Limitations of Granger Causality Analysis to assess the price effects from the financialization of agricultural commodity markets under bounded rationality

Stephanie-Carolin Grosche†

Abstract
Modern agricultural commodity markets are simultaneously governed by a physical and a financial market element. Whether financial “index trading” activity influences price levels on the futures markets has been investigated by empirical studies using Granger Causality Analysis. A critical review of these studies reveals inconclusive results. Based on sensitivities of the method, reasons for limited interpretability of results may be omitted determinants of financial trading activity, failure to consider the informational efficiency of markets, time-varying and feedback effects of boundedly rational heterogeneous trading strategies and limitations in specifying adequate theoretical variables from existing data.

Keywords: Granger Causality, Commodity speculation, Financialization of commodity markets, Agricultural commodity markets

1 Introduction
Today’s agricultural commodity markets appear to be simultaneously governed by a physical and a financial market element. The physical market element comprises fundamental factors of supply and demand as well as the activities of “traditional” market participants. The latter are predominantly producers and other actors along the supply chain of the physical commodity and its derived products. They are buying and selling the physical commodity on the spot and hedging their input and output prices on the derivative markets. The financial market element is the result of the perception of agricultural commodities as a financial asset class, an emergence of financial traders on the market and a growth in derivative markets (cf. REDrado et al. 2009; Erb and Campbell 2006; Gorten and Rouwenhorst 2006; Miffre and Rallis 2007). As opposed to “traditional”

† University of Bonn, Institute for Food and Resource Economics, Bonn, Germany, stephanie.grosche@ilr.uni-bonn.de. The author thanks Thomas Heckelei, University of Bonn for various helpful comments and suggestions.
market participants, financial investors have no interest in the physical underlying commodity but merely a financial interest in gaining exposure to its value and price changes. Consequently, their market activity concentrates on the derivative markets, such as exchange-traded futures and options or the non-exchange traded over-the-counter (OTC) markets.

While the physical market element has at all times been acknowledged as a market governing factor, interest in the influence of the financial element has first emerged at a significant scale as the result of the 2007-2008 food price crisis. A perceived correlation of price movements with position holdings of financial investors sparked interest in potential underlying causalities. A particular point of concern was “index trading” of financial investors. According to the Commodity Futures Trading Commission (CFTC) (n.d.a; n.d.c) definitions, index traders are “[…] replicating a commodity index by establishing long futures positions in the component markets and then rolling those positions forward from future to future using a fixed methodology”. Index trading is“[…] a passive strategy designed to gain exposure to commodity price movements as part of a portfolio diversification strategy”. According to the CFTC (2008, p. 17), an “index fund” is an investment product that has the obligation to replicate the performance of a specific index or sub-index, most commonly either via direct future-based replication or synthetic swap-based replication. In the latter case, the index fund enters into a swap agreement with a financial intermediary from which the fund receives the index return in exchange for some asset basket (RAMASWAMY 2011, p. 5). Commodity index funds are typically either mutual funds, Exchange Traded Funds (ETF) or Exchange Traded Products (ETP).

A range of empirical studies have attempted to find evidence for a cause-and-effect relation running from financial index trading to price levels observed on the agricultural commodity futures markets. These studies predominantly rely on the method of Granger-Causality analysis (GCA). The objective of this paper is to conduct a critical review of their analysis, to interpret and synthesize results and discuss potential limitations. In doing so, this paper contributes to existing research by explicitly considering findings from financial market theory to increase understanding of the emerging financial market element and to augment the theory underlying the null hypothesis. Thereby, the focus will be on implications of

\[ \text{\footnotesize 2 Common commodity indices are the Thompson Reuters Jefferies Commodity Research Bureau (TRJ-CRB) Index, or the Standard & Poors Goldman Sachs Commodity Index (S&P GSCI).} \]

\[ \text{\footnotesize 3 Following the definitions used in BLACKROCK (2011), the term ETF is referring only to structures similar to an index-type mutual fund. ETPs include all products that share similarities with ETFs but are debt securities.} \]
the informational efficiency of markets and effects of bounded rationality of traders.

The subsequent section presents a short overview of the methodology of GCA as employed by the empirical studies and describes the data sources used in their analysis. Section 3 reviews the models and findings in the empirical studies. Section 4 presents necessary theoretical extensions based on sensitivities of the GCA methodology. Section 5 concludes the analysis and gives an outlook on promising areas of future research.

2 Methodology and data

The concept of Granger Causality (GC) has been frequently applied since the seminal paper by Granger (1969). The operational definition of GC states that a variable $X$ will cause another variable $Y$ if the probability of correctly forecasting $Y_{t+1}$, with $t = 1, \ldots, T$, will increase by including information about $X_t$ in addition to other information contained in a specific information set at time $t$ ($\Omega_t$). Specification of the information set requires measurement of variables from real world data, which renders data availability a crucial ingredient for the correct application of GCA.

2.1 GCA methodology and tests

Underlying the definition of GC are three axioms (Granger 1980, pp.330, 335-336, 338): Axiom I: An event can only be the cause of another event if it precedes it in time, a future event can thus never be the cause for an event in the past; Axiom II: There should not be any redundancy in $\Omega_t$; Axiom III: All causal relationships remain constant in their direction over time. Only the strengths of the relations or the time lags may change but never the general direction.

More formally, $X_t$ is not a prima facie cause of $Y_{t+1}$ if:

\[ F(Y_{t+1} \mid \Omega_t) = F(Y_{t+1} \mid \Omega'_t) \]

where $F(.)$ is the conditional distribution function of $Y_{t+1}$, $\Omega_t$ is an information set containing a series $Z_t$ and $\Omega'_t$ is an extended information set containing $X_t$ in addition to $Z_t$.

If a point forecast of $Y_{t+1}$ is used instead of the whole distribution then $X_t$ is not a prima facie cause in mean for $Y_{t+1}$ if:

\[ E[Y_{t+1} \mid \Omega_t] = E[Y_{t+1} \mid \Omega'_t] \]

and

\[ MSE(E[Y_{t+1} \mid \Omega_t]) = MSE(E[Y_{t+1} \mid \Omega'_t]) \]
where $E[.]$ denotes the conditional expectation of the point forecast of $Y_{t+1}$ given $\Omega_t$ and $\Omega_t'$, respectively and MSE is the mean squared error of the forecast of $Y_{t+1}$ used to measure quality of the prediction (e.g. GRANGER 1980, p. 337; HAMILTON 1994, p. 303; LÜTKEPOHL 2007, p. 42). If GC runs in both directions, there is feedback between the two variables $X$ and $Y$ (GRANGER 1969, p. 428) and if $X_{t+1}$ helps forecast $Y_{t+1}$ then *prima facie instantaneous causality or instantaneous feedback* is said to occur (GRANGER 1980, p. 340; HIEMSTRA AND JONES 1994, p. 1644; LÜTKEPOHL 2007, p. 47).

Statistical tests for non causality test the null hypothesis that a variable $X$ does not Granger-cause another variable $Y$. In the bivariate linear case, a typical way to test for GC is in the context of bivariate Vector Autoregressive (VAR) models of the form:

$$Y_t = c + \alpha_i Y_{t-i} + \ldots + \alpha_i' Y_{t-i} + \beta_j Y_{t-j} + \ldots + \beta_j' Y_{t-j} + \varepsilon_t, \quad t = 1, 2, \ldots, T$$

where $i$ and $j$ denote the lag-length, $c$ is any constant, $\alpha$ and $\beta$ are regression coefficients, and $\varepsilon_t$ is a white noise process (cf. HAMILTON 1994, pp. 48–49). In bivariate linear models with stationary variables Wald tests are most widely employed. For the case of integrated or cointegrated variables, Toda and Yamamoto (1995) introduce a more robust modified Wald-test with a lag-augmented VAR. In addition, GC tests for multivariate or nonlinear models exist. Lütkepohl (2007) illustrates possible approaches to test for multi-step causality in multivariate models and Dufour et al. (2006) develop linear methods for short-run and long-run non-causality testing within multivariate VAR models. Yamamoto and Kurozumi (2006) show an approach to test for long-run GC in multivariate models with cointegrated variables. A bivariate nonlinear GC test was developed in Baek and Brock (1992), applied with some modifications by Hiemstra and Jones (1994) and further discussed in Diks and Panchenko (2005).

### 2.2 Data sources

Data availability is a key requirement for the specification of theory-consistent proxy variables. In order to test for GC between financial index trading activity and price level effects, proxy variables for both need to be defined.

#### 2.2.1 Price data

Historical price data for exchange-traded futures contracts can for example be obtained from the relevant exchanges. In the case of U.S. traded contracts, the most important exchanges are the Chicago Board of Trade (CBOT), the Chicago Mercantile Exchange (CME) the Kansas City Board of Trade (KBOT) or the Intercontinental Exchange (ICE). The major agricultural commodity futures traded
there are wheat (W-), corn (C-), soybeans (S-), rough rice (RR), soybean oil (BO), soybean meal (SM), live cattle (LC), feeder cattle (FC), lean hogs (LH), cocoa (CC), coffee (KC), sugar no. 11 (SB) and cotton (CT). Price data for customizable OTC products such as forwards, swaps and options is more difficult to obtain as these markets generally suffer from a lack of transparency.

The future price (FP) for a given contract specification depends on the expiration date of the contract. The price for the nearby contract (FP0) is the price for the contract closest to expiration. Deferred contracts have expiration dates further into the future. For example, in November, the C- December contract is the nearby contract, the March contract the first deferred with price FP1, the May contract the second deferred contract with price FP2, etc.

2.2.2 Index trading data

The exclusive source for detailed trading data on U.S. agricultural commodity futures markets split by trader categories are the Commitments of Traders Reports (COT), the Disaggregated Commitments of Trader Reports (DCOT) and the Commodity Index Trader (CIT) Reports, all published by the CFTC. No comparable data is available for non U.S. Dollar denominated futures contracts or for the OTC markets.

The COT Report is published every Friday at 3:30 p.m. Eastern Standard Time and gives a snapshot of traders’ long position open interest (LPOI), short position open interest (SPOI) in futures and in combined futures and options contracts for major agricultural commodities at market close of the previous Tuesday of the week. The difference between LPOI and SPOI for a single trader category is the net open interest (OI). The trader categories used to disaggregate the OI were before 4 September 2009 limited to “commercial” and “noncommercial” reporting traders and nonreporting traders. Commercial traders are “[…] commercially engaged in business activities hedged by the use of the futures or option markets”, which includes “traditional” market participants with business activities linked to physical commodities as well as financial traders who use the markets for financial hedging, (CFTC n.d.a). Noncommercial traders are all other reporting traders.

In order to increase market transparency and better reflect trading motives, the DCOT Reports have been compiled starting 1 December 2009 (CFTC 2008). Backdated historical data exists until 13 June 2006. DCOT Reports use four reportable trader categories: “Producer/Merchant/Processor/User”, “Swap Dealers” (SWAP), “Managed Money” and “Other Reportables” (CFTC n.d.b). The Produc-

Details on these reports are available directly from CFTC (n.d.b; n.d.a). Stoll and Whaley (2010) and Sanders et al. (2009) also give detailed summaries.
ers, Merchants and Processors are engaged in a business activity related to the physical commodity and are using the futures markets to hedge their commercial risk. Traders in the SWAP category are mostly financial institutions that hedge the risk stemming from OTC swap transactions with their clients. Managed Money refers to traders who trade on behalf of clients such as registered Commodity Trading Advisors (CFA) or Commodity Pool Operators (CPO)\(^5\) or unregistered actively managed funds, including hedge funds or pension funds (c.f. Stoll and Whaley 2010, p. 9). “Other Reportables” includes all traders that report but do not match one of the above categories.

The CFTC publishes the CIT Report on index trading as a supplement to the (D)COT Reports. It is a combined report for futures and options positions, for which backdated historical data exists until 2006. Reporting traders in the CIT Report are split into “commercial”, “noncommercial” and “index traders (INDEX)” and the CIT Report lists the associated LPOI, SPOI and spreading positions.

3 Review of empirical studies

Selection criteria for the sample of empirical studies are a focus on the price level effects from financial index trading, the application of GCA, and the inclusion of the 2007-2008 price spike. Figure 1 shows the empirical studies included in the sample along with the covered time period and temporal data aggregation. The focus markets are grains and oilseeds (W- (CBOT, KCBT), C-, S-, RR, BO, SM), livestock (LC, FC, LH) and soft commodities (CC, KC, SB, CT). Robles et al. (2009) only use W-, C-, S- but include RR. Sanders and Irwin (2011) only analyze W-, C-, S-. Gilbert (2010) takes an index approach and includes all commodities in the IMF food price index.

Since the weekly CIT Reports are only available starting 2006, they constrain the time period of observation and feasible temporal disaggregation levels for studies based on their data. The exception is Sanders and Irwin (2011) who use non-public weekly data on INDEX positions for the years 2004-2005 on selected markets and Aulerich et al. (2010) who also use non-public daily data from the CFTC starting Jan 2004 and are able to split the sample period in two sub-periods (2004-2005 and 2006-2008).

\(^5\) CFAs give investment advice in commodity futures, which can include exercising the trading and managing the accounts of their customers; CPOs manage funds that are pooled from investors for futures and/or options trading. Both CFAs and CPOs have to register with the U.S. National Futures Association (NFA) (n.d.a, n.d.b).
The studies mostly focus on direct price level effects from futures trading on the futures market. The exceptions are ROBLES et al. (2009) and GILBERT (2010) who use FAO spot price data or the IMF food price index respectively. Nevertheless, with a no-arbitrage assumption leading to a spot-future parity at expiration of the futures contract, these spot prices could still be interpreted as the prices on future markets for an expiring contract (cf. HULL 2006, p. 26; GEMAN 2005, p. 37).

### 3.1 General model

The studies within the sample use bivariate linear frameworks with autoregressive specifications. For price level effects of financial index trading for a given commodity ($Com$), the price level ($Level$) at time $t$ is explained by its own lagged values and the lagged values of a variable representing index trading activity ($Activity$). A generic version of the models used in the empirical studies is given by the following equation.

$$Level_{t,Com} = c_1 + \alpha_1 Level_{t-i,Com} + \ldots + \alpha_i Level_{t-i,j,Com} + \beta_1 Activity_{t-i,j,Com} + \ldots + \beta_j Activity_{t-i,j,Com} + \varepsilon_t$$

for $t=1, \ldots, T$

where $c_1$ is a constant, $\alpha$ and $\beta$ are the coefficients, $i$ and $j$ denote the lag length and $\varepsilon_t$ is a white noise process. The null hypothesis of no causality is $H_0: \beta_1 = \beta_2 = \ldots = \beta_j = 0$ that is tested with standard F-tests.
Models that directly forecast volatility based on own lagged values and an *Activity* variable as presented in AULERICH et al. (2010) and IRWIN and SANDERS (2010a, 2010b) are omitted from this analysis due to the specific issues related to the measurement and forecasting of volatility, which cannot be covered within the scope of this paper (see e.g. GRANGER 2002 for a discussion).

### 3.2 Individual specifications

Apart from their choice of time period and the chosen temporal data aggregation, the sampled studies most notably differ in their exact specification of the *Level* and *Activity* variables included in their models. The *Level* variables are calculated based on absolute FP levels at time t, the changes in the FP between two time periods, which is the return at time t (R<sub>t</sub>), or as spreads between FP0 and FP1. Table 1 summarizes the different specifications.

**Table 1: Variable specification for Level**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Specification</th>
<th>Data source</th>
<th>Description</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln R&lt;sub&gt;t&lt;/sub&gt;</td>
<td>ln(&lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt;/&lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;) = ln &lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt; - ln &lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>Relevant exchange</td>
<td>Log-relative futures price changes of nearby futures contracts from one observation to the next, AULERICH et al. (2010) also use first-deferred contracts</td>
<td>AULERICH et al. (2010)/SANDERS and IRWIN (2011)</td>
</tr>
<tr>
<td>R&lt;sub&gt;t&lt;/sub&gt;</td>
<td>&lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt; - &lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>Relevant exchange</td>
<td>Difference in futures prices between one observation and the next</td>
<td>IRWIN and SANDERS (2010a, 2010b)</td>
</tr>
<tr>
<td>∆R&lt;sub&gt;t&lt;/sub&gt;</td>
<td>&lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt; - &lt;span class=&quot;lato&quot;&gt;FP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;</td>
<td>Relevant exchange</td>
<td>Price spread between first deferred and nearby contract</td>
<td>STOLL and WHALEY (2010)</td>
</tr>
<tr>
<td>LogSP&lt;sub&gt;t&lt;/sub&gt;</td>
<td>ln &lt;span class=&quot;lato&quot;&gt;SP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt;&lt;sup&gt;DEF&lt;/sup&gt; - &lt;span class=&quot;lato&quot;&gt;SP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;&lt;sup&gt;DEF&lt;/sup&gt;</td>
<td>FAO</td>
<td>Logarithm of spot prices</td>
<td>ROBLES et al. (2009)</td>
</tr>
<tr>
<td>∆LogSP&lt;sub&gt;t&lt;/sub&gt;</td>
<td>ln &lt;span class=&quot;lato&quot;&gt;SP&lt;/span&gt;&lt;sub&gt;t&lt;/sub&gt;&lt;sup&gt;DEF&lt;/sup&gt; - &lt;span class=&quot;lato&quot;&gt;SP&lt;/span&gt;&lt;sub&gt;t-1&lt;/sub&gt;&lt;sup&gt;DEF&lt;/sup&gt;</td>
<td>IMF</td>
<td>Difference in logarithm of IMF food price index</td>
<td>GILBERT (2010)</td>
</tr>
</tbody>
</table>

*Activity* variables are either specified as position holdings (*Pos*) of the traders, in absolute terms, relative to overall holdings in the market or to holdings of other traders, or as changes in position holdings (*Flow*). Table 2 summarizes the specifications used in the respective studies.
### Table 2: Variable specification for Activity

<table>
<thead>
<tr>
<th>Variables</th>
<th>Specification</th>
<th>Description</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta LPOI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>( LPOI_{t} - LPOI_{t-1} )</td>
<td>Change in LPOI of index traders, measured in contracts</td>
<td>STOLL and WHALEY (2010)</td>
</tr>
<tr>
<td>Delta Net OI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>( NetOI_{t} - NetOI_{t-1} )</td>
<td>Change in Net OI of index traders, measured in contracts, options positions excluded</td>
<td>AULERICH et al. (2010)</td>
</tr>
<tr>
<td>Relative magnitude Pos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent total Net OI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>( \frac{NetOI_{INDEX}^{INDEX}}{\sum NetOI^{ALL}_{ALL}} )</td>
<td>Net OI of index traders relative to total OI of all traders, options positions excluded</td>
<td>AULERICH et al. (2010)</td>
</tr>
<tr>
<td>Percent total LPOI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>( \frac{LPOI_{INDEX}^{INDEX}}{LPOI_{ALL}^{ALL}} )</td>
<td>LPOI of index traders relative to total OI of all traders</td>
<td>SANDERS and IRWIN (2011)</td>
</tr>
<tr>
<td>Absolute magnitude Pos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of Net OI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>Not available</td>
<td>Quantum index of Net OI of index traders in all ag. commodities included in CIT Report over period Jan06-Jun09, weights are prices as of 3 Jan 06</td>
<td>GILBERT (2010)</td>
</tr>
<tr>
<td>Net OI&lt;sub&gt;t&lt;/sub&gt;</td>
<td>( LPOI_{t} - SPOI_{t} )</td>
<td>Net OI of index traders, measured in contracts</td>
<td>ROBLES et al. (2009)/ SANDERS and IRWIN (2011)</td>
</tr>
</tbody>
</table>

### 3.3 Synthesis of results

Differences in individual model specifications naturally impede comparability of results. A summary of results on GC from Activity to Level are presented in Table 3. Overall, the studies find little evidence of GC, the null hypothesis of no GC can only be rejected for selected commodities with a level of significance \( \leq 5\% \). Most often, GC on future price levels is found with a time lag of one day (AULERICH et al. 2010). Both ROBLES et al. (2009) and GILBERT (2010) find evidence of GC causality (positive direction) on the FAO spot prices or the IMF food price index respectively. However, the direction of the impact is inconsistent among the studies. While some identify GC in a positive direction, i.e. an increase in index trading Granger-causes an increase in price levels, an almost equal amount of results
suggests a negative direction, i.e. an increase in index trading Granger-causes a decrease in price levels.

STOLL and WHALEY (2010) also use bi-directional GC tests and investigate whether GC runs from Level to Activity and find that they reject the null of no causality for W- (KCBT) (positive direction, 5% significance level) but not for any other commodity in their sample.

Table 3: **Results on Activity Granger-causing Level**

<table>
<thead>
<tr>
<th>Commodity (Exchange)</th>
<th>Time period (aggregation)</th>
<th>Activity variable</th>
<th>Price level variable</th>
<th>Significance level</th>
<th>Direction</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.</td>
<td>08/2006-12/2009 (weekly)</td>
<td>Net OI</td>
<td>Delta futures prices (nearby contract)</td>
<td>1%</td>
<td>–</td>
<td>IRWIN and SANDERS (2010a, 2010b)</td>
</tr>
<tr>
<td></td>
<td>01/2006-05/2008 (monthly)</td>
<td>Net OI</td>
<td>Spot prices</td>
<td>5%</td>
<td></td>
<td>ROBLES et al. (2010a)</td>
</tr>
<tr>
<td>S.</td>
<td>2004-2005 (daily)</td>
<td>Delta Net OI</td>
<td>Relative futures return (nearby contract)</td>
<td>1%</td>
<td>+</td>
<td>AULERICH et al. (2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Percent total Net OI</td>
<td>Relative futures return (nearby contract)</td>
<td>1%</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Percent total Net OI</td>
<td>Relative futures return (nearby contract)</td>
<td>5%</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2006-2008 (daily)</td>
<td>Percent total Net OI</td>
<td>Relative futures return (nearby contract)</td>
<td>1%</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Percent total Net OI</td>
<td>Relative futures return (first deferred contract)</td>
<td>5%</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Percent total Net OI</td>
<td>Relative futures return (first deferred contract)</td>
<td>1%</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>IMF Index</td>
<td>03/2006-06/2009 (monthly)</td>
<td>Index of Net OI</td>
<td>Delta spot price</td>
<td>5%</td>
<td>+</td>
<td>GILBERT (2010)</td>
</tr>
</tbody>
</table>

4 **Interpretation of results based on an extended theoretical background**

Since GC is a statistical concept (cf. PEARL 2010, p. 32) that relies on testing the $H_0$ of no causality given a pre-specified information set $\Omega$, its results need to be interpreted by taking into account the theory underlying the null hypothesis. This is especially important to assess the vulnerability of a model to any of the sources of sensitivity for GCA results. To this end, our understanding of the financial market element and any price effects from the financialization of agricultural commodity markets can be improved by integrating theory from financial economics.
4.1 Sources of sensitivity for GCA results

GRANGER (1980) gives one of the most extensive summaries of potential result sensitivities of standard GC tests that need to be respected. For the purpose of this paper, these have been extended with other relevant points where considered necessary. Due to the focus of this paper on the theoretical underlying of the null hypothesis, sensitivities that are not discussed in more detail are those arising from time series model specification such as lag lengths included in the underlying model (cf. ÖSTERMARK and AALTONEN 1999) and from the selection of proper statistical tests.

Omission of relevant variables

$\Omega_t$ needs to contain all relevant variables without containing redundant information. Once relevant variables are omitted, results on GC can be spurious or measure wrong feedback relations (cf. GRANGER 1980; LÜTKEPOHL 1982). Due to the frequent occurrence of omitted variable problems, LÜTKEPOHL (1982, p. 375) even generally doubts the reliability of bivariate frameworks for testing GC between economic variables.

Variable specification

Results on GC are always contingent on the definition of $\Omega_t$ and on the acceptance of the included proxy variables to represent the desired theoretical variables (cf. GRANGER 1980, p. 339). Once the measured variable does not contain all desired information, the actual information set $J_t$ deviates from the theoretical set $\Omega_t$.

Forward-looking behavior

If agents consider a variable $X$ important for forecasting a variable $Y$ and they form expectations about the future development of both variables and then act accordingly, expectations about $X_{t+1}$ may be present in time series data for $Y_t$. GC may then appear to run in the “wrong” direction from $Y$ to $X$ and any causal interpretation may be misleading (HAMILTON 1994, p. 307).

Temporal data aggregation

If the time lags between which causality occurs would require higher temporal data disaggregation than available from measured data, the true causal relationship will likely remain uncovered or, alternatively, instantaneous causality may be observed (GRANGER 1980, p. 340; BREITUNG and SWANSON 2002, p. 651).

Feedback relations

In difference to feedback between variables, GC only runs in one direction. However, if there is reason to assume that feedback relations exist in the data and the temporal disaggregation is insufficient to capture them, GC could be measured spuriously, depending on the chosen period of measurement.
*Time-varying effects*

Axiom III underlying the GC concept states that causal relations remain constant in their direction over time. This implies that time-varying effects where feedback relations or forward-looking behavior in prices would play a role during sub-periods of the analysis could disturb the results on GC if these sub-periods cannot ex ante be identified.

In order to assess the susceptibility of the results from the empirical models to any of these sources of sensitivity, the theory underlying the null hypothesis will in the following be augmented with relevant findings.

4.2 *Incomplete information sets due to omitted determinants of trading activity*

The given information set $J_t (\text{Level}_{t-i}; \text{Activity}_{t-j})$ used in the empirical studies could potentially be subject to an omitted variable bias if the determinants of Activity are not fully considered. If any variable simultaneously influences price levels and index trading activity, GC between the two can be spurious. Figure 2 illustrates important determinants of index trading activity. Associated with the ultimate trading activity, i.e. the observed position change on the futures markets, is a trading strategy, underlying which is again a genuine trading motive. The trading strategy can be realized either directly by trading positions on the future market or indirectly by making use of an instrument such as index funds, certificates or index return swaps. Index trading could be linked to a variety of motives. Apart from portfolio diversification, which will likely play an important role, alternative motives could be (financial) hedging to avoid risk incurred somewhere else, pure risk/return investment, where risk is taken to capture a potential return from a favorable price movement or arbitrage, which implies the search for an effectively riskless profit. The trading strategy will then incorporate decisions concerning e.g. the timing of the trade, the exact selection of indices to replicate, products in which to invest, or the size of the trade. Thus, the trading strategy consists of all behavioral rules necessary to translate the trading motive into actual trading activity.

In case of realization of the trading activity via an instrument, the chosen investment product will also have to replicate the benchmark index. Replication schemes can differ widely depending on the actual product. As mentioned above, index funds, for example, can directly take long positions in the index asset, i.e. the (agricultural) commodity futures, or use a synthetic replication scheme, e.g. via swaps. In the latter case, the swap dealer would then also have to replicate the index performance and hedge their swap position with the index fund. Such index replication activities and financial hedges conducted by index funds or swap dealers on the futures markets could be interpreted as derived trading motives realized
with a derived trading strategy. Derived thereby refers to the fact that they stem from the need to gain exposure to the index returns due to a previous inflow of liquidity into the investment products linked to a genuine motive and strategy.

Figure 2: Determinants of index trading activity in agricultural commodity futures markets

Consequently, the actual position changes on the futures markets will result from a mixture of genuine and derived trading motives and strategies. The determinants of the trading activity will therefore also vary – and might in parts be direct determinants of Level. For example, any genuine trading strategy that considers fundamental factors of supply and demand of agricultural commodities before investing into an index product would lead to problems if fundamentals are omitted from the information set since they influence both Activity and Level. This is by far no negligible exception but would hold true for any trading strategy with an underlying behavioral rule consistent with the rational expectations hypothesis (REH), as formally described by MUTH (1961). Any rational trader will consider fundamental factors in addition to all relevant other information when formulating expectations about future price movements (cf. CUTLER et al. May, 1990; FLOOD and HODRICK 1990). Consequently, in order to avoid omitted variable problems within the information set, a first requirement is that the trading activity can be disaggregated according to motives and strategies based on real world data.

4.3 Limitations to correct variable specification from data

Even though the empirical studies use a wide array of specifications both for the Activity variable, all are measured based on the INDEX positions in the CIT Report, i.e. the combined futures and options positions in the U.S. exchange-traded future markets. Any trading activity on the OTC market cannot be measured.

---

6 AULERICH et al. (2010) use futures-only INDEX positions.
Observable from the CIT Report data are the position holdings and changes associated with index trading, but only for those traders who exceed a specific position limit. The CFTC includes in this category all trading positions with an underlying passive strategy that only replicates a benchmark index via investment into the future markets. There are no active trading decisions involved. However, this passive strategy is essentially a derived trading strategy from a previous investment of liquidity into e.g. index funds or index return swaps. Therefore INDEX positions could include direct replication of index funds as well as hedging positions of swap dealers. In fact, it can be observed that INDEX LPOI (CIT Report) co-moves with SWAP LPOI (DCOT Report), which suggests that there is a strong overlap between these two, as shown in Figure 3 (CFTC 2011a).

Figure 3: Co-movement of swap dealer and index trader LPOI

The genuine trading strategies and motives that originally led to the derived index trading positions remain obscured in the CFTC reports. It is therefore possible to measure position changes asociable with passive index trading but not those variables that originally determined these position changes. Thus, the lack of ability to properly differentiate Activity variables according to their determinants can cause omitted variable problems in the information set as described above.

4.4 Consequences from informational efficiency and forward-looking behavior

Trading strategies consistent with the REH could not only cause omitted variable problems as explained above, but also problems related to forward-looking behavior, especially if markets are informationally efficient. In its strong form, the Efficient Market Hypothesis (EMH) states that prices should be fully aggregating public and private information. Under the semi-strong form, they only transmit public information and under the weak form, they only transmit information con-

Informationally efficient markets do not reveal by which mechanism the information was incorporated in the price (BROWN 2011, p. 82). However, accepting the REH for all traders suggests that markets should at least behave in accordance with the semi-strong form of the EMH, if no other factors distort the process of information aggregation (CUTLER et al. 1990, p. 65). If this can be accepted for agricultural commodity markets it would imply that it is reasonable to assume that expectations about future price movements and about the development of other factors relevant in forecasting returns would be incorporated in the price already in period t.

Thus, if including past values of Activity in an autoregressive model for Level would improve the forecast, it would be rational for market agents to formulate expectations about future developments of Activity and trade accordingly. If the forecasts of Activity_{t+i} are based on more information than past values of Activity and the market is informationally efficient, the prediction of Activity_{t+i} would be incorporated into Level, and GC could be measured as running from Level to Activity. Once this can be suspected, the test for GC can only be interpreted as a test for informational efficiency of the market (HAMILTON 1994, p. 307). Unfortunately, despite the fact that the EMH has been tested extensively for various financial markets, little research exists that systematically investigates the informational efficiency of agricultural commodity markets.

Critics of the EMH state that informational efficiency can be disturbed not only by imperfect institutional setups of the markets (BEJA and GOLDMAN 1980, pp. 235–236) or costly information gathering (cf. GROSSMAN 1976, p. 574) but also by a “psychological and behavioral element” once the REH is relaxed (MALKIEL 2003, p. 60). If these informational inefficiencies exist within a market, they are hypothesized to be predominantly present in short-term as transitory states of disequilibria (BEJA and GOLDMAN 1980, pp. 235–236; TIMMERMANN and GRANGER 2004). If this can be confirmed for agricultural commodity markets forecasts of Activity may not be fully incorporated in Level. However, once these inefficiencies are assumed to stem from boundedly rational behavior of market agents, there may be other problems arising from the interaction of such strategies on the market.

4.5 Time-varying or feedback effects from boundedly rational trading strategies

A common criticism of the REH refers to its failure to sufficiently incorporate human boundaries to rational expectation formation such as the judgmental biases and heuristics shown e.g. by KAHNEMAN and TVERSKY (1979). The notion of bounded rationality is described in SIMON (1991, p. 132) as “[…] the limits upon
the ability of human beings to adapt optimally, or even satisfactorily, to complex environments”. The bounded rational expectations hypothesis (BREH) considers such limits and psychological boundaries and allows for the possibility of agents to only selectively consider information in their expectation models. Therefore, accepting the BREH for market agents gives rise to the possibility of heterogeneous boundedly rational trading strategies. An early example of the consideration of such strategies is the “noise trader” who is trading on “pseudo-signals” or “popular models” (Black 1986; Shleifer and Summers 1990). Building on earlier examples such as De Long et al. (1990), Chiarella (1992) and Brock and Hommes (1998), the possible design and market effects of two stylized boundedly rational financial trading strategies have since been considered in a range of more recent studies.

A fundamental value trader (FVT) bases the trading strategy on price estimates about the fundamental value of an asset and thus considers all available information that can determine this price. The underlying expectation model is regressive in that market prices of an asset are assumed to fluctuate around, but eventually mean revert to their fundamental long-run equilibrium price ($P^E$) (Chiarella et al. 2002, p. 175; Farmer and Joshi 2002, p. 157; Westerhoff and Reitiz 2005, p. 645). The model by which $P^E$ is calculated is a potentially differentiating factor among FVTs, but in agricultural commodity markets can be assumed to incorporate at least some of the fundamental factors of supply and demand. The demand from a FVT ($D^{FVT}$) will thus be related to the perceived inequality between $P^E$ and the current market price ($P^M_t$), for example:

$$D^{FVT} = \gamma (P^E - P^M_t)$$

where $\gamma$ is a positive reaction coefficient. If $P^E > P^M_t$, the FVT will buy and if $P^E < P^M_t$, the FVT will sell the agricultural commodity (cf. Reitiz and Westerhoff 2007, p. 234; Farmer and Joshi 2002, p. 157).

The technical trader (TT) or chartist, on the other hand, only considers information that is implied in past and current prices (Farmer and Joshi 2002, p. 154; Chiarella et al. 2002, p. 174; Westerhoff and Reitiz 2005, p. 642). A special type of TT is the trend follower (TF) who trades based on an underlying extrapolative expectation model. A TF would then investigate the difference between the market price in period $t$ ($P^M_t$) and in period $t-n$, with $n$ representing the timescale ($P^M_{t-n}$) (e.g. Chiarella et al. 2002, p. 175; Farmer and Joshi 2002, p. 154), such that the demand of the TF ($D^{TF}$) could be represented as:

---

7 E.g. Farmer and Joshi (2002), Chiarella et al. (2002), Westerhoff and Franke (2009), Dieci and Westerhoff (2010), Tramontana et al. (2011) and specifically for the commodity market Westerhoff and Reitiz (2005), He and Westerhoff (2005) and Redrado et al. (2009).
\[ D^{TF} = \delta(P^M_t - P^M_{t-1}) \]

where \( \delta \) is a positive reaction coefficient. If \( P^M_t - P^M_{t-1} > 0 \), the TF will buy, if \( P^M_t - P^M_{t-1} < 0 \), the TF will sell (cf. REITZ and WESTERHOFF 2007, p. 234).

Market effects from the interaction of such trading strategies under the BREH are most commonly assessed with heterogeneous agent-based models. In isolation, FVTs should have a stabilizing and TFs a destabilizing effect on prices. However, interaction of both trader types can add an “endogenous component” to the market with cyclical swings and time-dynamic changes (e.g. REITZ and WESTERHOFF 2007, p. 233). Whether FVTs or TTs dominate the market and its price mechanism can vary. WESTERHOFF and REITZ (2005) and REITZ and WESTERHOFF (2007) investigate possible determinants for the market power of each trader group for the commodity markets by endogenizing the share of either FVTs or TTs in the market. Once TTs drive \( P^M \) away from \( P^E \), they increase the incentive for the FVTs to trade on the delta until at some point in time FVTs would dominate the market and \( P^M \) would closely fluctuate around \( P^E \). Conversely, the closer market prices fluctuate around their fundamental values, the lower will be the incentive for FVTs to trade in the market and a dominant group of TFs could drive prices away from their equilibrium. Some models allow for traders to switch between strategies, for example based on past performance, expected future performance of another strategy or simple “swings in opinion” (WESTERHOFF and FRANKE 2009, p. 2). Also, if traders believe that in the long-run prices will mean revert but in the short run prices are driven by TFs, they may employ a FVT strategy for their long-term positions but a TF strategy for their short-term positions (cf. DELONG et al. 1990, p. 393).

Since both time-varying and feedback effects can cause results of GCA to become unreliable and hinder interpretation, accepting the BREH and allowing for boundedly rational trading strategies would mean that markets can become informationally inefficient (at least in the short-term) and time-varying and feedback effects may be present. The latter will predominantly occur if TFs have a dominant market position. In addition, if FVTs have sufficient market power, the same omitted variable problems may occur with respect to fundamental factors as already discussed in the previous sections.

4.6 Expectable time lags of causality

The time lags of causality that can be assumed based on theory need to be carefully assessed in order to capture it correctly. One the one hand, informational efficiency of the markets with respect to Activity would imply that any information and expectation on Activity would immediately be reflected in Level. On the other hand, allowing for the possibility of market inefficiencies, an alternative theory about time lags of causality is needed. It could for example be possible that it is a
gradual built up of INDEX positions that affects Level. Thus, it would not be the weekly or daily cause-and-effect relations that are relevant but rather a more long-term relation where only a specific level of INDEX positions above a threshold or a relative size of INDEX positions with respect to some other factor would affect Level. In order to capture such effects, alternative models would be needed.

AULERICH et al. (2010) and SANDERS and IRWIN (2011) also employ the long horizon regression fads model from JEGADEESH (1991). While AULERICH et al. (2010) find more long-term evidence of Activity Granger-causing Level compared to the short-run results discussed above, it is still comparatively weak. SANDERS and IRWIN (2011) do not find any difference compared to the short-run results except for the case of S-. However, these models still cannot account for the above mentioned threshold or relative size effects and therefore, alternative approaches based on augmented null hypotheses are still desirable.

In summary, all of the above factors could potentially affect reliability of the results from the reviewed empirical models and could help to explain the ambiguity e.g. concerning the direction of causality (price increasing or decreasing effect), the identified reverse causality or the generally inconclusive results regarding the significance of the Activity variable. However, agricultural commodity markets in general and the interaction between their physical and financial elements are still under-researched. More research is required to better formulate theoretical hypotheses and test these with adequate empirical models. Most importantly, the trading strategies used not only by financial investors but also by the traditional market participants who are hedging their physical commodity exposure should be better understood and adequate models developed that enable formulation of hypotheses on their interaction effects. These could as well take into account boundedly rational trading strategies.

5 Conclusion

This paper has conducted a critical review of existing empirical models of the price level effects of index trading activity in agricultural commodity markets based on GCA. Thereby the underlying research hypotheses have been augmented by including findings from financial market theory. Based on sources of sensitivities of GCA results, it could be shown that the determinants of trading activity, the informational market efficiency, or expectation models of the market agents could affect GC test reliability and interpretability and in part explain the inconclusive results we obtain from a synthesis of existing empirical studies.

First, measurement of the theoretical Activity variables from existing data is problematic. Index trading activity in particular could either be the result of a genuine trading strategy and motive or due to a derived strategy of an investment product resulting from a previous inflow of liquidity into that product. The determinants
of the genuine trading strategy remain obscured. Specification of the information set may therefore lead to an omitted variable bias, especially in light of trading strategies that consider fundamental factors of supply and demand. Better data availability from the CFTC such as the publication of the onetime reports on Large Traders Net Position Changes and Trading Account Net Position Changes on 30 June 2011 is therefore unlikely to significantly improve variable specification.

Second, a high degree of informational efficiency of the markets could lead to expectations about Activity being incorporated in Level and therefore measurement of GC running in the “wrong” direction. Such high informational efficiency can for example be expected if the REH is accepted for all traders. Third, boundedly rational heterogeneous trading strategies, on the other hand, could lead to time-varying and feedback effects over the time period of observation, which could also result in spurious measurement of GC. Fourth, expectable time lags of causality may vary if for example threshold or relative-size-effects exist.

Any problems resulting purely from the time series model specification or GC test selection have not been addressed. The sampled studies only focused on bivariate linear models. Multivariate or nonlinear models may be more adequate in light of potential omitted variable problems or nonlinear effects. However, given the implications of the above discussed theoretical problems, it seems necessary to first improve our theoretical hypotheses before we increase the specificity of our empirical models.

Consequently, open questions concerning the relevance and importance of specific trading motives and strategies on agricultural commodity futures markets or the informational market efficiency need to be answered. In addition, spillover effects from futures to spot markets and from other financial to agricultural commodity markets should be investigated to obtain a more comprehensive picture of today’s market environment. However, most importantly, the interaction between the financial and physical market elements, which is a unique feature of (agricultural) commodity markets, has to be closer examined as it will likely be the interplay between these two that is responsible for the price mechanisms we observe on today’s markets. To this end, research combining the agricultural and financial sub-disciplines of economics is needed to augment our theoretical hypotheses and choose adequate empirical models.
References

AULERICH, N.M. et al. (2010): The price impact of index funds in commodity futures markets: Evidence from the CFTC's daily large reporting system, Working paper, Department of Agricultural and Consumer Economics, University of Illinois.


www.blackrockinternational.com/content/groups/internationalsite/documents/literature/etfl_globalhandbook_q111.pdf, last accessed 10.05.2011.


CFTC (n.d.c): Index Investment Data, Explanatory Notes.


http://cftc.gov/MarketReports/LgTraderExplanatory#, last accessed 08.11.2011.


