1992. Is Normal Backwardation Normal? *Journal of Futures Markets* 12 1:75-90.

Labys W.C., Thomas H.C. 1975. Speculation, Hedging and Commodity Price Behavior: An International Comparison. *Applied Economics* 7 4:287-301.

Main S., Irwin S.H., Sanders D.R., Smith A. 2018. Financialization and the Returns to Commodity Investments. *Journal of Commodity Markets* 10:22-28.

Moskowitz T.J., Ooi Y.H., Pedersen L.H. 2012. Time Series Momentum. *Journal of Financial Economics* 104 2:228-250.

Nussbaum A., Javier L.A. 2016. Buy, Hold & Lose: How a Commodities Roll Undercuts Investors. *Bloomberg.com*, September 21. (http://www.bloomberg.com/news/articles/2016-09-21/buy-hold-lose-how-commodities-roll-is-undercutting-investors)

Pindyck R.S. 2001. The Dynamics of Commodity Spot and Futures Markets: A Primer. *Energy Journal* 22 3:1-29.

Samuelson R.J. 2015. Behind the Commodities Bust. *Washington Post*, August 26. (https://www.washingtonpost.com/opinions/behind-the-commodities-bust/2015/08/26/a56b35d4-4c07-11e5-902f-39e9219e574b_story.html?utm_term=.4f035bc43824)

Sanders D.R., Irwin S.H. 2012. A Reappraisal of Investing in Commodity Futures Markets. *Applied Economic Perspectives and Policy* 34 3:515-530.

Shumsky T. 2014. Commodity ETFs Get a Performance Boost This Year. *Wall Street Journal*, July 16. (http://blogs.wsj.com/totalreturn/2014/07/16/commodity-etfs-get-a-performance-boost-this-year/)

Summers, L.H. 1986. Does the Stock Market Rationally Reflect Fundamental Values? *Journal of Finance* 41 3:591-601

Szymankowska M., de Roon F., Nijman T., Goorbergh R.V.D. 2014. An Anatomy of Commodity Futures Risk Premia. *Journal of Finance* 69 1:453-482.

Telser L.G. 1958. Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy* 66 3:233-255.

Tomek W.G. 1997. Commodity Futures Prices as Forecasts. *Review of Agricultural Economics* 19 1:23-44.

Willenbrock S. 2011. Diversification Return, Portfolio Rebalancing, and the Commodity Return Puzzle. *Financial Analysts Journal* 67 4:42-49.

Williams J.C., Wright B.D. 1991. *Storage and Commodity Markets*. Cambridge University Press: Cambridge, UK.

Table 1. Recent Single Commodity Exchange Traded Fund (ETF) and Note (ETN) Performance

Symbol	Fund Name	Commodity	($ millions) Total Assets	Total Return 3 Year	Total Return 5 Year	Expense Ratio
USO	United States Oil Fund	WTI Crude Oil	1,976.5	-32.3%	-58.7%	0.77%
DBO	PowerShares DB Oil Fund	WTI Crude Oil	372.9	-20.6%	-54.3%	0.75%
UNG	United States Natural Gas Fund	Natural Gas	364.1	-59.2%	-74.4%	1.30%
DGL	PowerShares DB Gold Fund	Gold	199.3	5.6%	-17.0%	0.75%
BNO	United States Brent Oil Fund	Brent Crude Oil	98.7	-14.7%	-47.4%	0.90%
USL	United States 12 Month Oil Fund	WTI Crude Oil	88.1	-12.6%	-39.1%	0.86%
CORN	Teucrium Corn Fund	Corn	79.5	-22.7%	-54.8%	1.00%
WEAT	Teucrium Wheat	Wheat	67.1	-34.0%	-63.9%	1.00%
UGA	United States Gasoline Fund	Gasoline	45.9	-15.3%	-40.5%	0.75%
NIB	iPath Dow Jones-UBS Cocoa ETN	Cocoa	42.0	-8.7%	7.7%	0.55%
DBS	PowerShares DB Silver Fund	Silver	20.1	-5.4%	-38.5%	0.75%
SOYB	Teucrium Soybean	Soybeans	16.6	-5.8%	-19.9%	1.00%
CANE	Teucrium Sugar	Sugar	12.8	-24.3%	-52.9%	0.50%
CPER	United States Copper Index Fund	Copper	10.8	-0.6%	-14.0%	0.80%
PTM	E-TRACS UBS Bloomberg CMCITR Long Platinum ETN	Platinum	10.1	-22.5%	-43.7%	0.65%
UHN	United States Diesel-Heating Oil Fund	Diesel-Heating O	8.3	-17.5%	-35.3%	0.75%
OLO	DB Crude Oil Long ETN	Crude Oil	8.0	-21.2%	-52.7%	0.75%
UNL	United States 12 Month Natural Gas Fund	Natural Gas	5.9	-30.6%	-53.6%	0.90%
UBG	E-TRACS UBS Bloomberg CMCI Gold ETN	Gold	3.7	8.8%	-13.3%	0.30%
OLEM	iPath Pure Beta Crude Oil ETN	Crude Oil	2.8	-19.0%	-46.7%	0.45%
UBC	E-TRACS UBS Bloomberg CMCI Livestock ETN	Cattle and Hogs	2.4	-16.0%	2.0%	0.65%
USV	E-TRACS UBS Bloomberg CMCI Silver ETN	Silver	2.1	-5.8%	-37.2%	0.40%
LD	iPath Dow Jones-UBS Lead ETN	Lead	0.4	4.9%	1.7%	0.75%
		Average		**-16.1%**	**-36.8%**	**0.75%**
		Total	**3,438.2**			

Notes: All data are from the ETF Database (etfdb.com) as of May 7, 2018 and includes long-only, futures-based, mechanical, unlevered funds that focus on single market or a very narrow market mix. Funds had to have a 5 year track record to be included.

Table 2. Commodity Futures Markets and Sample Periods

	Exchange	Start Date	End Date	No. Daily Obs.
WTI Crude Oil	NYMEX/CME	3/30/1983	6/30/2017	8,548
Heating Oil	NYMEX/CME	6/5/1979	6/30/2017	9,504
RBOB (Gasoline)	NYMEX/CME	12/3/1984	6/30/2017	8,095
Natural Gas	NYMEX/CME	6/1/1990	6/30/2017	6,772
Gold	COMEX/CME	1/2/1975	6/30/2017	10,622
Silver	COMEX/CME	8/5/1963	6/30/2017	13,375
Copper	COMEX/CME	7/1/1959	6/30/2017	14,390
Corn	CBOT/CME	7/1/1959	6/30/2017	14,613
SRW Wheat	CBOT/CME	7/1/1959	6/30/2017	14,601
HRW Wheat	KCBOT/CME	4/9/1976	6/30/2017	10,378
Soybeans	CBOT/CME	7/1/1959	6/30/2017	14,597
Soybean Meal	CBOT/CME	7/1/1959	6/30/2017	14,565
Soybean Oil	CBOT/CME	7/1/1959	6/30/2017	14,525
Rough Rice	CBOT/CME	8/20/1986	6/30/2017	7,760
Oats	CBOT/CME	7/1/1959	6/30/2017	14,564
Cotton	NYBOT/ICE	12/4/1959	6/30/2017	14,247
Cocoa	NYBOT/ICE	8/4/1969	6/30/2017	11,845
Coffee	NYBOT/ICE	8/17/1973	6/30/2017	10,859
Sugar	NYBOT/ICE	1/4/1961	6/30/2017	13,927

Notes: CBOT: Chicago Board of Trade; CME: Chicago Mercantile Exchange; COMEX: Commodity Exchange; ICE: Intercontinental Exchange; KCBOT: Kansas City Board of Trade; NYMEX: New York Mercantile Exchange.

Table 3. Average Daily Returns (annualized) for Commodity Futures over Full Sample and by Decade, July 1959 - June 2017

	Full Sample		Average Return By Decade					
	Average Return	t-statistic	1960s	1970s	1980s	1990s	2000s	2010s
WTI Crude Oil	1.8%	0.30				5.6%	8.0%	-18.0%
Heating Oil	1.4%	0.28			4.9%	-0.7%	10.2%	-8.7%
RBOB (Gasoline)	7.0%	1.18				6.0%	12.4%	-4.0%
Natural Gas	-23.7%	-2.66				-16.2%	-24.6%	-32.3%
Gold	-0.8%	-0.28			-12.5%	-6.8%	9.0%	1.0%
Silver	-2.1%	-0.53		17.9%	-27.9%	-4.2%	5.3%	-2.5%
Copper	7.1%	2.11		0.6%	0.4%	3.3%	11.3%	-3.6%
Corn	-5.3%	-1.90	-5.3%	5.8%	-7.0%	-8.9%	-11.3%	-5.0%
CBOT Wheat	-5.0%	-1.59	-5.7%	11.9%	-6.8%	-10.6%	-8.8%	-12.2%
KCBT Wheat	-4.1%	-1.18		-1.3%	-3.5%	-3.3%	-2.1%	-10.2%
Soybeans	2.5%	0.88	4.3%	10.5%	-8.8%	-3.8%	9.7%	4.0%
Soybean Meal	1.9%	0.59	4.1%	26.8%	-10.0%	-6.8%	3.5%	-6.9%
Soybean Oil	6.2%	1.88	9.0%	8.4%	-5.4%	-1.9%	15.3%	12.2%
Rough Rice	-7.8%	-1.90				-9.1%	-6.7%	-13.7%
Oats	-3.4%	-0.98	-7.9%	4.8%	-7.3%	-18.2%	5.2%	3.6%
Cotton	-1.1%	-0.38	-2.8%	8.3%	3.6%	-2.1%	-15.7%	3.0%
Cocoa	-0.6%	-0.15	-29.3%	29.2%	-18.9%	-14.1%	9.0%	-8.9%
Coffee	-2.9%	-0.57			-5.0%	1.6%	-18.8%	-9.5%
Sugar	-6.3%	-1.23		13.4%	-23.9%	-5.9%	6.1%	-13.0%
All	-0.3%	-0.19	0.4%	13.6%	-7.1%	-5.4%	1.3%	-6.6%

Notes: All entries are daily annualized returns (log price changes). Not all markets have data for the full July 1959 through June 2017 sample period. See Table 1 for details.

Table 4. Decomposition of Daily Commodity Futures Returns (annualized), July 1959 - June 2017

	Futures		Spot		Roll	
	Average Return	**t-statistic**	**Average Return**	**t-statistic**	**Average Return**	**t-statistic**
WTI Crude Oil	1.8%	0.30	1.9%	0.31	-0.1%	-0.09
Heating Oil	1.4%	0.28	1.8%	0.34	-0.4%	-0.28
RBOB (Gasoline)	7.0%	1.18	2.9%	0.45	4.2%	1.95
Natural Gas	-23.7%	-2.66	3.4%	0.35	-27.1%	-7.32
Gold	-0.8%	-0.28	4.1%	1.39	-4.9%	-12.85
Silver	-2.1%	-0.53	3.9%	0.99	-5.9%	-14.60
Copper	7.1%	2.11	3.6%	1.05	3.4%	3.88
Corn	-5.3%	-1.90	2.0%	0.67	-7.3%	-7.01
CBOT Wheat	-5.0%	-1.59	2.0%	0.58	-6.9%	-5.81
KCBT Wheat	-4.1%	-1.18	1.2%	0.34	-5.4%	-4.89
Soybeans	2.5%	0.88	2.6%	0.89	-0.1%	-0.13
Soybean Meal	1.9%	0.59	2.4%	0.71	-0.5%	-0.54
Soybean Oil	6.2%	1.88	3.5%	1.03	2.7%	2.84
Rough Rice	-7.8%	-1.90	3.7%	0.85	-11.6%	-7.11
Oats	-3.4%	-0.98	2.5%	0.67	-6.0%	-4.22
Cotton	-1.1%	-0.38	2.4%	0.77	-3.5%	-3.06
Cocoa	-0.6%	-0.15	2.9%	0.66	-3.5%	-3.27
Coffee	-2.9%	-0.57	1.2%	0.23	-4.1%	-3.14
Sugar	-6.3%	-1.23	1.6%	0.28	-7.9%	-4.40
All	-0.3%	-0.19	2.9%	1.63	-3.3%	-7.19

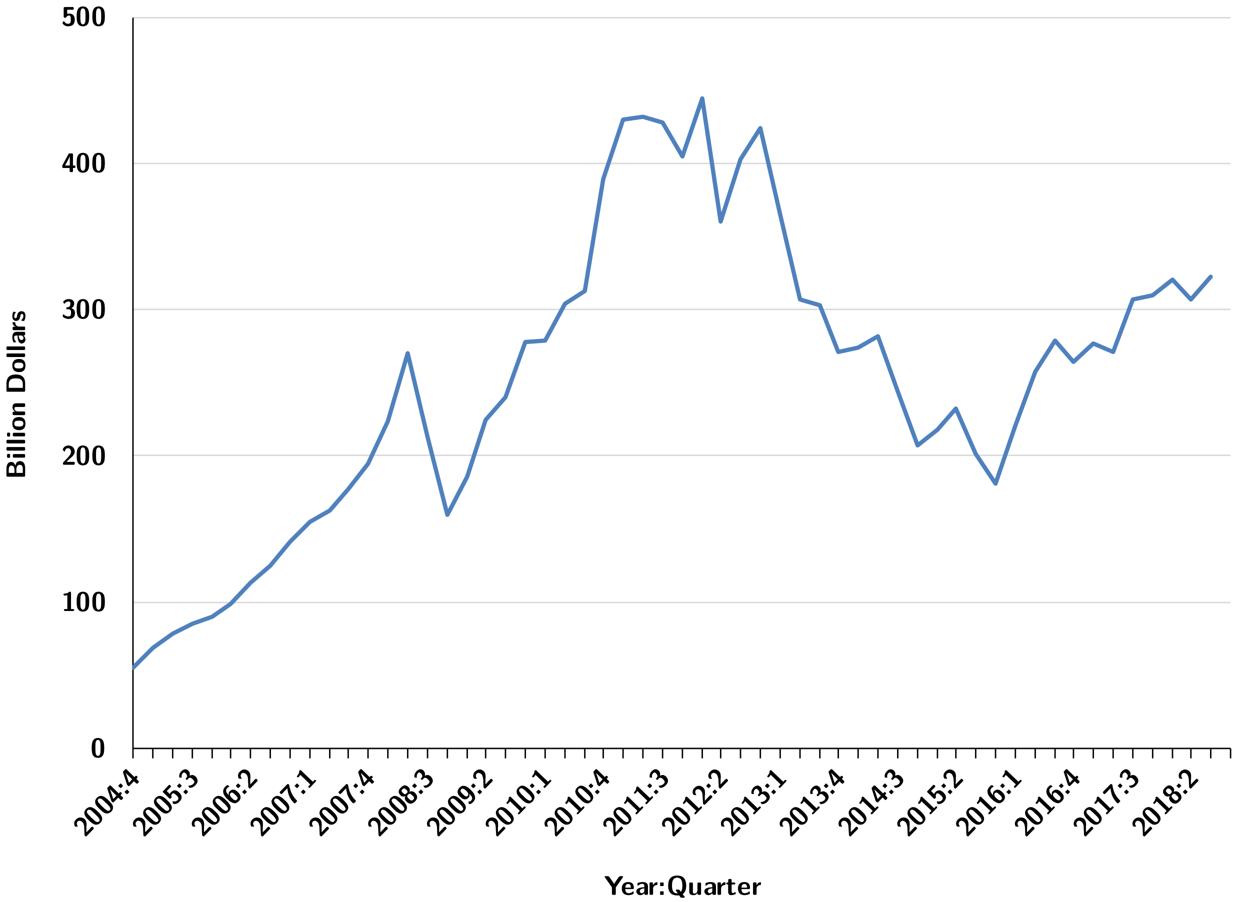
Notes: All entries are daily annualized returns (log price changes). Not all markets have data for the full July 1959 through June 2017 sample period. See Table 1 for details. The numbers presented in the table may not add up precisely (futures return = spot return + roll return) due to rounding.

Table 5. Regression Results for Difference in Average Daily Return (annualized) for Commodity Futures Markets in Contango versus Backwardation, July 1959 - June 2017

	July 1959 - June 2017			July 1959 - December 1988			January 1989 - June 2017		
	Dummy Coefficient	t-statistic	Proportion in Contango	Dummy Coefficient	t-statistic	Proportion in Contango	Dummy Coefficient	t-statistic	Proportion in Contango
WTI Crude Oil	-23.6	1.98	54%	-14.0	0.44	24%	-27.7	2.05	61%
Heating Oil	-14.2	1.29	68%	3.9	0.22	56%	-22.3	1.64	71%
RBOB (Gasoline)	1.5	0.12	43%	-7.0	0.21	31%	2.6	0.20	44%
Natural Gas	1.5	0.07	82%				1.5	0.07	82%
Gold	-14.1	0.56	99%				-12.7	0.59	98%
Silver	22.9	0.56	99%	-2.2	0.03	99%	36.2	0.73	99%
Copper	-26.9	3.99	57%	-46.0	4.98	56%	-7.1	0.72	57%
Corn	-2.2	0.32	79%	-11.0	1.53	70%	22.3	1.57	88%
CBOT Wheat	-8.2	1.01	82%	-4.2	0.47	76%	-10.0	0.66	87%
KCBT Wheat	-19.7	2.39	76%	-24.8	2.35	68%	-16.9	1.48	80%
Soybeans	0.5	0.09	70%	5.3	0.61	72%	-4.2	0.47	69%
Soybean Meal	-19.9	2.76	72%	-21.0	2.12	56%	-8.4	0.64	89%
Soybean Oil	-8.1	1.23	56%	-22.4	2.26	64%	5.6	0.62	48%
Rough Rice	42.9	3.39	88%	102.4	1.93	85%	37.6	2.90	88%
Oats	-10.2	1.35	69%	-0.1	0.01	64%	-21.4	1.71	74%
Cotton	-2.4	0.39	68%	-2.9	0.40	64%	-0.2	0.02	73%
Cocoa	-27.6	2.81	74%	-53.3	3.83	59%	9.4	0.62	85%
Coffee	-22.6	2.00	70%	-19.4	1.30	47%	-13.6	0.76	82%
Sugar	-4.8	0.45	62%	-14.5	0.76	71%	5.8	0.49	54%

Notes: Not all markets have data for the full July 1959 through June 2017 sample period. See Table 1 for details.

Figure 1. Total Global Commodity-Linked Investment, Fourth Quarter 2004 – Third Quarter 2018.



Source: Barclays

Figure 2. Nearby Corn Futures Prices and Teucrium Corn ETF Share Price, January 2015-May 2018.

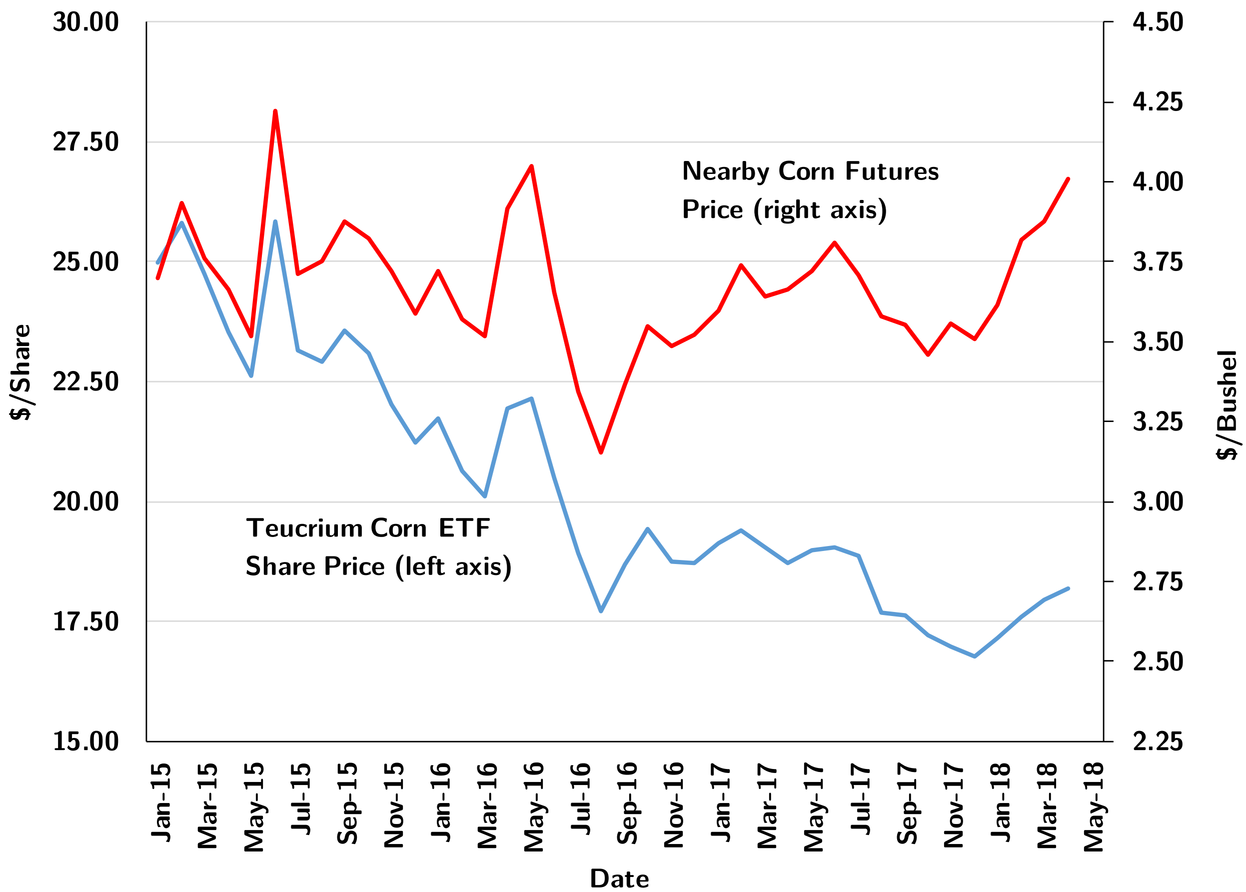
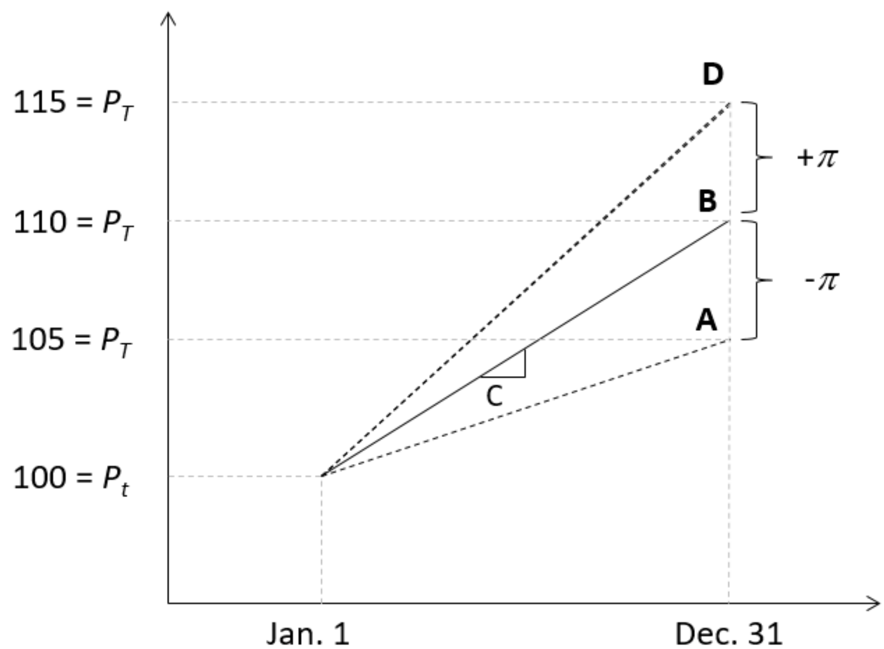


Figure 3. Pricing of Storable Commodities under Contango and Backwardation Term Structure

Panel A: Contango Term Structure



Panel B: Backwardation Term Structure

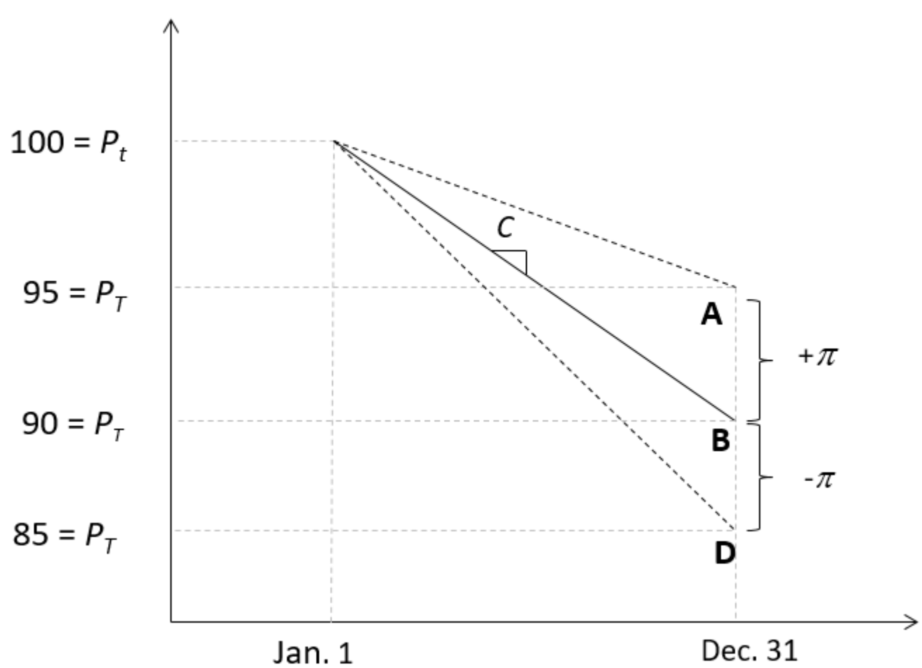


Figure 4. Daily U.S. Oil Fund (USO) Share Price Compared to WTI Crude Oil Price and a Simulated ETF, April 10, 2006 – June 30, 2017

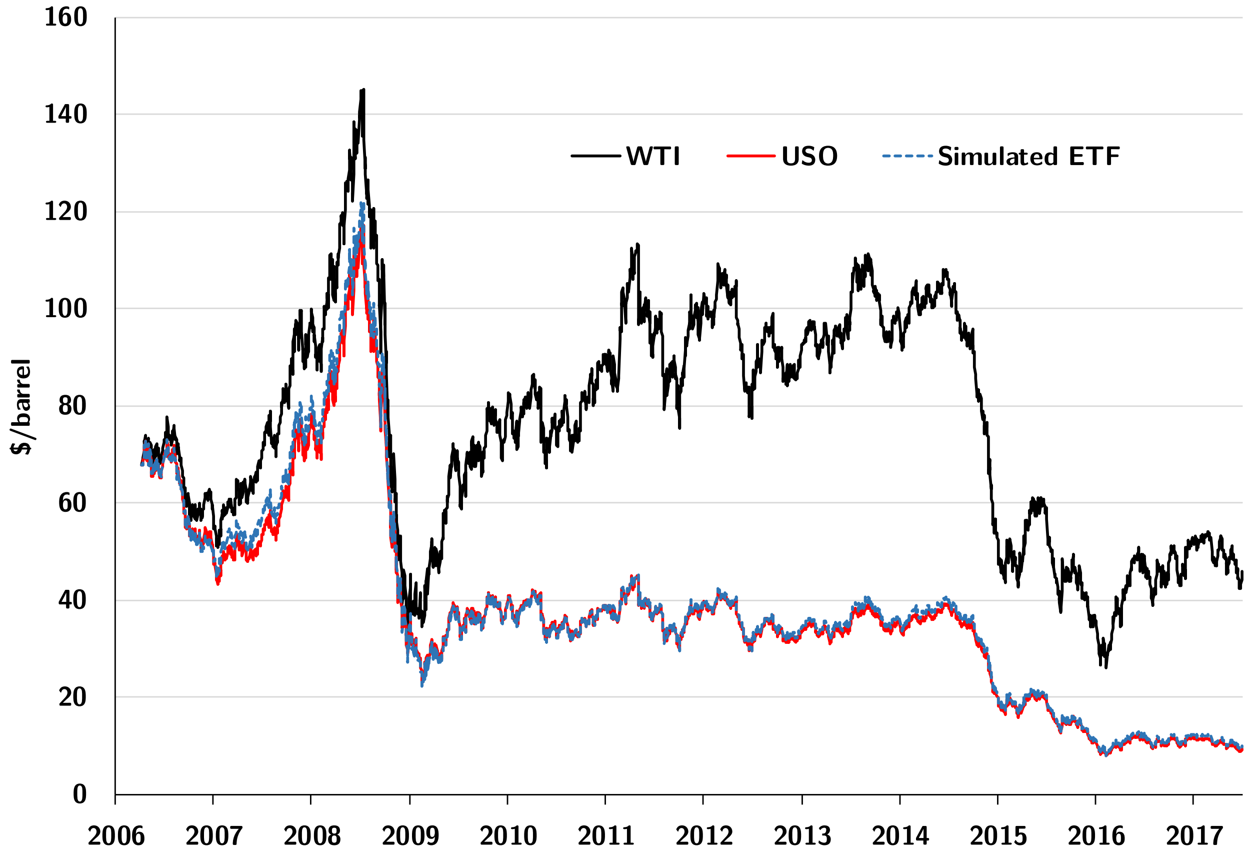


Figure 5. Daily Powershares Gold Fund (DGL) Share Price Compared to LBMA Gold Price and a Simulated ETF, January 5, 2007 – June 30, 2017

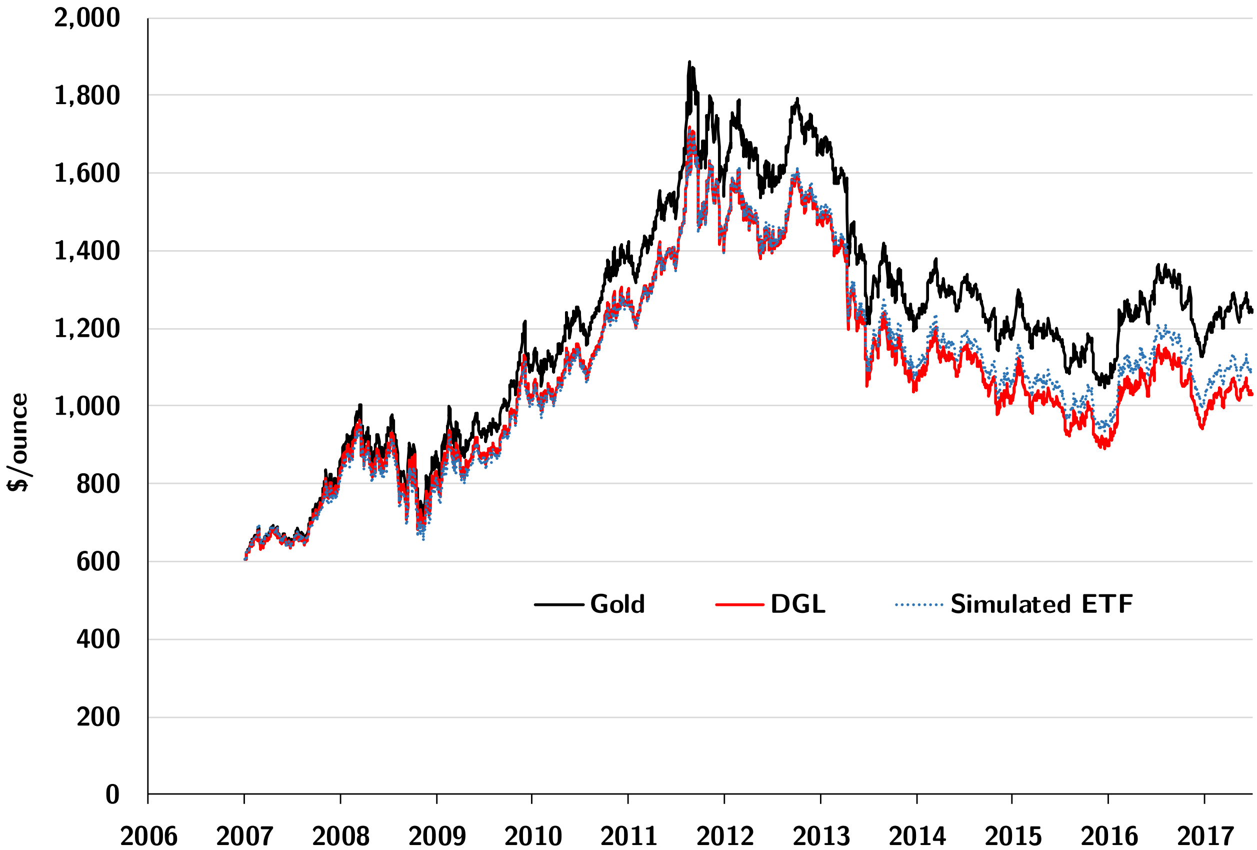


Figure 6. Daily U.S. Natural Gas Fund (UNG) Share Price Compared to Henry Hub Natural Gas Price and a Simulated ETF, April 18, 2007 – June 30, 2017

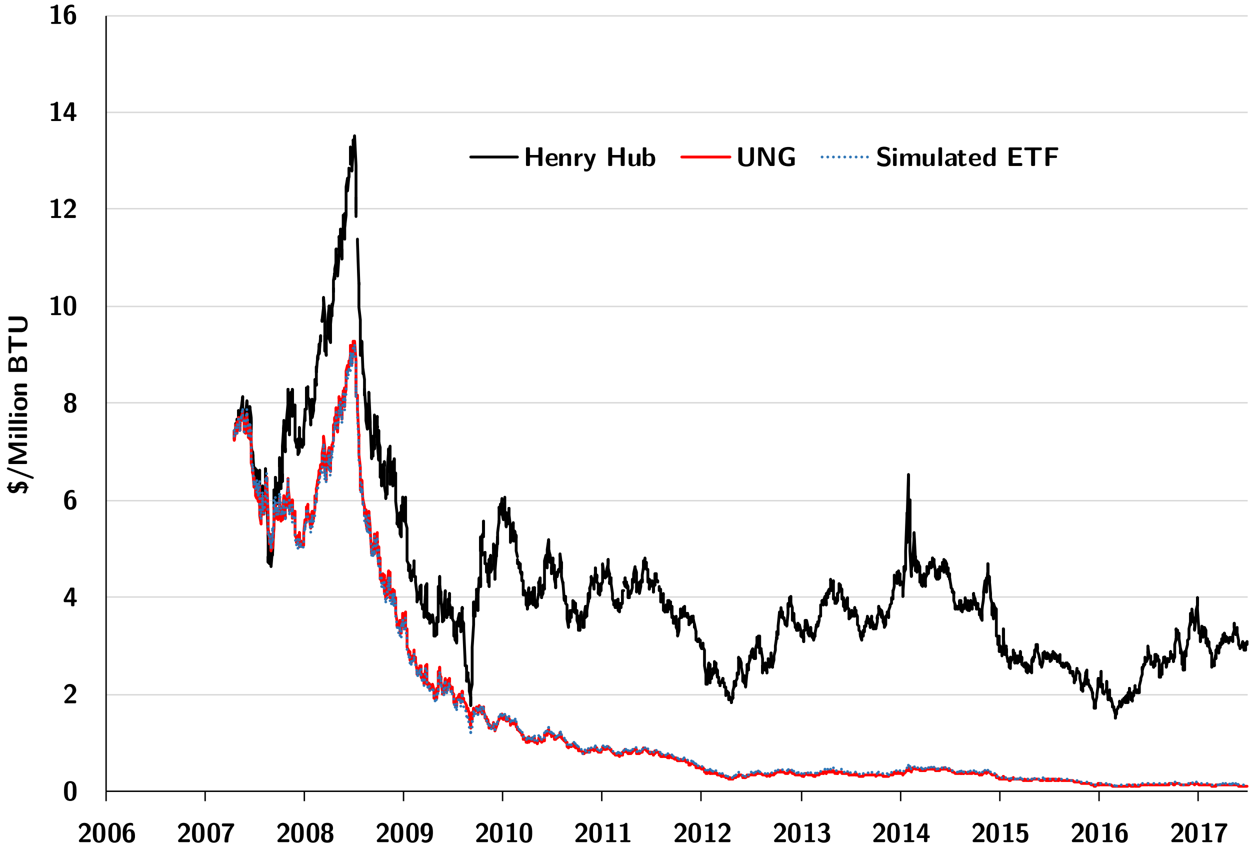


Figure 7. Bootstrap Simulation of the Probability of the Average Futures Return over a 10-Year Horizon Being Greater than Zero for 19 Storable Futures Markets, July 1959 – June 2017

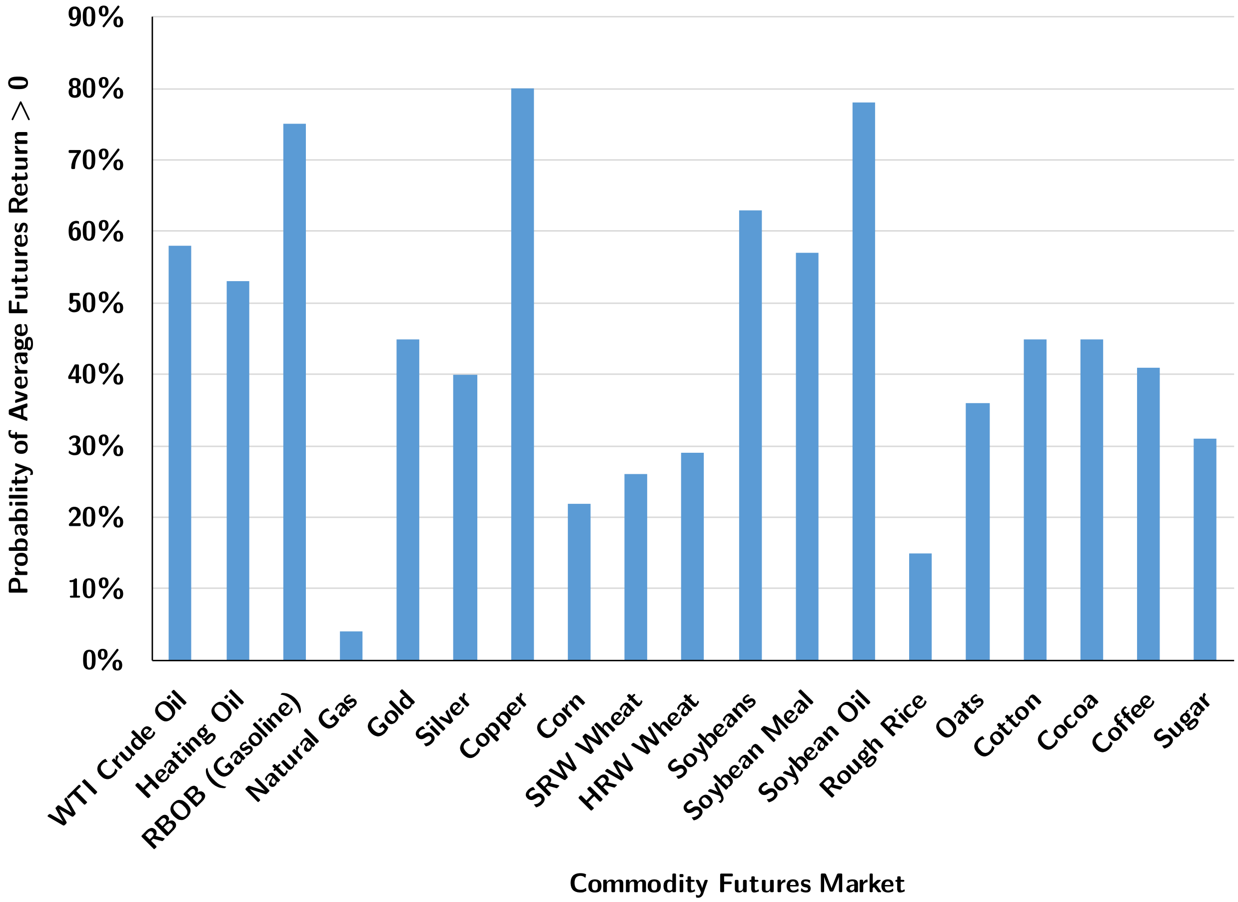
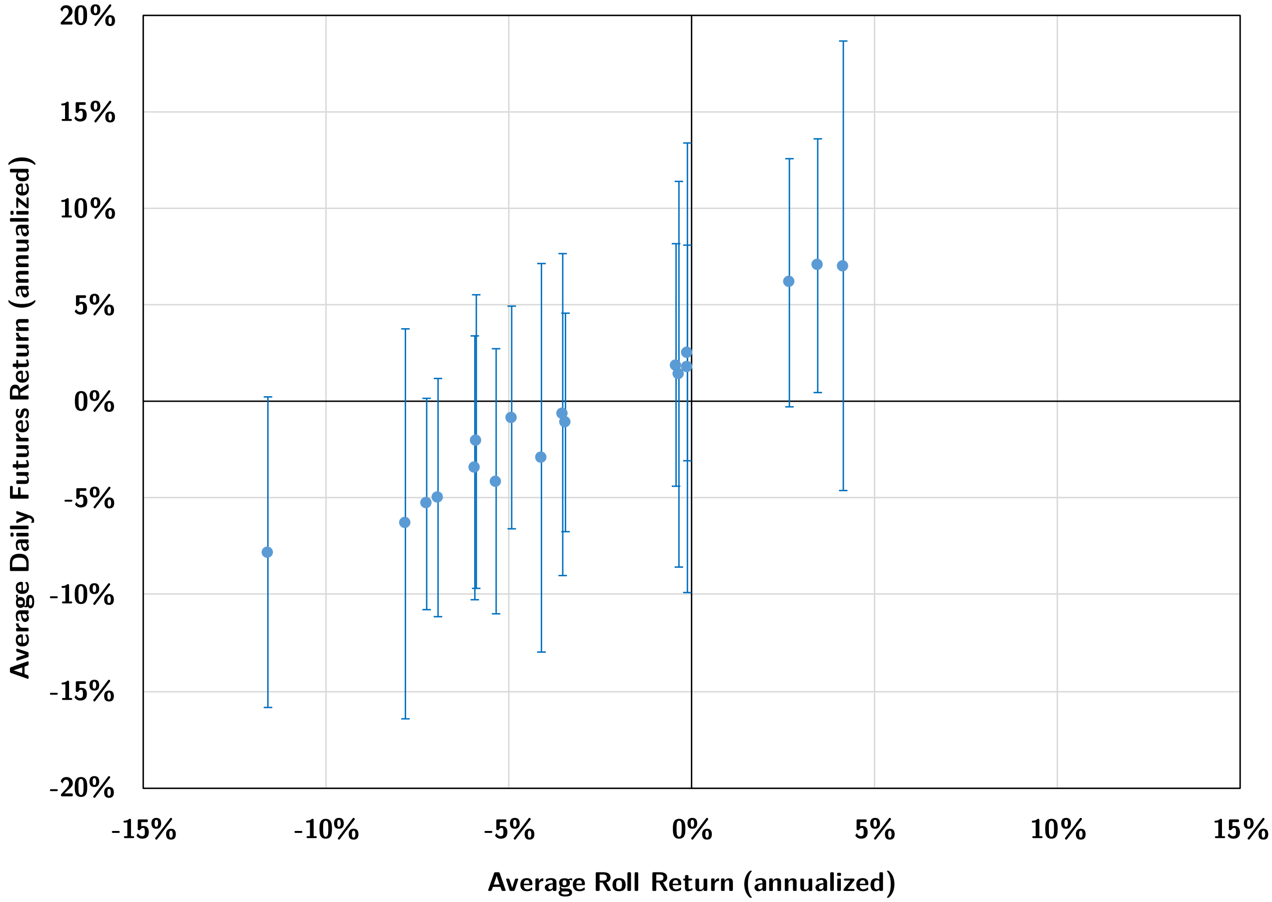


Figure 8. Average Daily Roll and Futures Returns (annualized) for 18 Storable Futures Markets (natural gas excluded), July 1959 – June 2017



Notes: The observations denoted with filled circles represent full sample averages for daily futures and roll returns (annualized). The error bars correspond to 95% confidence intervals for the mean of each futures return series.