The valuation effects of index investment in commodity futures

Michel Dubois* and Loïc Maréchal†

January 15, 2020

Abstract

We identify and date a significant surge in the amount of investment tracking commodity futures indices, a phenomenon identified heretofore with anecdotal or visual evidences. Using a difference-in-differences setting on cumulative abnormal log price changes, computed with several benchmarks during the roll window of the SP-GSCI, we first find that the uncovered break in the speculative investment structure had an alleviating effect. Second, we explain the abnormal nearby and first deferred contracts price changes by measures of risk (liquidity) premium required at long (short) term horizon by speculative (hedging) activity. Finally, in a cruder market efficiency framework, we find that transaction costs incurred by an arbitrager (price taker) explain most of the abnormal term-structure change with a coefficient close to unity. In addition, this abnormal change -which is of 17 basis points at most- is never significant once we adjust the standard errors for event-induced variance and cross-correlation.

JEL classification: G24, G28, K22, K42.

Keywords: commodity index investment, price pressure, risk premium, liquidity premium, hedging pressure, sunshine trading, predatory trading.

*University of Neuchâtel, michel.dubois@unine.ch

†University of Neuchâtel, loic.marechal@unine.ch

We thank Adrian Fernandez-Perez for his comments and suggestions as discussant during the Queenstown Derivatives Market Conference, as well as the participants for helpful discussions. We also thank the participants of the Non-Standard Investment Choices Workshop held at ESSEC, in particular Andrea Roncoroni, Romeo Tedongap and Sofia Ramos.

1. Introduction

Academic research shows that commodity investment improves portfolio diversification (Bodie and Rosansky, 1980), as commodity prices are contracyclical (Gorton and Rouwenhorst, 2006), permit to hedge against unexpected inflation (Greer, 2000), are negatively correlated with the U.S. Dollar (Rezitis, 2015) and are positively skewed (Deaton and Laroque, 1992).[1] These characteristics make this asset class particularly attractive. However, a direct investment in commodities is not easily achievable. In particular, long-term holdings are difficult to maintain because of the physical nature of the underlying (storage costs, decay and logistics). Exchange Traded Funds (ETFs), Exchange Traded Notes (ETNs) and mutual funds (commodity index traders, hereafter, CITs) are popular vehicles to track commodity futures indices such as "first generation" indices, which are long-only, passive and arithmetically weighted upon the relative importance of the commodity in the economy or on the liquidity of the futures contract.

As the primary objective of index investment is to obtain an exposure to the underlying commodity, these indices mimic the performance delivered by long positions in the nearest to maturity (hereafter, nearby) futures contracts to minimize the term-structure effects and maximize the liquidity. Hence, these indices roll their positions at fixed time intervals before the underlying futures contracts mature. The products that replicate these indices represent a significant part of the long open interest (OI) of commodity futures. As they must fully roll each position onto the next contract (hereafter, first deferred) before maturity, there are suspicions that such investment distorts the market. In several news articles, Dizard (2007, 2009) points out that the beneficiaries of the "congestion roll trade" or "date rape" are speculators on the floors of commodities exchanges or, more likely, the banks selling securities based on these indices.[2] In a public hearing before the U.S. Senate, Masters (2008) declared that "Index speculators have driven futures and spot prices higher". To support his view, he shows that the proportion of index speculators exceed that of traditional speculators. He also highlights that index speculators' demand is unrelated to the supply and demand of the underlying commodities.

In this paper, we examine whether the increase of CITs in market participation, a phenomenon coined as "financialization" of commodity markets, has a material effect on commodity futures prices at the date the index rolls positions from the nearby to the first

[1] For an exhaustive review of these stylized facts, see also Gorton and Rouwenhorst (2006) and Erb and Harvey (2006).

[2] "Speculators profit from commodity investors", Financial Times, January 22, 2007. See also, "Goldman Sachs and its magic commodities box", Financial Times, February 5, 2007, and "U.S. oil fund finds itself at the mercy of traders", Wall Street Journal, March 6, 2009.

deferred contract. To address these questions, we focus on the constituents of the Standard and Poor's Goldman Sachs Commodity Index (SP-GSCI).[3]

The financialization should translate into a significant change in OI of the contracts included in the SP-GSCI and more specifically in the proportion of OI resulting from index investment measures.[4] However, there is no official date for this change in market participation. Therefore, we look empirically for a structural change in the index investment share of total OI over the period from 1999 to 2010. We find a significant change occurring in December 2003, similar to the date used in Boons, Roon *de*, and Szymanowska (2014), and five years before the breaks identified on crude oil futures alone by Hamilton and Wu (2015). Moreover, we identify an up break in the amount of arbitrage capital in January 2006. Finally we confirm that both dates do not overlap with the staggered introduction of electronically traded contracts between August 2006 and October 2008.

If the financialization has a permanent effect on commodity futures markets, it should be visible when funds roll their positions from the nearby to the first deferred contract. The SP-GSCI rolls its position from the fifth to the ninth business day of the month preceding maturity (hereafter, the roll). Every day, the index transfers 20% of its positions from one contract to the next.[5] We focus on this particular roll window (hereafter, the roll), because it overlaps most of the other windows of first generation indices such as the BCOM or the US Funds on diversified and single commodities. We analyze the term structure changes during the roll, and estimate the valuation effect of the financialization of commodity markets with an event study. We use several parametric and non-parametric benchmarks to obtain counterfactuals of futures prices, which all deliver consistent empirical results. We also control for event clustering, as at least 16 over 27 constituents of the SP-GSCI roll every month. With a difference in differences setting, we look for the effect of financialization on the cumulative abnormal price changes (CAPCs). Next, we investigate the fact that if the financialization has an impact on the overall commodity market structure we should be able

[3]The three major cross-sector indices in terms of tracking OI are the SP-GSCI (formerly GSCI), the Bloomberg Commodity Index (BCOM, formerly Dow Jones UBS Commodity Index or DJ-UBSCI) and the Deutsche Bank Commodity Index (DBCI). Other diversified indices exist, such as the Reuters-Jefferies CRB Index (RJ-CRB), the Rogers International Commodity Index (RICI), the Chase Physical Commodity Index, Pimco, Oppenheimer or Bear Sterns. Other non-diversified funds directly track subsectors (energy, agricultural futures) or single futures (crude oil, natural gas, gold).

[4]We focus on the SP-GSCI contracts, but our main tests use the value of the total index investment targeting these contracts. In robustness tests, we use in turn the investment level imputed for the SP-GSCI alone, as well as the assets under management (AuM) of the iShares SP-GSCI, the largest ETF tracking the SP-GSCI available since 2006. We do not find any significant differences in our results.

[5]The roll occurs at the same time for every contract, although the actual contract maturity varies. For instance, the last trading day of the sugar NY#11 contract traded on the NYMEX is the last business day of the month preceding the quoted month, whereas for the grains contracts traded on the CME, it is on the 15th business day of the quoted month.

to explain the CAPCs around the roll with the measures of risk and liquidity premia of Kang, Rouwenhorst, and Tang (forthcoming). Finally, in complement of former studies (see, *e.g.*, Mou, 2011) we integrate market frictions in our analysis of market efficiency and focus on the cumulative abnormal term structure changes (CATCs) during the roll. We explain it with low frequency measures of transaction costs and liquidity.

We confirm that, with no correction for event-induced variance and cross-correlation, the average cumulative term structure changes are positive, of 15 basis points (bps) during the roll and 17 bps during the pre-roll, and statistically significant at the 1% level up until 2003. However, once we correct for clustering, the significance becomes void. When we consider individual futures contracts, we find that CAPCs are not statistically significant in any period and even before adjustement. In contradiction with previous research, we find that the financialization alleviates the magnitude of both CAPCs and CATCs. Nonetheless, the variables ought to proxy for the insurance and liquidity premia explain correctly these abnormal changes. Lastly, we uncover that transaction costs are of similar magnitude to that of the CATCs. In brief, we find that the residual profit left by arbitragers is of the size of their potential transaction costs, would they enter the trade. Our results contribute to the view that index investment has not been detrimental for the functioning of the commodity futures markets themselves. Yet, we cannot discard the possibility of intraday predatory trading achieved by funds managers ahead of their hedging trades, penalizing the performance from the point of view of a CIT investor.

2. Prior literature and hypotheses development

2.1. Dating the financialization

Following Master's hearing before the U.S. Senate, a series of articles examine whether the financialization of commodity futures affects commodity prices. In their survey, Cheng and Xiong (2014) list the consequences of this structural break: the increase of price pressure phenomena, the effects on risk sharing between hedgers and speculators, and the distortion of stock prices resulting from speculation. In this research, we explain how and when the financialization should materialize, what are the effects of the roll on futures prices and what are their determinants. Masters (2008) shows that the proportion of CITs in the total open interest changed dramatically from 1998 to 2008, which makes this ratio a natural candidate to look for a potential break caused by index investors. Naturally, our first alternative hypothesis is,

Hypothesis 1a. *The share of index investment in total open interest of commodity futures contracts, constituents of the SP-GSCI, show a common and significant break over the 1999–2010 period.*

A second and subsequent element of the financialization is the rise of hedge funds and other active management-related positions in commodity markets. Two reasons may underlie this phenomenon. First, as for passive investment, active funds are willing to get exposure to commodity markets to improve their returns and benefit from a better timing. Second, they may want to benefit from the arbitrage opportunities generated during the roll by passive investment. In fact, Stoll and Whaley (2010), Mou (2011) and Irwin and Sanders (2012) find that limits to arbitrage are the reasons for price pressure effects during the roll. Hence the hypothesis is,

Hypothesis 1b. *The share of arbitrage capital in total open interest of commodity futures contracts, constituents of the SP-GSCI, show a common and significant break over the 1999–2010 period.*

Finally, Raman, Robe, and Yadav (2017) study the "electronification" of commodity futures markets, an event they coin as the "third dimension" of financialization, which is the introduction, for nearly all the indexed commodities, of electronically traded futures contracts between August 2006 and October 2008. Focusing, in particular, on the U.S. crude oil contract, they show that the electronification leads to improved market quality measures, such as bid-ask spread sizes and market depths, in combination with a larger intraday activity. Other research on the effects of electronification includes Martinez, Gupta, Tse, and Kittiakarasakun (2011) who find that side by side trading of pit and electronically traded contracts lead to a shift of market quality from the former to the latter and Gousgounis and Onur (2018) who study the converse effects of pit closures following the side by side trading, although only from the perspective of pit trading. Because of the potential overlap of the electronification with the aforementioned financialization dimensions, we write our third hypothesis as,

Hypothesis 1c. *The electronification dates do not overlap with the breaks in index investments and arbitrage capital measures.*

2.2. Informationless trading

2.2.1. Price pressure

Grossman and Miller (1988) present a three-period model in which there are two types of agents, market makers and outside customers. They derive the demand function assuming that prices are normally distributed, and investors maximize their expected (exponential) utility. In their setting, they analyze the consequences of a liquidity shock that creates a temporary order imbalance, and show how market makers are compensated for bearing the risk during the holding period. They show that the (absolute) expected returns are an increasing function of the order imbalance, and an inverse function of the number of market makers. In our case, assuming that the market cannot absorb instantaneously the positive demand (supply) shock on the first deferred (nearby) contract, we should observe an increase (decrease) in the current price. Brunetti and Reiffen (2014) derive a two period model where CITs' demand is exogenous, and additional assumptions similar to those of Grossman and Miller (1988). Consistent with their prediction, they empirically find that, as index traders roll their positions, the slope of the term structure between the two nearest contracts increases. In contrast, Aulerich, Irwin, and Garcia (2013) find a significant association for the index traders' roll with term structure change but of opposite direction. Henderson, Pearson, and Wang (2015) examine the impact of the flows of financial investors on commodity futures prices through commodity-linked notes (CLNs). These flows, generated by CLN issuers hedging their liabilities on the commodities futures markets are also informationless. Nevertheless, they increase (buying pressure) or decrease (selling pressure) commodity futures prices. These results are consistent with Bessembinder, Carrion, Tuttle, and Venkataram (2016) who study the roll of the United States Oil Fund (USO), a large fund invested solely in WTI crude oil futures contracts. They show that accumulated trading costs resulting from this price pressure amounts to 3% per year, we specify the following testable alternative hypothesis,

Hypothesis 2a. *CAPCs of the nearby (first deferred) SP-GSCI contracts are significantly negative (positive) over the roll and are larger in the post-financialization period.*

2.2.2. Predatory trading

Brunnermeier and Pedersen (2005) study a distressed trader, who reveals some information to predatory traders (*e.g.*, its broker dealer). In their setting, the predatory trader knows the trader's needs to liquidate quickly her position. Their model foresees a higher price impact as the predator trades on this information along or before the distressed trader. Thus, the liquidation value of the distressed fund decreases. It also predicts that the more predatory traders compete, the lower the (permanent) price impact. With an infinity of predatory traders the effect is similar to the one of price pressure, a polar case that we assume, as roll-related information is public. Mou (2011) empirically explores the performance of a strategy that front runs the SP-GSCI roll by five and 10 days from 2000 to 2010. Both strategies take a long (short) position in the first deferred (nearby) contract. These strategies generate an abnormal performance for the SP-GSCI constituents and not for non-indexed commodities. The best performing strategy, which takes positions ten days ahead the roll, delivers an average monthly 31 bps per roll, before transaction costs. He attributes these findings to the limits of arbitrage capital which implies remaining profitable front trading opportunities. Furthermore, he estimates the costs of this lack of arbitrage to USD 8.4 billion, for a total index investment of USD 211 billion, as of 2009. Conversely, Bessembinder et al. (2016) cannot identify systematic use of predatory strategies and find more liquidity providers during the USO roll.

Hypothesis 2b. *CAPCs of the nearby (first deferred) SP-GSCI contracts are significantly negative (positive) over the pre-roll and are larger in the post-financialization period.*

2.2.3. Sunshine trading

CITs mention their roll policy which closely follows that of the index they track. For the SP-GSCI futures, the roll is monthly for 13 commodities and occurs every two, three or more months for the remaining 14 commodities.[6] The amount transferred during the roll is trivial to estimate since most investment vehicle's AuM is publicly available. Therefore, roll-related information such as funds identity and size is known to market participants.[7] In their model, Admati and Pfleiderer (1991) assume that some liquidity traders preannounce the size and timing of their orders to let liquidity providers get into position and reduce their price impact. The preannouncement reduces informational asymetry and equalize the supply and demand for assets, in particular, when trades are not related to information. The trading costs of traders that preannounce their trades are reduced, but the effect on other traders is ambiguous (trading costs and welfare). This practice, known as "sunshine trading", is of particular interest because it shares many commonalities with CITs' roll. However, index fund managers neither explicitly announce their trades, nor the intraday timing within the window. Again, the results of Bessembinder et al. (2016), and the higher roll liquidity they uncover, support the sunshine trading hypothesis. We specify the following null hypothesis,

Hypothesis 2c. *CAPCs magnitude of the nearby and first deferred SP-GSCI contracts over the roll are significantly lower in the post-financialization period.*

[6] See the contract description in Table 1.

[7] At first glance, the roll looks very much like stock index inclusion (first deferred contract) and deletion (nearby contract). However, there are two notable differences. First, changes in stock index composition are less frequent (20 per year on average for the S&P 500) than changes in the SP-GSCI (more than 250 per year). Second, the index is built on stocks (positive net supply), and not on futures contracts (zero net supply).

2.3. Price effects and the commodity market structure

2.3.1. Hedging pressure

As producers sell their output via futures contracts to hedge against future price fluctuations, they push the futures price below the expected spot price, providing positive premium to long futures investors. This theory of "normal backwardation" (Keynes, 1930; Hicks, 1946) assumes that the bias is systematically negative as a producer is more likely to pay a premium than a buyer.[8] Empirical studies on commodity futures premium use hedging pressure measures, generally computed as the difference between short and long positions of hedgers scaled by the total open interest. This permits tests of normal backwardation in the strict Keynes (1930) sense, but also for bi-directional hedging pressure effects, allowing for negativity in both measures of hedging pressure and premium. Gray (1961) detects such negative risk premium from the point of view of a long investor in this framework. Using Hirshleifer (1988) risk premium model, Bessembinder (1992) shows that market-risk residual futures price changes are non-zero, conditional on hedging pressure. Roon *de*, Nijman, and Veld (2000) study the "cross-hedging pressure" *i.e.*, the hedging pressure effect of groups of futures contracts whose underlying are substitutable. They uncover that commodity risk premium is function of both own and related hedging pressure.

2.3.2. Hedging pressure and speculative demand

Acharya, Lochstoer, and Ramadorai (2013) hedging pressure model incorporates speculative capital constraints, in addition to the producer risk aversion to explain both spot price and commodity futures risk premium. The inclusion of speculative limits to arbitrage and the subsequent higher premium explains why the producer depresses its hedging demand. Instead of hedging, the producer reduces its inventories, thereby decreasing the spot price and modifying the risk premium through both channels of expected commodity spot prices and futures prices (see also, Cheng, Kirilenko, and Xiong, 2015). Similarly, Kang et al. (forthcoming) extend the hedging pressure framework and test the hypothesis that, as speculators have a shorter-term horizon than hedgers, they are themselves in need for liquidity, which generates a second risk premium. This clarifies why some empirical tests on hedging pressure fail to identify risk premium when there is no control for liquidity. For instance, Gorton, Hayashi, and Rouwenhorst (2012) cannot detect a link between risk premium and hedgers' positions. Within this framework, index funds investment must distort both pre-

[8] A producer (*e.g.*, a farmer) receives an income which depends entirely on its production and thus will be more likely to short futures at a discount than a processor (*e.g.*, an airline company), which can report part or all of the price fluctuation in its output price, to buy futures at a premium.

mia and both index investment measures and hedging pressure measures explain the CAPCs during the roll. Therefore our hypothesis is,

Hypothesis 3. *Liquidity (insurance) premium are positive (negative) determinants of the CAPCs during the roll.*

2.4. Market efficiency

In this final analysis of market efficiency, we discard the individual asymmetric effects and study the CATCs, that are the difference of CAPCs between the first deferred and nearby contracts. We hypothesize that the magnitude of CATCs are directly related to market frictions (in the form of liquidity and transaction costs). Indeed, Mou (2011) finds the magnitude of the term structure effects to be of 31 basis points at most. Consider an investor, price taker, willing to arbitrage the roll. Her investment strategy implies to open and close positions on two legs, which entails to pay twice the size of the bid-ask spread. Additional market depths limitations may also further increase this cost.[9] Whereas Mou (2011) estimates the costs to be as low as one bp per contract, taking the, most liquid, crude oil futures contract as example, we hypothesize that these costs are in reality much higher, in particular for less liquid contracts. Stoll and Whaley (2010) also relate the magnitude of the CATCs with their view of the bid-ask spread, with no formal test, however. Thus, in this framework, the market frictions incurred by the arbitrager are the main determinants of the CATCs. The roll generates market distortion, whose magnitude is bounded by the size of the costs of the arbitrage strategy: as soon as the CATCs exceeds the transaction costs, any additional distortion is arbitraged away. Hence our hypothesis is as follow,

Hypothesis 4. *Market frictions, in the form of liquidity and transaction costs, are positive determinants of the CATCs.*

3. Methodology

3.1. Dating the financialization

To test **H1a**, we define the variable $IND_{c,t}$, as $\frac{VCIT_{c,t}}{TOI_{c,t}}$, that is the share of total index investment in total open interest, for each contract c and day t.[10] As a result of the finan-

[9] In addition, we also do not include in our hypothesis, the additional costs of exchange fees and execution timing risk.

[10] We explicitely defined all variables of this study in Appendix D. In Appendix C we add a functional chart of the index investment vehicles.

cialization, we should observe a common break simultaneously affecting IND for the futures contracts constituents of the SP-GSCI.

Since 1986, the weekly "Commitment of Traders" (COT) report of the CFTC provides the long and short positions of the "commercial" and the long, short and spreading positions of the "non-commercial" categories for every U.S. futures contract.[11] In 2006 and 2009, the CFTC revised these categories, and added the "supplemental" and "disaggregated" reports, respectively. The CFTC disaggregated the positions because of the classifications of most swap dealers into the commercial category. Despite their role is non-speculative, they dedicate a large part of their activity to hedge indices-related positions.[12] The disaggregated report splits the original categories of commercial and non-commercial traders, in hedgers, swap dealers, managed money, and other reportables.[13] The supplemental report refers to 13 agricultural futures contracts, and adds a CIT category that aggregates all the positions reported in the aforementioned categories, that are managed for commodity investment vehicles: ETFs, ETNs and funds. We download the COT report data from the CFTC website and compute the variable IND for the SP-GSCI futures contracts. To compute the numerator of IND, we use the total index investment information of Masters and White (2008), the total OI, the SP-GSCI weights, the underlying quantities and the futures prices.[14] We then apply the Masters' algorithm (see, *e.g.*, Masters, 2008; Mou, 2011; Sanders and Irwin, 2013) and control our results in terms of magnitude and correlation with the 13 available futures contracts of the supplemental report.[15] We describe the procedure in Appendix A. In Figure 1, we plot the monthly total index investment, total arbitrage capital, total OI and total trading volume, for the 27 selected contracts and expressed in USD.[16] We also exploit the fact that the legacy COT report displays most of the swap dealer position along with the commercial long positions to reproduce the test with an alternative measure $IND^{control}$ defined as $\frac{COM^L_{c,t}}{TOI_{c,t}}$, which allows to look for a break in both SP-GSCI and non-indexed commodities

[11]Precisely for each contract that has at least 20 active trader positions above an individual threshold defined by the CFTC. See http://www.cftc.gov/MarketReports/CommitmentsofTraders/index.htm

[12]For the same reasons, swap dealers could claim position limits exemptions, as if they were producers or processors in needs of hedging.

[13]In the disaggregated COT, the official CFTC denomination for hedgers becomes: "Producer/Merchant/Processor/User".

[14]We thank S&P Global for providing us with the historical SP-GSCI weights.

[15]Sanders and Irwin (2013), argue that the Masters algorithm produces index investment figures that are sensitive to the low weighted contracts generally used to impute the overall index investment. In particular they find a low, sometimes negative, correlation between index investment figures of the CFTC and those of the Masters algorithm. Instead, we find that the lowest correlation is 62% for the Kansas wheat contract and that all other coefficients are above 77%. Moreover, the index investment absolute values lie in a close range.

[16]Arbitrage capital approximated by spreading positions is only available for 19 contracts as the contracts traded on the London Metal Exchange (LME) and ICE-UK are not covered by the CFTC.

covered by the CFTC, for comparison.

Next, following the literature (see, *e.g.*, Stoll and Whaley, 2010 and Irwin and Sanders, 2012), we define the ratio of spreading positions over total open interest $ARB_{c,t}$, as $\frac{SPEC^{SPREAD}_{c,t}}{TOI_{c,t}}$, where the numerator is the total speculative spreading position. This is a proxy for the arbitrage capital deployed by speculators ought to ease the index investors' activity during the roll (or profit from it). We test **H1b** and check for a common break affecting the constituents of the SP-GSCI.

We look for a break in the commodity futures contracts open interest ratios, IND, $IND^{control}$, and ARB, using the Bai, Lumsdaine, and Stock (1998) algorithm. We estimate a $VAR(1)$, and restrict the break to apply to the intercept only.[17] The estimated equation is,

$$y_t = (G'_t \otimes G_t)\,\theta + d_t\,(k)\,(G'_t \otimes I_n)\,S'S\delta + \epsilon_t, \tag{1}$$

where y_t is a vector of IND, $IND^{control}$ or ARB series in turn, stacked for all available commodities. G'_t is a row vector containing y_{t-1} and a constant, n is the number of equations, and S is a selection matrix such that only the intercept is allowed to break with $S = s \otimes I_n$ and $s = (1, 0, \ldots, 0)$. We estimate the system of equations for every potential breaking date k, such that $d_t\,(k) = 0$ for $t \leq k$ and $d_t\,(k) = 1$ for $t > k$. We identify the break date as the value of k that generates a maximum Wald statistic higher than the limiting χ^2 distribution. Then, we construct the confidence interval (in days) of the estimated break date $\hat{k}$, for a given level of statistical significance.

[Insert Table 1 here]
[Insert Figure 1 here]

To test **H1c**, our identifying hypothesis of no overlap, we use the electronification dates provided by the exchanges in their press releases and we confirm it using dates identified in former research.[18,19] We report them in Table 1. Testing the overlap with the two aforementioned identified dates is therefore trivial.

3.2. The valuation effect of the roll

While the benchmarks in event studies based on stock returns are well established, there is no consensus on how to compute CAPCs in the existing literature. Mou (2011) does

[17] We choose the $VAR(1)$ specification based on the minimum AIC and BIC computed from one to 10 lags.

[18] See, *e.g.*, https://www.cmegroup.com/media-room/

[19] See Raman et al. (2017); Gousgounis and Onur (2018); Martinez et al. (2011).

not adjust the CATCs, while Henderson et al. (2015) use a linear factor model for price changes of single futures contracts. The period over which CAPCs could materialize is not well identified either. Depending on the scenario, CAPCs are supposed to occur before or during the roll.[20] Because every month, between 16 and 24 futures commodity contracts are rolled, the residual terms, *i.e.*, the error in the estimation window and the CAPCs during the event window, could be cross-correlated because of missing variables in the benchmark model. Ignoring this cross-correlation shrinks their standard errors, which in turns leads to an over-rejection of the null hypothesis. We first examine the valuation effect on the individual nearby and first deferred contracts. The reason is that the market reaction could be asymmetric.

To test our hypotheses **H2a, H2b and H2c**, we first conduct an event study. We download the daily closing prices of the 27 commodity futures constituents of the SP-GSCI for the first five consecutive maturities m. The sample starts on January 2, 1999 and ends on December 31, 2010. We compute the daily log futures price changes for every maturity available as $r^m_{c,t} = \log\left(F^m_{c,t}\right) - \log\left(F^m_{c,t-1}\right)$, accounting for the actual contract rollover. $F^m_{c,t}$ is the futures price of commodity c, on day t and for maturity m. We define the change on the term structure as $sr^m_{c,t} = r^{m+1}_{c,t} - r^m_{c,t}$, for every maturity of the term structure. Finally, we download and compute the Henderson Pearson Wang (see Henderson et al., 2015, hereafter, HPW) variables: the log changes on the MSCI emerging market index, SP-500 index, USD index, VIX, T-Bond, Baltic Dry Index, and inflation indices. We download the data from Thomson Reuters and the Commodity Research Bureau. First, we estimate the valuation effects of the roll and pre-roll on individual nearby ($r^1_{c,t}$) and first deferred ($r^2_{c,t}$) contracts in a single pass estimation with the following parametric specification,

$$r_{c,t} = \alpha_{0,c} + CAPC_{c,e}\Delta^{window}_{c,t} + \mathbf{a}^{\intercal}_c\mathbf{HPW}_t + \epsilon_{c,t}, \tag{2}$$

where c is the commodity contract, t is day-time indicator, $\Delta^{window}_{c,t}$ are individual commodity-event dummy vectors equal to 0.2 for every day of the roll or pre-roll, $CAPC_{c,e}$ captures the cumulative abnormal price changes for every c and event-month e.[21,22,23] For

[20]We verify that the roll (and pre-roll) period we use are followed well by the various ETFs in analyzing their prospectuses. Moreover, we conduct a test on the trading volume of a futures contract directly written on the SP-GSCI performance. We report these results in the Appendices E and F.

[21]The dummy vector is coded 0.2 during the days of the roll to capture the CAPCs using only one vector for the five days of the window. For any alternative window length, the value is $\frac{1}{\#\text{event days}}$.

[22]We conduct robustness tests discarding the January roll in all specifications, to control for the possible confounding effect of the yearly SP-GSCI reweighting which overlaps. For the contracts that are rolled in January the index unwinds 20% of its allocation from the nearby contract, while it adjusts the weights of the first deferred contracts to reflect the new allocation. For all other commodities, the reweighting is similar to that of a stock index. Our results are virtually the same.

[23]Our robustness tests also include an OLS estimation with the covariates used by Bakshi, Gao, and Rossi

the roll (pre-roll) valuation effect estimations, we discard the observations of the pre-roll (roll), as we expect them to be "abnormal".

Next, we estimate the valuation effects on the term structure during the roll and pre-roll with an alternative parametric benchmark. The counterfactual of the nearby term structure change ($sr^1_{c,t}$) is a linear function of the further deferred term structure (the price change differences between the third and the second deferred), which is not subject to roll effects (see Appendix B). We estimate the CATCs with a Seemingly Unrelated Regression (SUR; see Zellner, 1962), since the benchmarks are unique for each commodity, in a single pass estimation with the following specification,

$$sr^1_{c,t} = \alpha_{0,c} + CATC_{c,e}\Delta^{window}_{c,t} + \alpha_{1,c} sr^3_{c,t} + \epsilon_{c,t}, \tag{3}$$

We also control for benchmark misspecifications, and re-estimate CAPCs and CATCs with model-free benchmarks. We first use peer futures contracts written on fungible commodities that are not constituent of any index, to our knowledge. We download the first three maturities of these eighteen matching peers (see Table 1). The equation for the individual contracts becomes,

$$NPCAPC_{c,e} = \sum_{t=5}^{9} \left(r_{c,t} - r_{p(c),t} \right), \tag{4}$$

where $p(c)$ is the peer commodity contract, $NPCAPC_{c,e}$ captures the non-parametric cumulative abnormal price changes of commodity c and event-month e. For the term structure, the equation is,

$$NPCATC_{c,e} = \sum_{t=5}^{9} \left(sr^1_{c,t} - sr^1_{p(c),t} \right), \tag{5}$$

To control for any misspecification we also include a zero-benchmark (raw price changes). Finally, we adjust the standard errors for event-induced variance (Boehmer, Musumecci, and Poulsen, 1991) and cross-correlation (Kolari and Pynnonen, 2010).[24]

3.3. Testing the impact of the financialization

To further test **H2a, H2b and H2c** we use a difference in differences setting to see whether the financialization, *i.e.*, the post- relative to the pre-financialization period, con-

(2019), with no noticeable difference.

[24]Kolari and Pynnonen (2010) cross-correlation adjustment comes in complement of the Boehmer et al. (1991) correction. Moreover, it allows for adjustments in non-parametric cases, where the residuals of the estimation periods are the differences between the series of interest and the counterfactual.

ditional on the contract inclusion in a commodity index, had an impact on the panel of CAPCs. We estimate the following equation,

$$CAPC_{c,t} = \beta_1 DGSCI_{c,t} \times DFIN_t + \beta_2 DGSCI_{c,t} \times ELEC_{c,t} + \mathbf{b}^{\intercal}\mathbf{X}_{c,t} + \mu_c + \tau_t + \epsilon_{c,t}, \quad (6)$$

where $CAPC_{c,t}$ is a panel of nearby or first deferred contracts CAPCs in time t as of the last day of the event e estimated with the different benchmarks during the roll and pre-roll in turn, $DGSCI_{c,t}$ is a vector dummy coded "1" for the contracts constituents of the SP-GSCI and "0" otherwise. $DFIN_t$ is a dummy vector coded "1" if $t \geq$ Jan 2004 and "0" otherwise. $X_{c,t}$ is an optional matrix of stacked control variables vectors that are, (i) the absolute level of the term structure (basis) on the day preceding the event $B_{c,t-1}$, (ii) the cumulative price changes over the five days preceding the event $CR_{c,t-1}$ and (iii) the average realized volatility (squared price changes) computed over the past month $V_{c,t-1}$. μ_c and τ_t are contract and time fixed effects, respectively. If the financialization has any effect on the CAPCs, we expect the coefficient β_1 to significantly departs from zero. In an alternative setting, we add the interaction term between $ELEC_{c,t}$, the dummy vectors coded "1" from the inception date of the electronic contract, and $DGSCI_{c,t}$. This is to verify whether the electronification had an additional effect on the CAPCs.

3.4. Explaining individual abnormal price changes

We now turn to the determinants of the CAPCs, in the light of the recent advances of Kang et al. (forthcoming). We define two variables to proxy for insurance and liquidity premia. The net hedging pressure is a conventional factor for the explanation of the long-term insurance premium in asset-pricing studies. Instead, we choose a more appropriate, dynamic measure of net hedging pressure over the roll. We define the variable $\frac{\Delta HN_{c,t}}{OI^m_{c,t}}$ which is the variation of available net hedging pressure (short commercial minus long commercial category) between the first days available before and after the event, that are the two, enclosing, non-overlapping Tuesdays providing those are working days. For each nearby and first differed contract, we scale the net hedging pressure, which encompasses the whole term structure, by the corresponding individual open interest $OI^m_{c,t}$, to obtain a contract specific variable. Next, we define a measure of liquidity premium determinant $\frac{VCIT^m_{c,t}}{OI^m_{c,t}}$, as the total volume of index investment, similarly scaled by the individual level of open interest. This variable appears static but is in fact dynamic due to the index investment transfer from one contract to the next over this precise period. For an unambiguous interpretation of the variable, we sign it conditional on the amount traded: sold (negative, on the nearby) or bought (positive, on the first deferred).[25] We test **H3** with two panel data regressions for the nearby and first deferred contracts, defined as follows,

$$CAPC_{c,t} = \gamma_0 + \gamma_1 \frac{VCIT^m_{c,t}}{OI^m_{c,t}} + \gamma_2 \frac{\Delta HN_{c,t}}{OI^m_{c,t}} + \mathbf{g}^\intercal \mathbf{Y}_{c,t} + \mu_c + \tau_t + \epsilon_{c,t}, \tag{7}$$

In both settings, because $\frac{VCIT^m_{c,t}}{OI^m_{c,t}}$ is signed and negative (positive) on the nearby (first deferred), we expect γ_1 to be significantly positive if the liquidity premium explains the CAPCs during the roll. Conversely, as $\frac{\Delta HN_{c,t}}{OI^m_{c,t}}$ is a measure for the whole term structure, we expect that the more the net hedging pressure increases, the more important the normal backwardation, so that γ_2 would be negative.

3.5. Market efficiency

We assume the predators and arbitragers to be price takers. This is because the timing, in the context of the roll, is an essential feature for such agents. Price takers face both transaction costs and liquidity restrictions. Transaction costs can be decomposed in a fixed component made of the exchange commissions and operational costs and a variable component which is the bid-ask spread and which is the only exploitable varying term, both in time series and in cross section. We use the low frequency estimation of the bid-ask spread of Abdi and Ranaldo (2017) and restrict the minimum estimation to be at least of the size of the effective tick,

$$TC^m_{c,t} = \max\left(\sqrt{4\frac{1}{N}\sum_{t=1}^{N}\left(C^m_{c,t} - M^m_{c,t}\right)\left(C^m_{c,t} - M^m_{c,t+1}\right)}; [\text{effective tick}]^m_{c,t}\right), \tag{8}$$

where $C^m_{c,t}$ is the closing price for the contract c in t, $M^m_{c,t}$ is the difference between the "high" and "low" price. $[\text{effective tick}]^m_{c,t}$ is the theoretical minimum bid-ask spread computed with the minimum fluctuation (tick) of the contract in t and divided by the price of the contract in t. We compute this measure for the two maturities m, nearby and first deferred and sum them to obtain a transaction cost measure $TC_{c,t}$ incurred by the arbitrager. Next, we compute the illiquidity measure of Amihud (2002) defined as,

$$ILLIQ^m_{c,t} = \frac{|r^m_{c,t}|}{[\text{dollar trading volume}]^m}, \tag{9}$$

for the nearby and first deferred contracts. Finally, we use these measures to explain the

[25] As for the computation of IND, we also compute the liquidity premium determinant with the index investment restricted to the imputed SP-GSCI as well as with the iShares SP-GSCI AuM (available since 2006), with no sensible difference.

CATCs in the following panel regression,

$$CATC_{c,t} = \lambda_0 + \lambda_1 TC_{c,t} + \lambda_2 ILLIQ^1_{c,t} + \lambda_3 ILLIQ^2_{c,t} + \mu_c + \tau_t + \epsilon_{c,t} \qquad (10)$$

4. Empirical results

4.1. Summary statistics and break test

We report the results of the break tests in Table 2 for the 27 available IND series and for the 19 available $IND^{control}$ and ARB series of the SP-GSCI contracts.[26] We identify a significant break as of December 17, 2003, with a +/- 15 days confidence interval at 1%. Based on this test, the index investment break happens around the previous research consensus (late 2003 to early 2004). We also identify a break in ARB which occurs in January 2006. Using CFTC legacy report data of commercial long share of total OI, another proxy for index investment we compare the index investment break for the SP-GSCI contracts with the equivalent series for contracts with no index investment by definition. The break in $IND^{control}$ occurs by October 2003 for the SP-GSCI contracts, close to the one of IND, while we cannot identify any change in the market structure for the six non-indexed contracts.[27] Finally, all identified dates occur at least six months before the inception of the first electronically traded SP-GSCI contracts. This makes us confident that what we identify as financialization is not related to a market structure change due to innovations at the exchange level.

[Insert Table 2 here]

In Table 3, we report summary statistics of the log price changes of the nearby futures contracts of the SP-GSCI contracts for the entire sample period. We report the same statistics for the peer contracts written on similar underlying products.[28]

[Insert Table 3 here]

[26] We exclude the two ICE-UK contracts and the industrial metals traded on the LME from the break test on $IND^{control}$ and ARB, since they are not under the CFTC supervision and for which we have no data.

[27] The test is adjusted for the sample size.

[28] Note that some SP-GSCI contracts share the same peer because a close contract does not exist or is too illiquid.

4.2. Valuation effect

In Table 4, we report the results of the event study for the individual nearby and first deferred contracts, in Panels A and B, respectively, estimated with the HPW parametric benchmark and for the term structure, estimated with further to maturity term structure. We split the sample at the financialization break date and report estimations for the pre-, post-financialization and full sample period. We first notice that there is no systematic abnormal price changes on the individual contracts. For instance, the nearby contract price changes abnormally of close to 50 bps in the pre-financialization period and of -20 bps in the post-financialization, which is in line with the potential expected detrimental effect. However, the pattern for the first deferred contract is similar when the opposite reaction would confirm the financialization effects. Moreover, the t-statistics of all results (as high as 5.60 for the nearby contract) vanishes, as soon as we correct for event-induced variance and cross-correlation. The results of the Appendices G and H, which use peer contracts and a zero-benchmark as counterfactuals display the same pattern, which makes us confident that these results are not driven by benchmark misspecifications. We cannot make any inference either during the pre-roll, with results that display the same non-monotonic patterns and void significance, after we apply the event study corrections. Despite the fact that we target CAPCs to firstly test our hypotheses, we also report the equal- and SP-GSCI weighted CATCs in the Panels C and D of Table 4. The results are more consistent, with positive CATCs of 17 bps in all periods during the pre-roll. Interestingly, the CATCs estimated during the roll are decreasing from 15 bps in the pre-financialization period, to 3 bps in the post-financialization period. We interpret this decrease as a support for **H2c**, or sunshine trading. Our results, however, do not permit to clearly favor one hypothesis over another.

[**Insert Table 4 here**]

4.3. Impact of the financialization

The difference in differences model we use on both HPW factors and raw returns, with additional covariates, gives better definition of the effect of the financialization, conditional on the commodity index inclusion. Results are significantly negative for both contracts which cannot permit to favor one hypothesis over the other, as the effect of the financialization is asymmetric depending on the contract. On the one hand, the difference in differences operator is negative for the nearby contract during both roll and pre-roll, in support of the price pressure and predatory trading hypotheses, respectively. On the other hand, we find a negative coefficient as well (and even more significant) for the first deferred contract, which favors the alleviating effect of the financialization and thus, the sunshine trading hypothesis.

In addition to these contrasted results, and despite the large statistical significance of the difference in differences operator, the adjusted R^2 never reaches 1%, a further indication that most of the variance source is latent in our setting. Finally, we do not find any evidence that the staggered introduction of electronically traded contracts had any effect on the SP-GSCI contracts, with respect to the non-indexed contracts.

[Insert Table 5 here]

4.4. Insurance and liquidity premium

We report the results in Table 6. The coefficients of interest are of the expected signs, except for ΔHN in the two nearby contract settings with controls. The sign of the coefficient of $VCIT$ is always positive and significant at 10% in all but one test. We also attribute importance to the consistency of the signs over the tests. We confirm this robustness in unreported tests where the CAPCs are computed with additional benchmarks, and for individual tests in the pre- and post-financialization. Moreover, the fact that we do not find strong support for any of our hypotheses **H2a,b,c** regarding the CAPCs, and the inconsistent results we get when we compare the nearby and first-deferred contract, tend to confirm that we now capture an essential feature of the CAPCs determination during the roll. Hence, the CAPCs of the nearby contract depend on the amount of index-related investment unwound and therefore to the increase of liquidity premium on this leg. They also positively depend on the decrease of hedging pressure, and therefore on insurance premium reduction, over the event. We find consistent results in the first deferred contract, with coefficients of the expected signs. Moreover, the significance of the $\overline{HN}$ (the hedging pressure estimated in the long run) coefficient, one of our control variable, also provides support of the time horizon differences in the realization of (short term) liquidity and (long term) insurance (See, *e.g.*, Kang et al., forthcoming). Altogether, our results support **H3**, and in contrast with our previous findings, they indicate that CITs could modify the commodity market structure, in taking the role of traditional speculators, collecting insurance and generating liquidity premia.

[Insert Table 6 here]

4.5. Market efficiency

We report our results in Table 7. We estimate eq. 10 using CATCs as dependent variables. Indeed, we only look for market frictions effects, and hence expect the reaction to be symmetrical on both nearby and first deferred contract. Second, we expect a magnified effect when looking at both contracts simultaneously. Moreover, CATCs avoid the concurrent effect of fundamental shocks that may shift the whole term structure at once. In Panels C and D of 4, we report the CATCs that we use in our tests and their adjustest t-statistics. We find no significance in the coefficient of the liquidity measures which is, however, of the expected sign (the more liquidity, the smaller the CATCs). In contrast, the coefficient of our measure of transaction costs is significant at 1% for the pre-roll, roll and the two periods combined. In addition to the significance, the coefficient is of an order of magnitude close to unity (from 56 to 86%). We interpret this coefficient size as an additional indication that an essential feature of the CATCs are transaction costs. Indeed, we hypothesize in **H4** that CATCs are bounded by the size of the transaction costs and hence should be of similar magnitude. The extra component of the CATCs (that we miss, given that our coefficient is below unity), can therefore be the additional transaction costs such as exchange fees or market execution risks.

[**Insert Table 7 here**]

5. Conclusion

We study the consequences of the financialization of commodity futures markets. First, we identify a structural break in the share of index investment in total open interest common to the 27 constituents of the SP-GSCI. This break occurs on December 17, 2003. We also identify a break in the amount of overcoming arbitrage capital in January 2006. While we cannot find supportive results for any of the hypotheses underlying the individual price changes, we uncover, however, a decrease in CATCs, both in magnitude and statistical significance in the post-financialization period, which weakly support the sunshine trading hypothesis. Hence, the financialization might have eased the activity of index investors due to the concurrent public awareness of these predictable trades. More specific tests of the effect of financialization on the abnormal price changes of individual contracts do not permit to conclude either. The financialization appears detrimental to the nearby contract and favorable to the first deferred contract. Despite these inconsistent results, we are able to explain the abnormal price changes by two measures of liquidity and insurance premia. Moreover, we confirm a change in the commodity market structure in identifying that index investment related positions generate the liquidity premium, a function previously identified to be that of the traditional speculators (see Kang et al., forthcoming). We also question the small size of the CATCs, and relate them to the transaction costs bore by a price taker arbitrager. This further supports the sunshine trading hypothesis as arbitragers provide liquidity, and ease the roll activity of hedgers as soon as the price impact overcomes their transaction costs. Hence, what is perceived as risk-free and profitable arbitrage in a back-test actually becomes a non-profitable or zero-profit strategy. Finally, this study reconciles two contrasting findings of the literature on commodity index investment. On the one hand, we document a significant effect almost surely caused by CITs. On the other hand, the size of the effect is limited to the transaction costs, and it is very unlikely that index investors' position roll have modified the term structure as it has been previously been advocated.

References

Abdi, F., Ranaldo, A., 2017. A simple estimation of bid-ask spreads from daily close, high, and low prices. Review of Financial Studies 30, 4437–4480.

Acharya, V. V., Lochstoer, L. A., Ramadorai, T., 2013. Limits to arbitrage and hedging: Evidence from commodity markets. Journal of Financial Economics 109, 441–465.

Admati, A. R., Pfleiderer, P., 1991. Sunshine trading and financial market equilibrium. Review of Financial Studies 4, 443–481.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Aulerich, N. M., Irwin, S. H., Garcia, P., 2013. Bubbles, food prices, and speculation: Evidence from the cftc's daily large trader data files. Available at: http://dx.doi.org/10.3386/w19065 .

Bai, J., Lumsdaine, R. L., Stock, J. H., 1998. Testing for and dating common breaks in multivariate time series. Review of Economic Studies 65, 395–432.

Bakshi, G., Gao, X., Rossi, A. G., 2019. Understanding the sources of risk underlying the cross-section of commodity returns. Management Science 65, 619–641.

Bekaert, G., Harvey, C. R., Lumsdaine, R. L., 2002. Dating the integration of world equity markets. Journal of Financial Economics 65, 203–247.

Bessembinder, H., 1992. Systematic risk, hedging pressure and risk premiums in futures markets. Review of Financial Studies 5, 637–667.

Bessembinder, H., Carrion, A., Tuttle, L., Venkataram, K., 2016. Liquidity, resiliency and market quality around predictable trades: Theory and evidence. Journal of Financial Economics 121, 142–166.

Bodie, Z., Rosansky, V. I., 1980. Risk and return in commodity futures. Financial Analysts Journal 36 (3), 33–39.

Boehmer, E., Musumecci, J., Poulsen, A. B., 1991. Event-study methodology under conditions of event-induced variance. Journal of Financial Economics 30, 253–272.

Boons, M., Roon *de*, F., Szymanowska, M., 2014. The price of commodity risk in stock and futures markets. Available at: http://dx.doi.org/10.2139/ssrn.1785728 .

Brunetti, C., Reiffen, D., 2014. Commodity index trading and hedging costs. Journal of Financial Markets 21, 153–180.

Brunnermeier, M. K., Pedersen, L. H., 2005. Predatory trading. Journal of Finance 60, 1825–1863.

Cheng, I.-H., Kirilenko, A., Xiong, W., 2015. Convective risk flows in commodity futures markets. Review of Finance 19, 1733–1781.

Cheng, I.-H., Xiong, W., 2014. The financialization of commodity markets. Annual Review of Financial Economics 6, 419–441.

Deaton, A., Laroque, G., 1992. On the behaviour of commodity prices. Review of Economic Studies 59, 1–23.

Erb, C. B., Harvey, C. R., 2006. The strategic and tactical value of commodity futures. Financial Analysts Journal 62 (2), 69–97.

Gorton, G., Hayashi, F., Rouwenhorst, K. G., 2012. The fundamentals of commodity futures returns. Review of Finance 17, 35–105.

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

Gousgounis, E., Onur, E., 2018. The effect of pit closure on futures trading. Journal of Commodity Markets 10, 69–90.

Gray, R. W., 1961. The search for a risk premium. Journal of Political Economy 69, 250–260.

Greer, R. J., 2000. The nature of commodity index returns. Journal of Alternative Investments 3 (1), 45–52.

Grossman, S. J., Miller, M. H., 1988. Liquidity and market structure. Journal of Finance 43, 617–633.

Hamilton, J. D., Wu, J. C., 2015. Effects of index-funds investing on commodity futures prices. International Economic Review 56, 187–205.

Henderson, B. J., Pearson, N. D., Wang, L., 2015. New evidence on the financialization of commodity markets. Review of Financial Studies 28, 1285–1311.

Hicks, J. R., 1946. In: *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*, Clarendon Press, Oxford.

Hirshleifer, D., 1988. Residual risk, trading costs, and commodity futures risk premia. Review of Financial Studies 1, 173–193.

Irwin, S. H., Sanders, D. R., 2012. Testing the masters hypothesis in commodity futures markets. Energy Economics 34, 256–269.

Kang, W., Rouwenhorst, K. G., Tang, K., forthcoming. A tale of two premiums: The role of hedgers and speculators in commodity futures markets. Journal of Finance .

Keynes, J. M., 1930. In: MacMillian (ed.), *A Treatise on Money: The Applied Theory of Money*, MacMillian, London, pp. 3–46.

Kolari, J. W., Pynnonen, S., 2010. Event study testing with cross-sectional correlation of abnormal returns. Review of Financial Studies 23, 3996–4025.

Martinez, V., Gupta, P., Tse, Y., Kittiakarasakun, J., 2011. Electronic versus open outcry trading in agricultural commodities futures markets. Review of Financial Economics 20, 28–36.

Masters, M. W., 2008. Testimony before the Committee on Homeland Security and Governmental Affairs United States senate, May 20. Available at: `https://www.hsgac.senate.gov/download/052008masters` .

Masters, M. W., White, A. K., 2008. The accidental hunt brothers, how institutional investors are driving up food and energy prices. Available at: `https://www.loe.org/images/content/080919/Act1.pdf` .

Mou, Y., 2011. Limits to arbitrage and commodity index investment: Front-running the goldman roll. Available at: `http://dx.doi.org/10.2139/ssrn.1716841` .

Raman, V., Robe, M. K., Yadav, P. K., 2017. The third dimension of financialization: Electronification, intraday institutional trading, and commodity market quality. Available at: `http://dx.doi.org/10.2139/ssrn.3068248` pp. –.

Rezitis, A. N., 2015. Empirical analysis of agricultural commodity prices, crude oil prices and us dollar exchange rates using panel data econometric methods. International Journal of Energy Economics and Policy 5, 851–868.

Roon *de*, F. A., Nijman, T. E., Veld, C., 2000. Hedging pressure effects in futures markets. Journal of Finance 55, 1437–1456.

Sanders, D. R., Irwin, S. H., 2013. Measuring index investment in commodity futures markets. Energy Journal 34, 105–127.

Schwartz, E. S., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. Journal of Finance 52, 923–973.

Stoll, H. R., Whaley, R. E., 2010. Commodity index investing and commodity futures prices. Journal of Applied Finance 1, 7–46.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association 57, 348–368.

Table 1: **Contracts description**

In Panel A, we report the specifications of the 27 SP-GSCI contracts as well as their peer futures contracts (non-indexed and with similar underlying commodities). The specifications include their trading venues, ticker, underlying commodities, unit and currencies if applicable. We also report their maturity months with the appropriate code letter. Finally, we report the inception date of the corresponding electronically traded contract. In Panel B, we report the same information for the six commodity contracts not part of the SP-GSCI but covered by the CFTC.

Panel A: SP-GSCI contracts

SP-GSCI contracts						Peer contracts					
Ticker	Trading venue	Underlying	Unit	Maturity	Electronification	Ticker	Trading venue	Underlying	Unit	Maturity	Currency
C	CBT	Corn	bu (5,000)	HKNUZ	01/08/2006	EMA	MATIF (EURONEXT)	Corn	MT (50)	FHMQX	EUR
CC	ICE-US	Cocoa	MT (10)	HKNUZ	02/02/2007	LCC	ICE-EUROPE	Cocoa	MT (10)	HKNUZ	GBP
CL	NYMEX/ICE	WTI	bbl (1,000)	FGHJKMNQUVXZ	05/09/2006	JCO	TOCOM	Crude oil	KL (50)	FGHJKMNQUVXZ	JPY
CT	ICE-US	Cotton	lbs (50,000)	HKNVZ	02/02/2007	CCF	ZCE	Cotton#1	MT (5)	FHKNUX	RMB
FC	CME	Feeder cattle	lbs (50,000)	FHJKQUVX	24/11/2008	BGI	BM&F Bovespa	Beef	net arrobas. (330)	FGHJKMNQUVXZ	BRL
GC	CMX	Gold	ozt (100)	GJMQVZ	04/12/2006	JAU	TOCOM	Gold standard	Kg (1)	GJMQVZ	JPY
HO	NYMEX	Heating Oil	gal (42,000)	FGHJKMNQUVXZ	05/09/2006	JKE	TOCOM	Kerosene	KL (50)	FGHJKMNQUVXZ	JPY
KC	ICE-US	Coffee	lbs (37,500)	HKNUZ	02/02/2007	0JAC	TGE (TOCOM)	Coffee	MT (3.45)	FHKNUX	JPY
KW	KBT	Kansas wheat	bu (5,000)	HKNUZ	13/01/2008	1MWE	MGE	Wheat	bu (5000)	HKNUZ	USD
LC	CME	Live cattle	lbs (40,000)	GJMQVZ	04/12/2006	BGI	BM&F Bovespa	Beef	net arrobas. (330)	FGHJKMNQUVXZ	BRL
LCO	ICE-UK	Brent	bbl (1,000)	FGHJKMNQUVXZ	12/06/2006	JCO	TOCOM	Crude oil	KL (50)	FGHJKMNQUVXZ	JPY
LGO	ICE-UK	Gasoil	MT (100)	FGHJKMNQUVXZ	12/06/2006	JGO	TOCOM	Gasoil	KL (50)	FGHJKMNQUVXZ	JPY
LH	CME	Lean hogs	lbs (40,000)	GJMNQVZ	04/12/2006	BGI	BM&F Bovespa	Beef	net arrobas. (330)	FGHJKMNQUVXZ	BRL
MAL	LME	Aluminum	MT (25)	FGHJKMNQUVXZ	04/12/2006	0JAL 1	TOCOM	Aluminium	MT (5)	GJMQVZ	JPY
MCU	LME	Copper	MT (25)	FGHJKMNQUVXZ	04/12/2006	HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ	USD
MNI	LME	Nickel	MT (6)	FGHJKMNQUVXZ	na	HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ	USD
MPB	LME	Lead	MT (25)	FGHJKMNQUVXZ	na	HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ	USD
MSN	LME	Tin	MT (5)	FGHJKMNQUVXZ	na	HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ	USD
MZN	LME	Zinc	MT (25)	FGHJKMNQUVXZ	na	HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ	USD
NG	NYMEX/ICE	Natural gas	MMBtu (10,000)	FGHJKMNQUVXZ	05/09/2006	NGLNM	ICE-UK	Natural gas	therms/day (1000)	FGHJKMNQUVXZ	GBP
OJ	ICE	Orange juice	lbs (15,000)	FHKNUX	02/02/2007	LSU	LIFFE	White sugar	MT (50)	HKQVZ	USD
PL	NYMEX	Platinum	ozt (50)	FGHJKMNQUVXZ	04/12/2006	JPL	TOCOM	Platinum	g (500)	GJMQVZ	JPY
RB	NYMEX	RBOB gasoline	gal (42,000)	FGHJKMNQUVXZ	05/09/2006	JGL	TOCOM	Gasoline	KL (50)	FGHJKMNQUVXZ	JPY
S	CBT	Soybeans	bu (5,000)	FHKNQUX	01/08/2006	JAS	TGE (TOCOM)	Soybeans	MT (25)	FGHJKMNQUVXZ	JPY
SB	ICE-US	Sugar#11	lbs (112,000)	HKNVZ	02/02/2007	LSU	LIFFE	White sugar	MT (50)	HKQVZ	USD
SI	CMX	Silver	ozt (5,000)	FHKNUZ	04/12/2006	JSV	TOCOM	Silver	Kg (1)	GJMQVZ	JPY
W	CBT	Chicago wheat	bu (5,000)	HKNUZ	01/08/2006	1MWE	MGE	Wheat	bu (5,000)	HKNUZ	USD

Panel B: Non-indexed contracts

1CB	CME	Butter	lbs (20,000)	FGHJKMNQUVXZ	17/09/2007
HG	COMEX	Copper	lbs (25,000)	FGHJKMNQUVXZ	04/12/2006
LB	CME	Lumber	m^3 (110,000)	FHKNUX	20/10/2008
NF	CME	Non fat dry milk	lbs (44,000)	FGHJKMNQUVXZ	17/09/2007
O	CBOT	Oats	bu (5,000)	HKNUZ	01/08/2006
PA	NYMEX	Palladium	ozt (100)	HMUZ	04/12/2006
RR	CBOT	Rough Rice	CWT (2,000)	FHKNUX	01/08/2006

Maturity month code: F = January, G = February, H = Mars, J = April, K = May, M = June, N = July, Q = August, U = September, V = October, X = November, Z = December.

Table 2: **Break detection in commodity index investment and arbitrage capital**

The table reports the results of the break detection algorithm, for multivariate stationary time series, of Bai et al. (1998). The algorithm loops the following estimation, $y_t = (G'_t \otimes G_t)\,\theta + d_t\,(k)\,(G'_t \otimes I_n)\,S'S\delta + \epsilon_t$, over the k possible break dates, trimmed to begin (end) at the first (last) decile of the sample. the model is a $VAR\,(1)$, a lag we choose according to AIC and BIC on uni- and multivariate autoregressions. We restrict the model to test for a break on the intercept only. We report the highest F-statistic ($max_k F\,(k)$), of the generated distribution, its significance according to the Bekaert, Harvey, and Lumsdaine (2002) critical values, the corresponding date, the dimension of the test (equal to the amount of series in the $VAR\,(1)$ case),the confidence intervals (CI) expressed in days at 10%, 5% and 1%, and the magnitude of the intercept shift. In Panel A, we report the results of the test on the SP-GSCI contracts total index investment: (i) OI of CITs imputed with the Masters (2008) algorithm, scaled by TOI, denoted IND and (ii) commercial long OI from the CFTC legacy report scaled by TOI, denoted $IND^{control}$. For the arbitrage capital ARB, we use the spreading positions of speculators, scaled by the total OI. In Panel B, we report the results for $IND^{control}$ of non-SP-GSCI contracts covered by the CFTC (butter, lumber, non-fat dry milk, oats, palladium and rough rice), excluding the copper COMEX contract, which is part of the BCOM index. The frequency is quarterly and the period of study is from January 1995 to December 2015.

Variable	# Contracts	Break date	CI (10%)	CI (5%)	(1%)	Intercept shift %	F-statistic	Significance
				Panel A: SP-GSCI contracts				
IND	21	17-Dec-03	$+/-9$	$+/-11$	$+/-15$	1.80	259.14	$***$
$IND^{control}$	19	15-Oct-03	$+/-30$	$+/-36$	$+/-54$	3.40	141.62	$***$
ARB	19	04-Jan-06	$+/-48$	$+/-64$	$+/-104$	5.28	58.45	$***$
				Panel B: Non-Indexed contracts				
$IND^{control}$	6	02-Feb-06	$+/-904$	$+/-1080$	$+/-1418$	7.21	18.78	

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3: **Summary statistics of the nearby contracts of the SP-GSCI and their peers**

We report the annualized mean, standard deviation, skewness, excess kurtosis of the daily log price changes of SP-GSCI futures and their matching peer contracts by underlying commodity. We also report the proportion of days during which the corresponding contract was in contango. The period is from January 1999 to December 2010.

SP-GSCI Ticker	Mean %	σ %	Skewness	Kurtosis	Contango %	Mean %	σ %	Skewness	Kurtosis	Contango %
Panel A: SP-GSCI and peer contracts										
	SP-GSCI contracts					Peer contracts				
C	−5.48	25.83	0.19	0.33	86.75	1.45	17.00	−0.39	0.68	11.21
CC	3.08	30.42	−0.78	0.30	76.30	−1.01	27.42	−0.38	0.37	20.97
CL	17.98	34.72	−1.10	0.22	63.57	9.64	18.11	−1.06	1.24	31.77
CT	−8.97	27.51	−0.72	0.30	74.49	1.55	7.22	1.25	3.09	13.72
FC	7.28	10.58	−0.65	0.47	45.39	5.08	16.21	−1.14	0.83	26.51
GC	11.52	16.25	−1.35	0.44	85.87	10.20	17.87	−1.42	0.50	35.69
HO	17.47	35.19	−0.08	0.08	71.86	16.56	23.78	−0.99	0.38	28.65
KC	−6.72	31.81	−0.34	0.40	85.73	−24.54	40.05	−1.06	0.51	13.49
KW	1.86	25.86	0.18	0.30	74.80	8.39	24.01	0.96	0.49	35.37
LC	8.51	13.37	−0.47	0.19	56.00	5.08	16.21	−1.14	0.83	26.51
LCO	19.47	33.01	−0.97	0.21	51.38	9.64	18.11	−1.06	1.24	31.77
LGO	18.73	31.30	−0.45	0.13	59.60	6.49	8.49	4.45	4.07	16.61
LH	−1.31	22.72	−1.06	0.29	57.75	5.08	16.21	−1.14	0.83	26.51
MAL	2.97	19.78	−0.14	0.71	83.62	1.22	21.08	−1.70	1.12	22.02
MCU	17.09	25.06	0.08	1.05	55.48	20.38	26.99	−0.65	0.55	33.74
MNI	13.47	33.51	0.02	0.62	75.92	20.38	26.99	−0.65	0.55	33.74
MPB	14.27	30.31	−1.61	0.82	70.82	20.38	26.99	−0.65	0.55	33.74
MSN	15.32	23.11	−1.58	1.19	61.28	20.38	26.99	−0.65	0.55	33.74
MZN	8.15	27.27	−0.26	1.21	80.57	20.38	26.99	−0.65	0.55	33.74
NG	−25.84	50.84	0.13	0.23	80.24	−52.65	41.88	0.31	0.85	30.97
OJ	−13.17	36.38	0.23	1.01	67.13	20.17	23.71	0.02	0.24	61.40
PL	18.46	19.38	−0.84	1.27	48.15	15.97	22.92	−2.55	0.81	58.57
RB	2.22	23.42	−2.48	1.67	41.68	11.90	23.24	−1.20	0.41	38.36
S	10.69	24.00	−1.58	0.47	65.12	3.84	29.13	−0.81	0.57	48.08
SB	12.30	33.26	−1.02	0.28	55.41	20.17	23.71	0.02	0.24	61.40
SI	17.54	26.49	−2.79	0.76	89.23	9.69	25.18	−3.03	0.87	41.66
W	−5.10	28.92	0.75	0.23	86.26	8.39	24.01	0.96	0.49	35.37
	Panel B: Non-SP-GSCI contracts (with CFTC reports)									
1CB	-4.61	4.35	-0.70	5.02	64.94					
HG	24.93	26.39	0.23	0.58	60.09					
LB	-16.08	27.28	0.89	0.16	67.39					
NF	-4.30	3.53	-14.88	10.86	56.63					
O	15.49	27.43	0.71	0.62	70.20					
PA	5.56	7.83	6.48	9.44	56.96					
RR	-0.47	24.55	0.38	0.22	86.90					

Table 4: **Cumulative abnormal changes of the nearby and first deferred prices, and of the term structure of SP-GSCI contracts**

In Panel A, we report the results of a parametric event study based on the HPW covariates of (Henderson et al., 2015) eq. 2, for the price changes of the nearby and first deferred contracts. In Panel B and C, we report the results of an event study based on the further to maturity term structure. (see Appendix B and eq. 3). Because the covariates are contract specific, we use a seemingly unrelated regression (SUR), for which the convergence of the feasible generalized least square is achieved over seven iterations. To capture the per-series average cumulative abnormal price changes, we use a dummy coded 0.2 during the five days of the window of interest for each month in which a roll occurs. We report the average cumulative abnormal changes for the five days, the unadjusted t-statistics as well as their adjustments: for event-induced variance (Boehmer et al., 1991, BMP) and abnormal return cross-correlation (Kolari and Pynnonen, 2010, KP).

	pre-roll			roll		
period	1999–2003	2004–2010	1999-2010	1999–2003	2004–2010	1999–2010
			Panel A: nearby contract			
CAPC (bps)	11.16	20.73	7.76	48.80	−20.57	8.90
unadj. t-stat.	1.12	0.55	0.97	3.98	−1.95	1.10
BMP	0.93	0.54	0.92	3.28	−1.85	0.88
KP	0.37	0.20	0.34	1.13	−0.66	0.34
			Panel B: first deferred contract			
CAPC (bps)	28.84	38.07	24.82	64.08	−17.37	16.94
unadj. t-stat.	2.63	3.09	2.86	5.60	−1.72	2.17
BMP	2.42	2.59	2.56	3.87	−1.60	1.64
KP	0.80	0.93	0.98	1.30	−0.58	0.62
			Panel C: term structure			
CATC (bps)	17.68	17.34	17.05	15.29	3.20	8.04
unadj. t-stat.	4.61	2.84	3.77	3.20	0.86	2.84
BMP	2.84	1.75	2.87	3.16	0.59	2.41
KP	1.10	1.00	1.62	1.05	0.22	0.96
			Panel D: term structure (GSCI weighted)			
CATC (bps)	22.94	18.47	18.90	20.37	3.57	9.47
unadj. t-stat.	4.56	2.90	3.82	3.41	0.96	2.92
BMP	2.88	1.75	2.96	3.31	0.66	2.50
KP	1.11	1.02	1.66	1.14	0.25	1.00

Table 5: **Difference in differences: effect of the financialization on CAPCs around the roll**

We report the results of the estimation of eq. 6, a difference in differences estimations with the interaction of the $DGSCI_{c,t}$ dummy variable, set to "1" when the commodity is part of the SP-GSCI and to "0" otherwise, and the $DFIN_t$ dummy, set to "1" after December 17, 2003 and to "0" before. We correct the standard errors with Newey-West method, using four lags of auto-correlation. When we estimate the model on raw returns, we include control variables that are (i) $B_{c,t-1}$ the basis computed as of the first available day preceding the event window, (ii) $CR_{c,t-1}$ the cumulated returns over the five days preceding the event window and (iii) $V_{c,t-1}$ the realized volatility (sum of squared returns) computed over the preceding month.

	nearby				first deferred			
Variables	HPW		raw		HPW		raw	
	Panel A: $CAPC_{c,t}$ (roll)							
$DGSCI_{c,t} \times DFIN_t$	-48.05^{*}	-58.43^{*}	-59.07^{*}	-62.77^{*}	-54.02^{**}	-43.43	-59.07^{**}	-45.23
	(-1.80)	(-1.76)	(-1.93)	(-1.68)	(-2.03)	(-1.34)	(-1.96)	(-1.23)
$DGSCI_{c,t} \times ELEC_{c,t}$		10.67		5.34		-20.19		-23.60
		(0.30)		(0.14)		(-0.58)		(-0.65)
$B_{c,t-1}$			22.20	16.25			20.91	3.82
			(0.08)	(0.06)			(0.08)	(0.01)
$CR_{c,t-1}$			78.70	72.02			62.14	9.75
			(0.33)	(0.30)			(0.37)	(0.05)
$V_{c,t-1}$			153.38	152.59			144.58	147.54
			(0.60)	(0.60)			(0.68)	(0.69)
$Adj.R^2\%$	0.47	0.49	0.54	0.57	0.47	0.49	0.51	0.53
#Obs: 3852 - #Contracts: 34								
	Panel B: $CAPC_{c,t}$ (pre-roll)							
$DGSCI_{c,t} \times DFIN_t$	-65.99^{**}	-56.72	-77.37^{**}	-70.63^{*}	-51.77^{*}	-65.55^{**}	-72.85^{**}	-77.60^{**}
	(-2.13)	(-1.54)	(-2.46)	(-1.72)	(-1.65)	(-1.96)	(-2.40)	(-2.09)
$DGSCI_{c,t} \times ELEC_{c,t}$		-8.23		-9.35		13.87		7.46
		(-0.20)		(-0.22)		(0.35)		(0.18)
$B_{c,t-1}$			4.80	15.29			-342.78	-337.72
			(0.01)	(0.04)			(-1.32)	(-1.30)
$CR_{c,t-1}$			-80.00	-70.65			-92.16	-82.99
			(-0.35)	(-0.31)			(-0.59)	(-0.42)
$V_{c,t-1}$			-284.59	-283.86			-25.39	-26.04
			(-1.16)	(-1.16)			(-0.10)	(-0.10)
$Adj.R^2\%$	0.49	0.50	0.52	0.67	0.50	0.60	0.53	0.72
#Obs: 3852 - #Contracts: 34								

*p<0.1; **p<0.05; ***p<0.01
Newey-West standard errors (four lags)

Table 6: **Insurance and liquidity premium**

We report the results of eq. 7. $VCIT_{c,t}$ is the ratio of the OI tracking total index investment, spread over the SP-GSCI commodity futures as per the SP-GSCI historical allocation, over the total OI of each contract. $\Delta HN_{c,t}$ is the share of the changes in hedging pressure, that we compute between the first available days before and after the event window of interest, in to the total OI. The control variables are (i) $\overline{HN}_{c,t-1}$, the average past hedging pressure in absolute value computed over the past 52 weeks, (ii) $B_{c,t-1}$ the basis (log term structure) computed as of the first available day preceding the event window, and (iii) $V_{c,t-1}$, the realized volatility (sum of squared returns) computed over the preceding month. For a better lecture of the relative magnitude of the coefficients, we scale the measures individually.

Variables	$CAPC_{c,t}$ (nearby)				$CAPC_{c,t}$ (first deferred)			
$VCIT_{c,t}(\%)$	2.34*	2.04**	1.35	3.74	2.36*	2.16*	2.67*	1.96
	(1.80)	(2.03)	(0.80)	(1.05)	(1.73)	(1.66)	(1.84)	(0.88)
$\Delta HN_{c,t}(\%)$	−4.96	−3.45	1.47	1.40	−0.69	−0.99	−2.51	−3.02
	(−1.68)	(1.50)	(0.39)	(0.55)	(−0.87)	(−0.88)	(−0.98)	(−1.03)
$\overline{HN}_{c,t-1}(\%)$			−7.46	−7.77*			3.39	3.98*
			(−1.50)	(−1.70)			(1.05)	(1.73)
$B_{c,t-1}(\%)$			−1.39	−1.47			1.22	1.28
			(−0.49)	(−0.26)			(0.70)	(0.59)
$V_{c,t-1}(\%)$			−1.5	−1.27			1.54	1.68
			(−0.15)	(−0.27)			(0.27)	(0.48)
FE		y		y		y		y
$Adj.R^2\%$	0.94	1.02	0.70	0.83	0.99	1.12	0.80	1.01
#Obs: 2377 - #Contracts: 19								

*p<0.1; **p<0.05; ***p<0.01

Newey-West standard errors (four lags) and standard errors clustered at the contract level

Table 7: **Market efficiency**

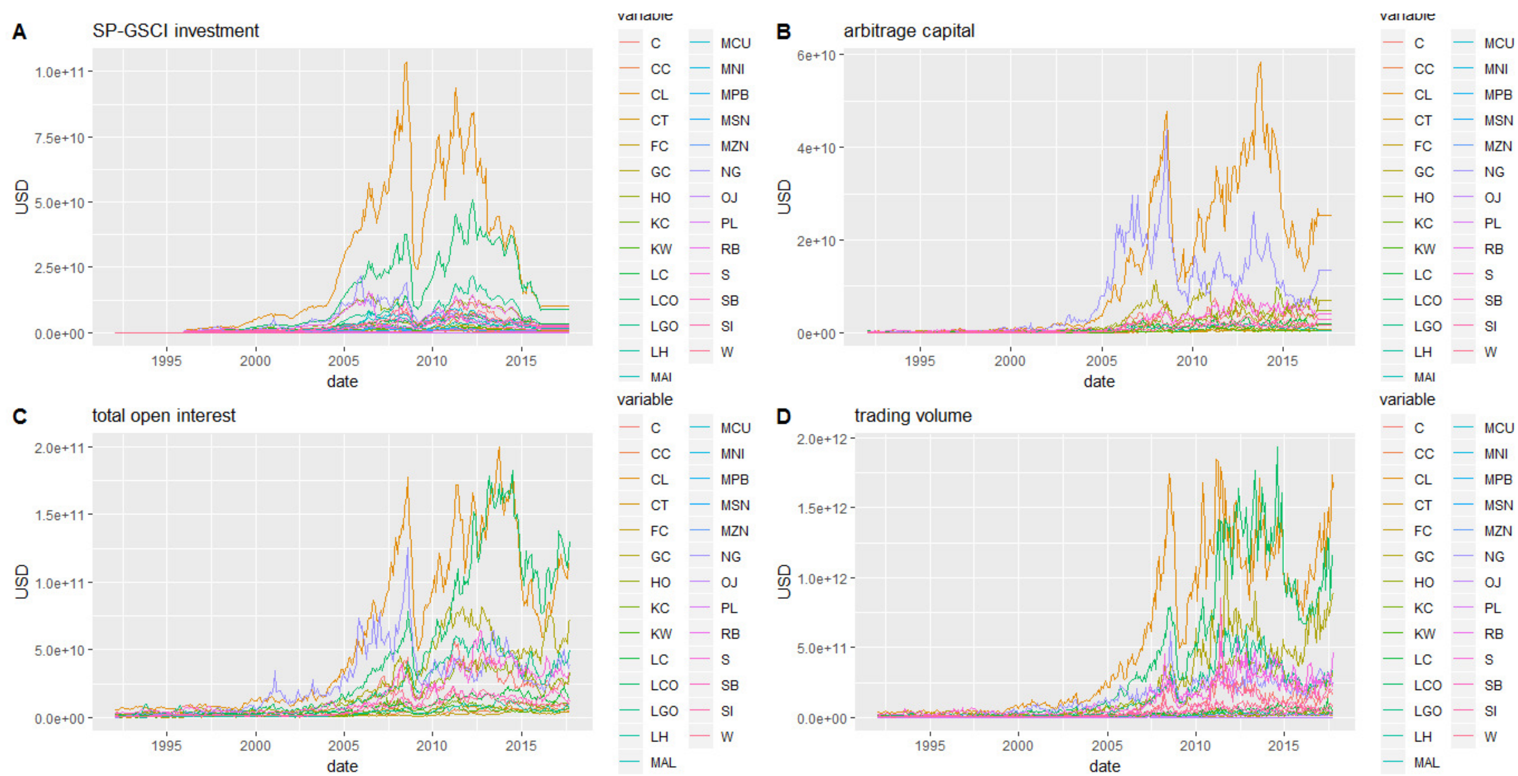
We report the results of the panel regression of eq. 10, where we explain the CATCs with the sum of the bid-ask spreads for the nearby and first deferred contracts that we estimate using the modified measure of Abdi and Ranaldo (2017) defined in eq. 8. Hence the variable TC approximates the transaction costs incurred by a price-taker, arbitrager. We also add the measure of illiquidity of Amihud (2002) that we compute for both nearby and first deferred contracts in eq. 9.

	$CATC_{c,t}$		
	pre-roll	roll	pre-roll and roll
$TC_{c,t}$	0.60^{***}	0.56^{***}	0.86^{***}
	(3.27)	(3.87)	(3.70)
$ILLIQ^1_{c,t} \times 10^4$	-3.04	-2.20	-1.51
	(-0.11)	(-0.49)	(-0.38)
$ILLIQ^2_{c,t} \times 10^4$	-4.68	-2.36	-2.35
	(-0.14)	(-0.67)	(-0.68)
$Adj.R^2\%$	9.24	8.32	8.56
#Obs: 2661 - #Contracts: 27			

*p<0.1; **p<0.05; ***p<0.01
Newey-West standard errors (four lags)
Standard errors clustered at the contract level

Fig. 1: **SP-GSCI futures contracts investment, arbitrage capital, open interest and trading volume**

In this figure we plot the monthly series of index investment IND (Panel A), total open interest (Panel C) and trading volume (Panel D) for the 27 futures contracts constituents of the SP-GSCI. In Panel B, we plot the spreading positions of speculators, proxy for arbitrage capital ARB, for the 21 contracts covered by the CFTC (excluding the metal contracts of the LME that are combined with the horizontal axis).



Appendices

A. Masters (2008) procedure

- We compute the value of each futures contract at each date, in multiplying the underlying deliverable quantity (available from the contract specifications) by the contract price.

- Masters and White (2008) report the total, yearly $VCIT$ for the period 1995-2006 for all CITs with the following breakdown, SP-GSCI, BCOM and "other funds". We reallocate these values in each futures contracts based on the weights of the SP-GSCI.

- The Masters and White (2008) data are sampled annually. We assume a constant growth rate and convert it monthly. This leaves the overall impact of index investment unchanged but allows for a more precise allocation given the frequencies of our SP-GSCI (monthly) and BCOM (daily) weights.

- We assume zero index investment before 1995 as Masters and White (2008) data already indicate that the index investment before 1998 is negligible.

- From 2006 onwards, the CFTC publishes weekly the CIT report that precisely allocates the $VCIT$ of all index investors for 13 agriculture contracts. We convert the reported $VCIT$ in USD.

- To extrapolate the total index investment, We select the Cocoa contract among the 13 commodities of the CIT report. Indeed, the Cocoa contract is indexed only by the SP-GSCI and the BCOM (and no other major index funds to our knowledge). Our results are unchanged if another unique SP-GSCI/BCOM contract is chosen such as Coffee or Cotton).

- We divide the selected series of $VCIT$ by the contemporaneous percent weight of the contract in the indices to obtain 100% of the index investment tracking these indices.

B. Further deferred contracts benchmark

We start from the conventional commodity futures value equation in logs,

$$f_{c,t}^{m} = s_{c,t} + \left(r + w_c - y_{c,t}^{m}\right)\left(T^{m} - t\right), \tag{11}$$

and assume that r, the per-period common interest rates and w_c the individual storage costs are constant over time and over maturity, whereas $y_{c,t}^{m}$, the convenience yield is maturity and time varying (see, *e.g.*, Schwartz, 1997). T^{m} stands for the stopping time, of maturity m. If we assume the equation to be valid along the term structure and taking the usual nearby contract approximation for spot price, we write any contract of the term structure in term of another one as follows,

$$f_{c,t}^{m+k} = f_{c,t}^{m} + \left(r + w_c\right)k - \sum_{i=m+1}^{m+k} y_{c,t}^{i}, \forall k \geq 0, \tag{12}$$

For the spreading position benchmark, we relate the difference of the nearby and first deferred prices $f_{c,t}^{2} - f_{c,t}^{1}$ with the second and third deferred prices, $f_{c,t}^{4} - f_{c,t}^{3}$, such that, $f_{c,t}^{2} - f_{c,t}^{1} = f_{c,t}^{4} - f_{c,t}^{3} + y_{c,t}^{4} - y_{c,t}^{2}$. Hence the log term structure change is,

$$\left(f_{c,t+1}^{2} - f_{c,t}^{2}\right) - \left(f_{c,t+1}^{1} - f_{c,t}^{1}\right) = \left(f_{c,t+1}^{4} - f_{c,t}^{4}\right) - \left(f_{c,t+1}^{3} - f_{c,t}^{3}\right) + \left(y_{c,t+1}^{4} - y_{c,t}^{4}\right) - \left(y_{c,t+1}^{2} - y_{c,t}^{2}\right),$$

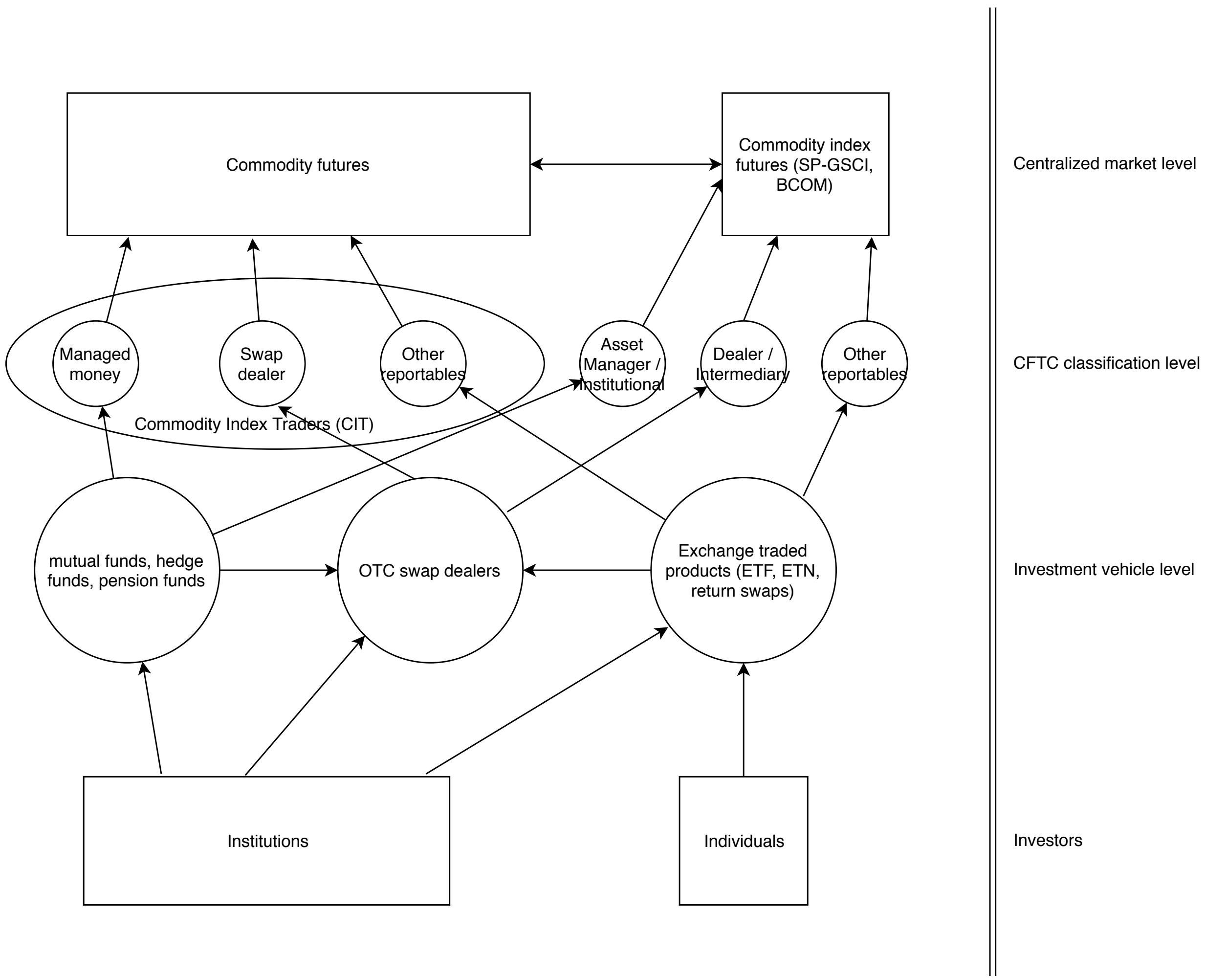
or simplified is,

$$sr_{c,t+1}^{1} = sr_{c,t+1}^{3} + \left(y_{c,t+1}^{4} - y_{c,t}^{4}\right) - \left(y_{c,t+1}^{2} - y_{c,t}^{2}\right) \tag{13}$$

We estimate the relationship between $sr_{c,t+1}^{1}$ and $sr_{c,t+1}^{3}$ with a regression whose intercept is composed of the convenience yield terms on the right hand side of the equation. Such regression predicted values are a parametric benchmark free of any first generation fund activity (SP-GSCI and BCOM tracking investment is absent of further deferred contracts by definition). The model for the benchmark is, $sr_{c,t}^{1} = \alpha_{0,c} + CATC_{c,e}\Delta_{c,t}^{window} + \alpha_{1,c} sr_{c,t}^{3} + \epsilon_{c,t}$, as in eq.(3).

C. Index investment functional chart

In this functional chart, we describe the various vehicles that commodity index investment may use, and how these vehicles hedge on the commodity futures markets, through a portfolio of futures or directly with a futures contract written on the performance of this portfolio. We add the CFTC classification of the categories involved.



D. Variable definition

We define all variables used throughout the article, with their symbols, plain definition, how we compute them (if applicable), their units, frequency, period of availability and additional descriptions if needed.

Symbol	Title	Computation	Unit	Frequency	Period	Description
$F^m_{c,t}$	Futures price	–	USD	D	1999–2010	m: maturity, c: commodity, t: time
$f^m_{c,t}$	Log futures price	–	–	D	1999–2010	–
$R^m_{c,t}$	Arithmetic return	$R^m_{c,t} = \frac{F^m_{c,t}}{F^m_{c,t-1}} - 1$	%	D	1999–2010	We use "return" for simplicity but it is a misnomer, "percent price change" is more appropriate.
$r^m_{c,t}$	Log price change	$r^m_{c,t} = f^m_{c,t} - f^m_{c,t-1}$	%	D	1999–2010	–
$SR^m_{c,t}$	Arithmetic spreading return	$SR^m_{c,t} = R^{m+1}_{c,t} - R^m_{c,t}$	%	D	1999–2010	–
$sr^m_{c,t}$	Log term structure change	$sr^m_{c,t} = r^{m+1}_{c,t} - r^m_{c,t}$	%	D	1999–2010	–
$CATC^m_{c,e}$	Cumulative abnormal term structure change	$CASR^m_{c,e} = \sum_{t=1}^{t+\tau} \left(SR^m_{c,t} - Benchmark_{c,t}\right)$	%	M	1999–2010	e: event month, τ: number of event days
$B^m_{c,t}$	Basis	$B^m_{c,t} = \left(\frac{F^{m+1}_{c,t}}{F^m_{c,t}} - 1\right) \times \frac{1}{\#\text{interval days}}$	%	D	1999–2010	We use "Basis" for simplicity but it is a misnomer as we use the nearby contract to proxy for the spot price. "Maturity ratio" is more appropriate.
$b^m_{c,t}$	Log basis	$b^m_{c,t} = \left(f^{m+1}_{c,t} - f^m_{c,t}\right) \times \frac{1}{\#\text{interval days}}$	%	D	1999–2010	–
$V^m_{c,t}$	Volume	–	# contracts	D	1999–2010	–
CQ_c	Deliverable quantity	–	–	–	1999–2010	–
$CVAL^m_{c,t}$	Contract value	$CVAL^m_{c,t} = CQ_c \times F^m_{c,t}$	USD	D	1999–2010	–
$XVAL^m_{c,t}$	Contract value	$XVAL^m_{c,t} = CQ_c \times X_{c,t}$	USD	D	–	For any value $X_{c,t}$ denominated in # of contracts, there is a $XVAL^m_{c,t}$ denominated in USD.
$TO^m_{c,t}$	Turnover	$TO^m_{c,t} = CVAL^m_{c,t} \times V^m_{c,t}$	USD	D	1999–2010	–
$OI^m_{c,t}$	Open interest	–	# contracts	D	1999–2010	Total # of contracts opened in the exchange for each maturity.
$TOI_{c,t}$	Total open interest	–	# contracts	D	1999–2010	Total # of contracts opened in the exchange for all maturities.
$COM^P_{c,t}$	Commercial position	–	# contracts	W	1999–2010	Total # of contracts held by the "commercial" category, for all maturities. P stands for the position either long (L), short (S) or spreading ($SPREAD$).
$SPEC^P_{c,t}$	Speculative position	–	# contracts	W	1999–2010	Total # of contracts held by the speculative category for all maturities.
$ARB_{c,t}$	Arbitrage capital ratio	$ARB_{c,t} = \frac{SPEC^{SPREAD}_{c,t}}{TOI_{c,t}}$	$\% \in [0,1]$	W	1999–2010	Total net spreading position (always positive) scaled by total open interest.
$TVCIT_t$	Total SP-GSCI type investment	see Appendix A	USD	M	1999–2010	–
$WGSCI_{c,t}$	Weights of SP-GSCI	–	$\% \in [0,1]$	M	1999–2010	–
$VCIT_{c,t}$	Amount tracking SP-GSCI type funds	$VCIT_{c,t} = WGSCI_{c,t} \times TVCIT_t$	USD	D	1999–2010	–
$IND_{c,t}$ (1)	Index investment ratio	$IND_{c,t} = \frac{VCIT_{c,t}}{TOI_{c,t}}$	$\% \in [0,1]$	W	1999–2010	–
$IND^{control}_{c,t}$ (2)	Index investment ratio (alternative)	$IND^{control}_{c,t} = \frac{COM^L_{c,t}}{TOI_{c,t}}$	$\% \in [0,1]$	W	1999–2010	Alternative using total commercial long position from the legacy report, as this report includes swap dealers in this category.
$HP_{c,t}$	Net hedging pressure	$HP_{c,t} = \frac{COM^S_{c,t} - COM^L_{c,t}}{TOI_{c,t}}$	$\% \in [-1,1]$	W	1999–2010	–
$\Delta HP_{c,t}$	Change in net hedging pressure	$\Delta HP_{c,t} = HP_{c,e+1} - HP_{c,e-1}$	$\% \in [-1,1]$	W	1999–2010	–

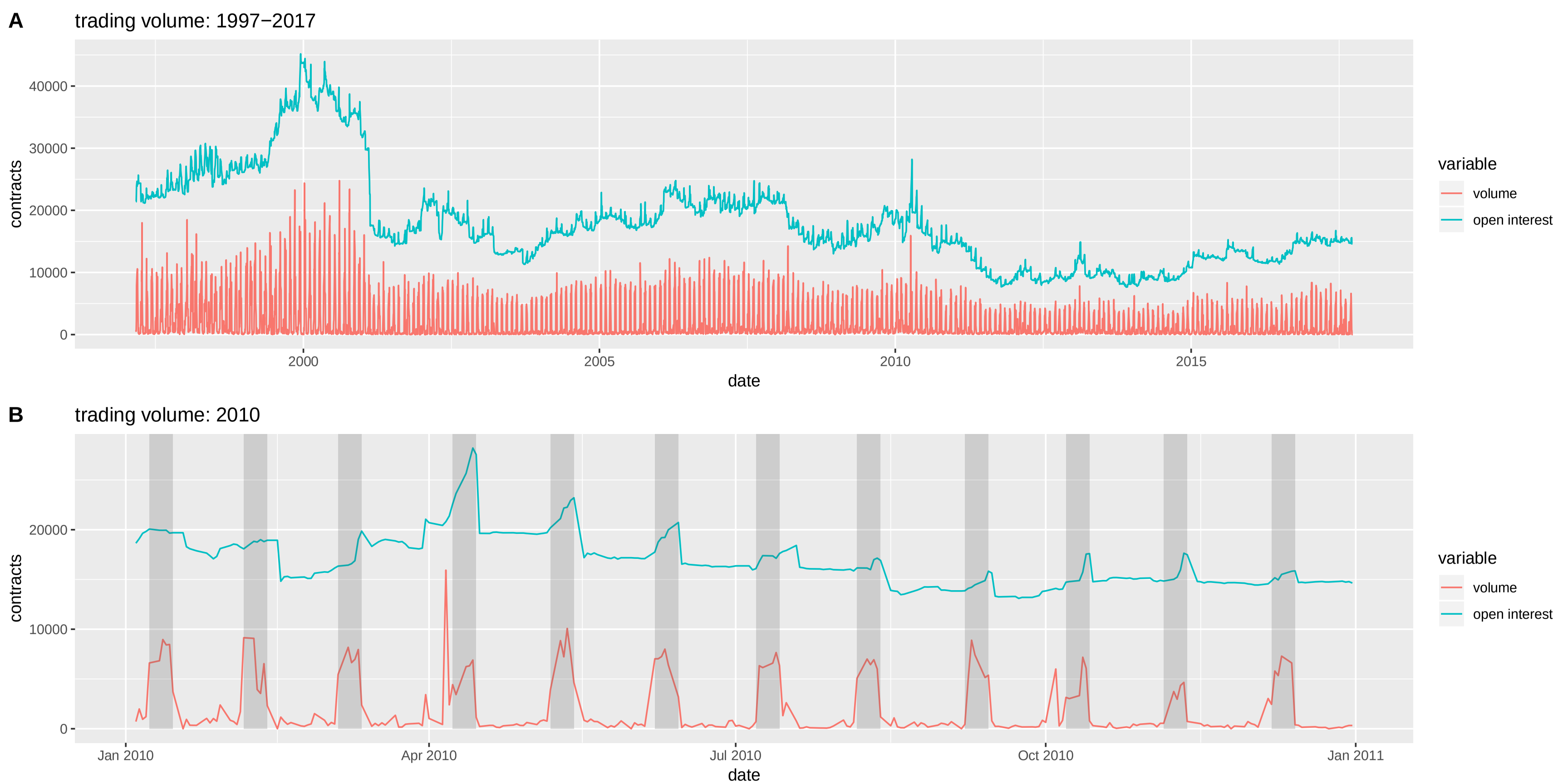
E. SP-GSCI futures roll-period validity (1)

One key issue of this study is to differentiate the positions taken by the various class of investors, their timing and on which contract they are held. To overcome this identification problem, the literature adopts sometimes the usage of the Large Trader Reporting System (LTRS), an internal database of the CFTC with a higher frequency (daily) and splits positions of each investor by maturity. We propose an alternative identification test with the study of a futures contract traded on the CME, that is directly written on the SP-GSCI performance, which was launched by the CME in January 1994 (see CME specifications of the SP-GSCI futures contract). Because this contract has no use for genuine commercial traders, it is likely that its daily OI and trading volume will indicate when SP-GSCI trackers, predators, or arbitragers take most of their positions. In this table, we display the results of a regression of the total trading volume of the nearby and first deferred contracts written on the SP-GSCI on a dummy coded "1" during the roll and "0" otherwise. We report the t-statistics in parenthesis. The t-statistic of 69.30 and the mean difference of 4812 contracts indicate that traders are active almost only during the roll.

Variables	SP-GSCI daily trading volume
intercept	695.08
	(20.58)
roll dummy	4812.34
	(69.30)
R^2 %	44.68
# observations	5968

F. SP-GSCI futures roll-period validity (2)

In Panel A, we plot the daily total trading volume of the future contract written on the SP-GSCI for the period 1997–2017. In Panel B, we focus on the year 2010, to highlight the concentration of trading activity during the five days of the official SP-GSCI roll, we shade the corresponding surfaces.



G. CAPC and CATC: peers benchmark

Peers benchmark: To capture the per-series average cumulative abnormal price changes, we use a dummy coded 0.2 during the five days of the window of interest for each month in which a roll occurs. We report the average cumulative abnormal changes for the five days, the unadjusted t-statistics as well as their adjustments: for event-induced variance (Boehmer et al., 1991, BMP) and abnormal return cross-correlation (Kolari and Pynnonen, 2010, KP).

	pre-roll			roll		
period	1999–2003	2004–2010	1999-2010	1999–2003	2004–2010	1999–2010
	Panel A: nearby contract					
CAPC (bps)	13.06	5.76	8.55	41.75	17.91	26.06
Unadj. t-stat.	1.07	0.55	1.10	3.32	1.56	3.11
BMP	1.00	0.54	0.89	2.88	1.55	2.73
KP	0.40	0.20	0.38	1.00	0.67	1.21
	Panel B: first deferred contract					
CAPC (bps)	31.58	18.83	23.29	50.95	27.15	34.62
Unadj. t-stat.	2.89	1.76	2.56	4.55	2.58	4.55
BMP	2.37	1.38	2.50	3.52	2.40	3.70
KP	0.89	0.65	1.01	1.26	0.88	1.51
	Panel C: term structure					
CATC (bps)	18.52	13.08	14.75	9.20	9.24	8.56
Unadj. t-stat.	3.74	1.31	2.30	1.53	2.05	2.16
BMP	2.68	1.23	2.28	1.07	1.72	2.08
KP	1.01	0.58	1.10	0.36	0.87	0.95
	Panel D: term structure (GSCI weighted)					
CATC (bps)	23.63	13.91	16.53	11.50	10.48	9.81
Unadj. t-stat.	3.73	1.30	2.36	1.54	2.14	2.20
BMP	2.68	1.23	2.34	1.14	1.77	2.17
KP	1.01	0.58	1.11	0.39	0.91	0.99

H. CAPC and CATC: Zero benchmark

Zero benchmark: To capture the per-series average cumulative abnormal price changes, we use a dummy coded 0.2 during the five days of the window of interest for each month in which a roll occurs. We report the average cumulative abnormal changes for the five days, the unadjusted t-statistics as well as their adjustments: for event-induced variance (Boehmer et al., 1991, BMP) and abnormal return cross-correlation (Kolari and Pynnonen, 2010, KP).

	pre-roll			roll		
period	1999–2003	2004–2010	1999-2010	1999–2003	2004–2010	1999–2010
	Panel A: nearby contract					
CAPC (bps)	29.04	25.87	27.17	65.71	−14.55	17.47
Unadj. t-stat.	3.09	2.07	3.25	5.31	−1.37	2.10
BMP	2.42	2.01	3.07	3.70	−1.27	1.72
KP	1.09	0.52	1.04	1.26	−0.47	0.64
	Panel B: first deferred contract					
CAPC (bps)	41.18	44.12	42.58	71.22	−9.43	22.63
Unadj. t-stat.	3.88	3.15	4.79	6.15	−1.01	2.82
BMP	3.76	2.83	4.24	3.78	−0.84	2.28
KP	1.32	0.81	1.49	1.24	−0.35	0.84
	Panel C: term structure					
CATC (bps)	12.14	18.25	15.40	5.50	5.12	5.16
Unadj. t-stat.	3.18	2.36	3.12	1.25	1.64	2.02
BMP	2.04	1.85	2.61	1.15	1.48	1.87
KP	0.77	0.83	1.44	0.41	0.59	0.80
	Panel D: term structure (GSCI weighted)					
CATC (bps)	15.80	19.51	17.09	7.63	5.73	6.13
Unadj. t-stat.	3.16	2.40	3.20	1.40	1.67	2.14
BMP	2.08	1.86	2.69	1.25	1.48	1.93
KP	0.79	0.84	1.49	0.47	0.60	0.85