



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search  
<http://ageconsearch.umn.edu>  
[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Discussion Paper 2014:5

# Price dynamics and financialization effects in corn futures markets with heterogeneous traders

*Stephanie-Carolin Grosche*

*Thomas Heckelei*

---

The series "Agricultural and Resource Economics, Discussion Paper" contains preliminary manuscripts which are not (yet) published in professional journals, but have been subjected to an internal review. Comments and criticisms are welcome and should be sent to the author(s) directly. All citations need to be cleared with the corresponding author or the editor.

Editor: Thomas Heckelei

Institute for Food and Resource Economics

University of Bonn

Nußallee 21

53115 Bonn, Germany

Phone: +49-228-732332

Fax: +49-228-734693

E-mail: [thomas.heckelei@ilr.uni-bonn.de](mailto:thomas.heckelei@ilr.uni-bonn.de)

# Price dynamics and financialization effects in corn futures markets with heterogeneous traders

Stephanie-Carolin Grosche<sup>a</sup> \*, Thomas Heckelei<sup>a</sup>

May 2014

<sup>a</sup> Institute for Food and Resource Economics, University of Bonn, Nussallee 21, 53115 Bonn, Germany

\* Corresponding Author. E-Mail address: stephanie.grosche@ilr.uni-bonn.de; Phone: +49 228 732326

## Abstract

Presumed portfolio benefits of commodities and the availability of index fund-type investment products increase attractiveness of commodity markets for financial traders. But resulting “index trading” strategies are suspected to inflate commodity prices above their fundamental value. We use a Heterogeneous Agent Model for the corn futures market, which can depict price dynamics from the interaction of fundamentalist commercial traders and chartist speculators, and estimate its parameters with the Method of Simulated Moments. In a scenario-based approach, we introduce index funds and simulate price effects from their inclusion in financial portfolio strategies. Results show that the additional long-only trading volume on the market does not inflate price levels but increases return volatility.

*JEL classification:* D84, G15, G17, Q02

*Key words:* Heterogeneous agents; Agent-based modeling; Commodity index trading; Financialization of commodity markets

# 1 Introduction

Alleged benefits of commodities in financial portfolio strategies (cf. Ankrum and Hensel 1993; Anson 1999; Gorton and Rouwenhorst 2006) have sparked interest in financial commodity investment and promoted the creation of commodity index funds.<sup>1</sup> In the period 2005-2010, assets under management of exchange traded commodity index funds increased from 1.2 to 45.7 billion U.S. Dollars (BlackRock 2011). These funds facilitate market entry for investors who are interested in the return of a diversified commodity portfolio, but are hesitant to trade single futures contracts. Nevertheless, the index funds need to replicate the index return, e.g. by engaging in “index trading” activities in the single futures markets, which corresponds to taking long (buy-side) positions and rolling these positions forward (cp. CFTC 2014a). This has consequences for trading volume on agricultural futures markets. In the case of Chicago Board of Trade (CBOT) corn futures, volume in the active contract more than doubled from 31 thousand contracts in the period 2000-2005 to 73 thousand contracts between 2005-2010 (Bloomberg data). And, U.S. Commodity Futures Trading Commission (CFTC) reports show that long position open interest in CBOT corn futures and options associated with index trading was at an average of 25% of total open interest over the period 2006-2013.

The influence of the commodity index trading volume on price levels and volatilities in agricultural commodity markets have been vividly discussed since the 2007/08 food price crisis, without reaching a definite consensus. According to the prominent “Master’s hypothesis”<sup>2</sup>, index trading drives commodity price bubbles by creating a constant artificial demand on the futures markets that is disconnected from market fundamentals (cf. Irwin and Sanders 2012; Will et al. 2012). Others reject this hypothesis, stating that an increase in long positions would only affect price levels if it were suspected to convey new information, due to the theoretical possibility to create an infinite amount of futures contracts at a given price (e.g. Irwin et al. 2009). Empirical studies have not succeeded in resolving this theoretical debate. The analysis of direct price level, return or volatility effects from a change in index trading volume on futures markets with help of Granger Causality tests (e.g. Robles et al. 2009; Gilbert 2010; Stoll and Whaley 2010; Sanders and Irwin 2011a,b; Gilbert and Pfuderer 2014) has led to inconclusive results, and further difficulties arise in

---

<sup>1</sup>We will in the following use the term “index funds” for all financial products that replicate a commodity index.

<sup>2</sup>Authors frequently use this term to refer to the statements of the U.S. hedge fund manager Michael W. Masters in front of Congressional hearings or the CFTC.

their interpretation as evidence of presence or absence of a price influence (Grosche 2014). On the other hand, the analysis of indirect effects such as changing return or volatility interdependencies between commodity and traditional asset markets (e.g. Diebold and Yilmaz 2012; Ji and Fan 2012; Silvennoinen and Thorp 2013; Mensi et al. 2013; Gao and Liu 2014; Grosche and Heckelevi 2014) or tests for rational bubbles (e.g. Gutierrez 2013; Liu et al. 2013; Etienne et al. 2014) do not allow a direct causal attribution of these effects to specific trading strategies.

We take an alternative approach and investigate price effects from index trading within a heterogeneous agent model (HAM) that simulates price dynamics emerging from the interaction of a few stylized heterogeneous trader types. These models have previously been applied to financial markets (see e.g. Hommes (2006) for a survey) but there has hitherto been scant application to agricultural commodity markets (exceptions are Westerhoff and Reitz (2005); He and Westerhoff (2005); Reitz and Westerhoff (2007); Redrado et al. (2009)) and only Redrado et al. (2009) specifically consider price effects from financialization. In our model, we first simulate a base scenario where index funds are unavailable. In a later “financialization” scenario, financial portfolio managers include commodities in their portfolio but only via index fund shares. Parameters for the base scenario are empirically estimated with the Method of Simulated Moments (MSM) (Lee and Ingram 1991; Duffie and Singleton 1993). Its use in HAM parameter estimation has recently been developed in e.g. Winker et al. (2007); Franke (2009); Franke and Westerhoff (2011, 2012). We complement these applications with refinements in parameter validation. The focus is on CBOT corn futures due the importance of corn in global agricultural production and its comparatively large futures market. Corn has the highest trading volume on the CBOT and the largest S&P Goldman Sachs Commodity Index (S&P GSCI) percentage dollar weight among the agricultural commodities.

In the remainder of the paper we first provide some background on commodity index funds and on the general setup of few-type HAMs. Second, we describe our Commodity HAM and the procedure we use for estimation and validation of the model parameters. We then proceed with a discussion of results and the final section concludes the analysis.

## 2 Background

A discussion of strategies and replication schemes of commodity index funds provides the necessary background to model the portfolio managers' trading activities. And, a brief overview of the general setup and previous applications of few-type financial market HAMs serves as the conceptual basis for our Commodity HAM.

### 2.1 Commodity index funds and index trading

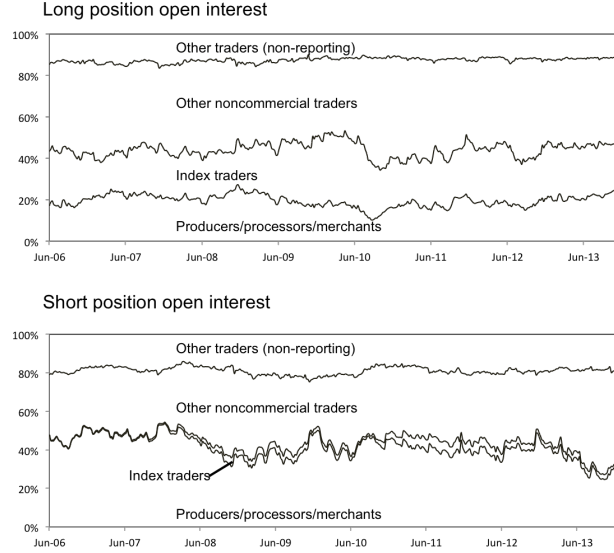
Index funds are in essence investment products that replicate the performance of a specific underlying index. Its investors gain exposure to the index return by buying a share in the fund and thus do not have to trade single futures contracts. The index fund itself then replicates the index either directly by taking adequate long positions in the futures markets or synthetically by engaging in an index return swap with a swap dealer. In the latter case, the swap dealer could then choose to hedge the open position by taking long positions in the futures market. Both direct or synthetic replication can thus ultimately lead to an increase in index trading positions in the single futures markets. The magnitude of these position holdings can be assessed with the CFTC weekly Commodity Index Trader report, which is a supplement to the (Disaggregated) Commitment of Traders report. The "index trader" category groups all positions associated with index trading strategies. Figure 1 shows the development of long and short position open interest for index traders and other trader types in the CBOT corn futures and options markets over the period 13 June 2006-31 December 2013. Thereby, "producers/processors/merchants" refers to those traders that deal with the physical commodity and hedge their positions on the futures markets. The "other noncommercial trader category" includes hedge funds or Commodity Trading Advisors and Commodity Pool Operators who trade on behalf of their clients (CFTC 2014b). Unsurprisingly, "index traders" hold a sizable share in long position open interest while their short position open interest share is negligible.

While the overall share of index trader open interest is at a relatively constant 25% level, changes in their position holdings will occur on a daily basis. Reweightings of the underlying index only play a minor role here.<sup>3</sup> The daily fluctuations primarily stem from changes in the desired replication volume. Such changes are the result of investors buying or selling shares in the index fund. The more liquidity flows into the

---

<sup>3</sup>E.g. substantial S&P GSCI reweightings only occur annually with smaller monthly reviews (S&P Dow Jones Indices 2014)

Figure 1: CBOT corn futures and options trader type volume shares



Source: CFTC

fund, the larger the return cash flow that has to be paid out to the investors and the larger the ultimate long position on the futures market used for return replication. Thus, even though the index fund itself has a passive strategy and only replicates the index, a higher (lower) attractiveness of commodities as financial investments will nevertheless increase (decrease) the size of the total index trader long position.

## 2.2 Few-type heterogeneous agent models for financial markets

In the past, few-type HAMs have frequently been applied to investigate price dynamics in financial markets. In difference to their many-type counterparts they do not attempt to explicitly model the multitude of possible real world trading strategies but rather focus on selected stylized trading rules. Existing HAMs differ with respect to their market focus. Some models concentrate on exchange rates and/or equities (e.g. Bauer et al. 2009; Manzan and Westerhoff 2005; Franke and Westerhoff 2011, 2012). For commodity markets, Westerhoff and Reitz (2005) and Reitz and Westerhoff (2007) focus on corn and on cotton, lead, rice, sugar, soybeans and zinc respectively. He and Westerhoff (2005) build a HAM for a general commodity market, Redrado et al. (2009) for a mixed commodity index and Ellen and Zwinkels

(2010) for the oil market. Finally, Alfarano et al. (2005) develop a model applicable to a broader range of asset markets and base their empirical estimation on gold and selected German stock (index) price series. The interaction of different markets is modeled e.g. in Westerhoff (2012) for a Keynesian goods and a stock market, in Chiarella et al. (2005) and Chiarella et al. (2007) for multiple risky and one risk free asset and in Dieci and Westerhoff (2010) for two international stock markets inked via a foreign exchange market. HAMs have also been used for policy analysis. For example, Anufriev and Tuinstra (2013) model the effect of short-selling constraints, Westerhoff (2003) investigate the effectiveness of price limits and resulting trading breaks, He and Westerhoff (2005) analyze the effects of a minimum and maximum price and Westerhoff and Dieci (2006) investigate transaction taxes.

In a few-type HAM, price dynamics arise from the interaction of selected stylized heterogeneous trading strategies. Commonly, these strategies are either of a fundamentalist or a chartist nature and their development goes back to e.g. Zeeman (1974), Beja and Goldman (1980) and Frankel and Froot (1990). While fundamentalists expect that market prices will revert back to their fundamental equilibrium value, chartists believe that prices follow a trend that can be extrapolated. Formally, a basic expression of the fundamentalist ( $V_t^F$ ) and chartist trading volume ( $V_t^C$ ) is given by:

$$V_t^F = \gamma_F(P_F - P_t), \quad (1)$$

$$V_t^C = \gamma_C(P_t - P_{t-1}), \quad (2)$$

where  $P_t$  is the *log* of the market price in period  $t$ ,  $P_F$  is the logged constant fundamental price of the asset and  $\gamma_F$  and  $\gamma_C$  are positive reaction coefficients, measuring how responsive a trader type is to the observed price movement.

Markets can be in disequilibrium and prices are determined from either excess supply or demand on the market. A positive trading volume equals demand and a negative volume supply. A simple price-impact function is defined by:

$$P_{t+1} = P_t + (\phi_t^F V_t^F + \phi_t^C V_t^C), \quad (3)$$

Where  $\phi_t^F + \phi_t^C = 1$  are the relative weights of the respective fundamentalist or chartist trader groups on the market. These weights are often assumed to be time-dependent and to vary according to a switching mechanism. The exact design of this switching mechanism is model-specific, depending on the underlying assumptions



and the desired degree of complexity.

The complexity of the price dynamics that emerge from the traders' interaction over time affect calibration and estimation of the model parameters. While some models may in part allow analytical derivations (e.g. Chiarella 1992; Lux 1997; Chiarella et al. 2002) or permit direct estimation of their parameters (e.g. Alfarano et al. 2005; Westerhoff and Reitz 2005; Reitz and Westerhoff 2007; Redrado et al. 2009), more complex model setups require a simulation-based solution approach. Thereby, parameters are sometimes set "by hand" (e.g. Westerhoff 2003; Manzan and Westerhoff 2005) and their simulated return properties *ex-post* compared to empirical returns. Recently, progress has been made in the area of simulation-based estimation with the MSM where a part of the model parameters is estimated by *simultaneously* setting parameter values and considering differences between simulated and empirical returns. Building on Gilli and Winker (2003), Winker et al. (2007) demonstrate how to set up an objective function, Franke (2009) extends the work by more explicitly considering the quality of moment-matching, and recently, Franke and Westerhoff (2011, 2012) demonstrate the use of measures of model fit in evaluating the quality of model parameters and comparing models.

### 3 Commodity market HAM

Our Commodity HAM is from a market perspective most closely related to the corn model in Westerhoff and Reitz (2005). But while their direct parameter estimation approach necessitated a relatively simple model setup, estimation with the MSM permits more complex dynamics. For our base scenario HAM we follow the "structural stochastic volatility" (SSV) approach developed in Franke and Westerhoff (2011, 2012). Our base scenario models the time period before 2006, i.e. before the strong growth of commodity index funds. Its setup closely follows the "DCA-TPM" model introduced in Franke and Westerhoff (2012) but in formulating the basic commodity trading strategies we draw some connection to the CFTC trader categories. The financialization scenario then simulates the market entry of a portfolio manager who uses commodities as portfolio diversifiers but does not trade directly on the futures markets but only via index funds.

### 3.1 The base scenario

The “producers/processors/merchants” from the CFTC reports can be interpreted as “commercial traders” (CO) who have some idea about the fundamental value of the commodity (from their primary business operations) and will use this knowledge to trade accordingly. Their trading volume ( $V_t^{CO}$ ) is generated by a fundamentalist strategy, such that:

$$V_t^{CO} = \gamma_{CO}(P_F - P_t) + \epsilon_t^{CO}, \quad \epsilon_t^{CO} \sim N(0, \sigma_{CO}^2). \quad (4)$$

The first term in the volume equation represents the deterministic volume that stems from deviations between the fundamental price and the current market price. The reaction coefficient  $\gamma_{CO}$  determines how strong the commercial traders’ volume reacts to such perceived deviations. The second term is a stochastic volume. In our Commodity HAM this component could capture random shocks due to the traders’ different estimates of the fundamental value.

The “other noncommercial traders” are assumed to be trading on price data rather than on fundamentals. They follow trends and can thus be interpreted as “speculators” (S). Their trading volume ( $V_t^S$ ) is represented by a chartist strategy:

$$V_t^S = \gamma_S(P_t - P_{t-1}) + \epsilon_t^S, \quad \epsilon_t^S \sim N(0, \sigma_S^2). \quad (5)$$

Again, the first term in the volume equation represents the deterministic volume, which depends on daily price changes and the reaction coefficient  $\gamma_S$  determines how strongly the speculators react to price trends. The second term adds a stochastic volume, which accounts for additional variation in the trading rules (cf. Westerhoff 2003). The commercial traders’ and speculators’ stochastic volumes are fully independent.

Both trader types trade directly on the commodity futures market.<sup>4</sup> The total contract trading volume is composed of fundamentalist volume ( $V_t^F$ ) from the

---

<sup>4</sup>Even though trading volume corresponds to contract holdings, we will not consider any rolling effects over the time period but assume that trading out of the active and into the first deferred contract can be achieved without any transaction costs, which is equivalent to holding an artificial active contract over the full simulation period.

commercial traders and chartist volume ( $V_t^C$ ) from the speculators, such that:

$$\begin{aligned} V_t^F &= \phi_t^F V_t^{CO}, \\ V_t^C &= \phi_t^C V_t^S, \end{aligned} \tag{6}$$

where  $\phi_t^F$  and  $\phi_t^C$  are the market weights of the respective trading strategies. The price-impact function is given by:

$$P_{t+1} = P_t + \gamma_{MM} (V_t^F + V_t^C). \tag{7}$$

The coefficient  $\gamma_{MM}$  is a positive reaction coefficient from a “market maker” who somewhat balances supply and demand by releasing inventory in case of excess demand and taking inventory in case of excess supply to avoid extreme spikes (cf. Westerhoff 2003; Franke and Westerhoff 2012).

We allow the market weights  $\phi_t^F$  and  $\phi_t^C$  to vary based on relative strategy attractiveness. For our commercial trader-speculator setting, we can imagine that a higher attractiveness of a fundamentalist strategy induces more commercial traders to enter the market and speculators to leave, and vice versa. In determining relative strategy attractiveness, we follow the “DCA-TPM” model approach in Franke and Westerhoff (2012) and compute an attractiveness index of a fundamentalist strategy ( $\alpha_t$ ) as:

$$\alpha_t = \alpha_p + \alpha_h(\phi_t^F - \phi_t^C) + \alpha_m(P_t - P_F)^2, \quad \alpha_h, \alpha_m > 0. \tag{8}$$

The first summand is the predisposition parameter ( $\alpha_p$ ), which measures whether traders have an à priori strategy preference, whereby a positive (negative) value indicates preference for fundamentalism (chartism). The second summand accounts for the tendency of the traders to follow the herd, i.e. join a group that is already dominating the market. Thus, if  $\phi_t^F > \phi_t^C$ , the attractiveness of fundamentalism increases and the parameter  $\alpha_h$  defines the strength of the increase. The last summand accounts for a potential fear of bubbles. The stronger the misalignment between the current and the fundamental price, the higher the attractiveness of fundamentalism. Speculators would leave the market in expectation of a bubble. The parameter  $\alpha_h$  measures how strongly price misalignment affects attractiveness of fundamentalism.

The functional relation between the market weights and the attractiveness index is modeled with a “Discrete Choice Approach” (DCA) (Brock and Hommes 1998),

where the attractiveness index directly affects the level of market shares:<sup>5</sup>

$$\begin{aligned}\phi_t^F &= \frac{1}{1 + \exp(-\beta\alpha_{t-1})}, \\ \phi_t^C &= 1 - \phi_t^F,\end{aligned}\tag{9}$$

where  $\beta$  is the “intensity of choice” parameter that could be used to scale the level of the attractiveness index in the above equation (Franke and Westerhoff 2012).

Inserting the above equations into equation (7), leads to:

$$\begin{aligned}P_{t+1} &= P_t + \gamma_{MM}(\phi_t^F(\gamma_{CO}(P_F - P_t) + \epsilon_t^{CO}) + \phi_t^C(\gamma_S(P_t - P_{t-1}) + \epsilon_t^S)), \\ \Leftrightarrow P_{t+1} &= P_t + \gamma_{MM}(\phi_t^F(\gamma_{CO}(P_F - P_t)) + \phi_t^C(\gamma_S(P_t - P_{t-1})) + \epsilon_t^P, \\ \epsilon_t^P &\sim N(0, \sigma_{P,t}^2), \\ \sigma_{P,t}^2 &= (\phi_t^F)^2 \sigma_{CO}^2 + (\phi_t^C)^2 \sigma_S^2.\end{aligned}\tag{10}$$

Thus, the variance of the stochastic trading volumes and the trader weights affect the *time-dependent* variance of the stochastic price component, which is key to the SSV model approach (Franke and Westerhoff 2011, 2012).

### 3.2 The financialization scenario

We assume that the portfolio managers’ decision on the level of investment in commodity index funds will depend on both idiosyncratic returns of the single commodities in the index and on commodity index return or volatility correlations with other portfolio assets. Trading volume associated with single commodity returns stems from an underlying weighted fundamentalist-chartist strategy, similar to the portfolio manager in Redrado et al. (2009), while trading volume as a result of portfolio correlations is modeled as a stochastic component. Total portfolio managers’ trading volume ( $V_t^{PM}$ ) is expressed as:

$$V_t^{PM} = \gamma_{PM}[\tilde{\phi}_t^F(P_F - P_t) + \tilde{\phi}_t^C(P_t - P_{t-1})] + \epsilon_t^{PM}, \quad \epsilon_t^{PM} \sim N(0, \sigma_{PM}^2)\tag{11}$$

where the first two summands show the deterministic volume and  $\tilde{\phi}_t^F$  and  $\tilde{\phi}_t^C$  represent the relative fundamentalist and chartists volume weights and  $\gamma_{PM}$  determines

---

<sup>5</sup>An alternative is the Transition Probability Approach (TPA) where the effect on the *rates of change* of the trader-type population shares is modeled. As demonstrated in Franke and Westerhoff (2012), if  $\alpha_t$  is composed of the same elements ( $\alpha_p, \alpha_h, \alpha_m$ ), there will be no major difference in results between DCA and TPA and we choose DCA due to its comparative popularity in the literature.

the reaction strength to price deviations.  $\epsilon_t^{PM}$  is the stochastic volume and assumed to be independent of either the stochastic commercial traders' and speculators' volumes.

Since portfolio managers' trading volume ultimately reaches the futures market via index fund replication volume, the total position has to be net long. Therefore, in any period  $t$  contract demand is equivalent to the volume derived from equation (11). But contract supply cannot exceed the total long position that has been built up until period  $t$ . Formally, this restricted volume ( $\tilde{V}_t^{PM}$ ) is expressed as:

$$\tilde{V}_t^{PM} = \begin{cases} \max \left[ V_t^{PM}, -\sum_{i=1}^{t-1} V_i^{PM} \cdot (1 - \chi) \right] & , \text{ if } V_t^{PM} < 0 \\ V_t^{PM} & , \text{ otherwise.} \end{cases} \quad (12)$$

As a total position holding of zero would strictly mean that index funds go out of business, the parameter  $\chi$  is introduced as a percentage minimum position holding leading to a moving lower bound for  $\tilde{V}_t^{PM}$ .

The combined deterministic trading volume of all three trader types is still only associated with either a fundamentalist or a chartist strategy, whereby total fundamentalist and chartist volumes are calculated as:

$$\begin{aligned} V_t^F &= \phi_t^F V_t^{CO} + \tilde{\phi}_t^F \tilde{V}_t^{PM}, \\ V_t^C &= \phi_t^C V_t^S + \tilde{\phi}_t^C \tilde{V}_t^{PM}. \end{aligned} \quad (13)$$

This assumes that the fundamentalist/chartist shares in  $\tilde{V}_t^{PM}$  are the same as in  $V_t^{PM}$  and that size of  $\tilde{\phi}_t^F$  and  $\tilde{\phi}_t^C$ , i.e. the weight of the fundamentalist and chartist components in the portfolio managers' volume, are also determined with a DCA approach from the attractiveness index  $\alpha_t$ . But, the herding component within  $\alpha_t$  now needs to take into account the additional portfolio managers' fundamentalist and chartist volume, which is why we now use absolute fundamentalist ( $\psi_t^F$ ) and chartist ( $\psi_t^C$ ) volume shares, calculated as:

$$\begin{aligned} \psi_t^F &= \frac{\phi_t^F |V_t^{CO}| + \tilde{\phi}_t^F |\tilde{V}_t^{PM}|}{(\phi_t^F |V_t^{CO}| + \phi_t^C |V_t^S| + |\tilde{V}_t^{PM}|)}, \\ \psi_t^C &= \frac{\phi_t^C |V_t^S| + \tilde{\phi}_t^C |\tilde{V}_t^{PM}|}{(\phi_t^F |V_t^{CO}| + \phi_t^C |V_t^S| + |\tilde{V}_t^{PM}|)}, \end{aligned} \quad (14)$$

such that  $\alpha_t$  becomes:

$$\alpha_t = \alpha_p + \alpha_h(\psi_t^F - \psi_t^C) + \alpha_m(P_t - P_F)^2, \quad (15)$$

Inserting the above trading volumes in the price impact function leads to:

$$\begin{aligned} P_{t+1} &= P_t + \gamma_{MM}(\phi_t^F(\gamma_{CO}(P_F - P_t)) + \phi_t^C(\gamma_S(P_t - P_{t-1})) + \epsilon_t^P + \tilde{V}_t^{PM}, \\ \sigma_{P,t}^2 &= (\phi_t^F)^2 \sigma_{CO}^2 + (\phi_t^C)^2 \sigma_S^2 + \tilde{\sigma}_{PM}^2, \end{aligned} \quad (16)$$

where the tilde in  $\tilde{\sigma}_{PM}^2$  indicates that the variance will be affected by the short-selling constraint as it truncates the distribution.

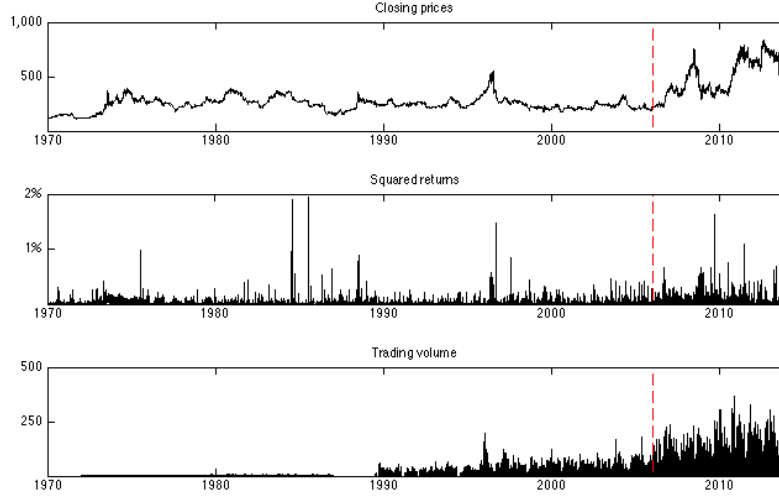
## 4 Base scenario parameter estimation

With the MSM, model parameters are chosen such that moments calculated from simulated returns come close to their empirical counterparts from daily relative returns of CBOT corn. We use price data for the trading days 01/05/1970-12/31/2013 from which we split off the base scenario sample ending on 12/31/2005 with a total number of 9,085 observations. This base period is used for later parameter estimation. Prices are Bloomberg's first generic contract prices where expiring active futures contracts are rolled to the next deferred contract on the last trading day of the active contract ("relative to expiration" rolling procedure). We calculate relative returns ( $R_t$ ) as  $R_t = \ln P_t - \ln P_{t-1}$ , where  $P_t$  and  $P_{t-1}$  are the closing prices. Squared returns ( $R_t^2$ ) are used to approximate short-term price volatility. Figure 2 shows the development of closing prices in U.S. Dollars, squared returns and trading volume (in thousand contracts) until 12/31/2013. The vertical dashed line indicates the end of the base period. The price level and short-term volatility markedly increase in the period after 2006 while trading volume surges.

### 4.1 Selection and calculation of moments

We first select and calculate the moments before we set up the objective function and continue with its minimization and parameter validation. Moments are chosen to capture important stylized facts of the corn futures prices. Commonly, financial market HAMs aim to replicate the overall volatility level in the price data, the non-autocorrelation property of returns, fat tails in the return frequency distribution, volatility clustering and long memory effects (cf. Franke and Westerhoff

Figure 2: Price, volatility and volume development



Source: Bloomberg

2011, 2012). While overall volatility is market specific, most of the stylized facts of other financial markets will also be found in commodity returns (Westerhoff and Reitz 2005). Our selection of moments follows suggestions in Franke and Westerhoff (2012) and Winker and Jeleskovic (2007) and calculated empirical moments for both the baseline and the full observation period are presented in Table 1.

The overall volatility level is captured by calculating the absolute value returns, which show a similar behavior to squared returns (Franke 2009) and can also be used to calculate the autocorrelation function, as illustrated below.

$$\bar{A}_t = \frac{1}{N} \sum_{t=1}^N A_t, \quad A_t = |R_t| \quad (17)$$

The first order autocorrelation coefficient of relative returns is close to zero (non-autocorrelation property of returns) while the autocorrelations of absolute returns should be slowly decaying over the time horizon to demonstrate the long memory effect in the return data. As pointed out by Franke and Westerhoff (2012), this property is also related to volatility clustering. Autocorrelations of lag  $k$  are computed

from the sample autocovariances ( $ACV$ ) as:

$$ACV_k = 1/T \sum_{t=k+1}^T (Q_t - \bar{Q}_t)(Q_{t-k} - \bar{Q}_t), \quad (18)$$

with  $Q_t = R_t, A_t$ . The sample autocorrelation for either relative ( $\rho_k^r$ ) or absolute returns ( $\rho_k^a$ ) is then computed as:  $ACV_k/ACV_0$ . For relative returns we only calculate  $\rho_1^r$  and rely on the finding in Franke and Westerhoff (2012) that autocorrelations at larger lags will vanish once  $\rho_1^r$  is close to 0. For absolute returns, we match the profile of the decaying autocorrelation function (ACF) by computing  $\rho_1^{ac}, \rho_5^{ac}, \rho_{10}^{ac}, \rho_{25}^{ac}, \rho_{50}^{ac}, \rho_{100}^{ac}$ . The exponent  $ac$  denotes that these are “centered” autocorrelation coefficients of absolute returns. The ACF of  $R_t$  may have an autocorrelation coefficient at one of our selected lags that contradicts the typical behavior (e.g. in our corn sample,  $\rho_5^a$  is larger than the  $\rho_1^a$ ). Since we would not expect our simulated returns to match these properties, we follow Franke and Westerhoff (2012) and smooth this effect by calculating the mean of the autocorrelation coefficients at the selected lag and the two lags surrounding it (or one lag in case of  $\rho_1^{ac}$ ).

The fatness of the tail of the frequency distribution of returns is commonly measured with the Hill-estimator ( $\xi$ ) and its corresponding tail index ( $tail = 1/\xi$ ) (Hill 1975). The smaller its value, the fatter the tail of the distribution. Also, moments above the value of the tail index are no longer defined. To compute the tail index for  $(R_t)$ , the returns are first ordered and we write the order statistics from the return sample  $R_1, R_2, \dots, R_T$  as  $R_1^T \geq R_2^T \geq \dots \geq R_T^T$ . The right tail index is computed as:<sup>6</sup>

$$\xi = \frac{1}{k} \sum_{i=1}^k \ln R_i^T - \ln R_k^T \quad (19)$$

$$tail = 1/\xi.$$

The Hill estimator is sensitive to the choice of  $k$ . Standard choices are either 5% or 10% of the total observations. We follow Winker and Jeleskovic (2007) and use the mean of the tail indices for  $k = 0.05 \cdot N$  and  $k = 0.1 \cdot N$ . Thus,  $tail = 1/2(tail_5 + tail_{10})$ .

The moment vector  $m = [\bar{A}_t \rho_1^r \text{tail} \rho_1^{ac} \rho_5^{ac} \rho_{10}^{ac} \rho_{25}^{ac} \rho_{50}^{ac} \rho_{100}^{ac}]'$  collects the single mo-

---

<sup>6</sup>For our model it is only necessary to include a one-sided tail index as the simulated models will produce symmetric positive and negative extreme returns (c.f. Franke and Westerhoff 2012).



Table 1: Empirical moments

	Baseline			Full period		
	Corn	Soy	S&P 500	Corn	Soy	S&P 500
$\bar{A}_t$	0.0102	0.0114	0.0077	0.0112	0.0116	0.0080
$\rho_1^r$	0.0503	0.0568	-0.0381	0.0431	0.0515	-0.0580
tail	2.5209	2.8034	2.7516	2.6944	2.9013	2.5272
$\rho_1^{ac}$	0.2239	0.2933	0.1807	0.2248	0.2660	0.2545
$\rho_5^{ac}$	0.2168	0.2755	0.1969	0.2150	0.2506	0.2705
$\rho_{10}^{ac}$	0.1995	0.2512	0.1685	0.1978	0.2309	0.2424
$\rho_{25}^{ac}$	0.1421	0.2163	0.1342	0.1585	0.1948	0.1770
$\rho_{50}^{ac}$	0.0834	0.1495	0.1338	0.1091	0.1365	0.1417
$\rho_{100}^{ac}$	0.0356	0.0846	0.0835	0.0617	0.0790	0.0979

*Notes:* Soybean prices are available from 05/20/1970 and S&P 500 prices from 04/21/1981.

ments. In Table 1, we have included the respective moments for CBOT soybeans and the S&P 500 U.S. equity index next to those for CBOT corn. As indicated by the level of  $\bar{A}_t$ , commodity volatility is higher than that of U.S. equities and is higher for the full compared to the base period. The other stylized facts are similar across markets and time periods.

## 4.2 Objective function

The objective function used to choose the model's parameter values is based on a loss function ( $J$ ) that considers the squared difference between empirical ( $m^{emp}$ ) and simulated moments ( $m^{sim}$ ):

$$J = (m^{sim}(\theta) - m^{emp})'W(m^{sim}(\theta) - m^{emp}), \quad (20)$$

where the vector ( $\theta$ ) collects all model parameters. The optimal parameter vector ( $\theta^*$ ) will minimize  $J$ .  $W$  is a weighting matrix of the deviations between empirical and simulated moments that considers their estimated variance (Winker and Jeleskovic 2007). We follow Winker et al. (2007) and Franke and Westerhoff (2011, 2012) and use a block bootstrap (due to the long-memory property of the data) to estimate the sample covariance matrix of the moments ( $\Sigma$ ), which also holds advantages for later model validation.

First, we block-bootstrap the baseline returns by dividing the  $R_t$  series into blocks of appropriate length. Franke and Westerhoff (2012) propose blocks of 250 days for short-memory moments and 750 days for long-memory moments. We choose to test 250, 500 and 750 day blocks for each moment and then select the bootstrap window length that leads to the smallest deviation between the mean of the bootstrapped moments and the empirical moments, which builds on the window selection procedure in Franke and Westerhoff (2011). Based on these results we select a 250 day window for  $\bar{A}_t$ , a 500 day window for  $\rho_1^r$ ,  $tail$ ,  $\rho_1^{ac}$ ,  $\rho_5^{ac}$ ,  $\rho_{50}^{ac}$ ,  $\rho_{100}^{ac}$  and a 750 day window for  $\rho_{10}^{ac}$ ,  $\rho_{25}^{ac}$ . In dividing the baseline  $R_t$  series by the block size, any residual observations are cut off at the beginning of the series, leading to 9,000 bootstrapped outcomes of  $R_t$  and 36, 18 and 12 blocks for 250, 500 and 750 day windows, respectively. Random block draws with replacement (number of draws = number of blocks) are used to construct a new series of  $R_t$  from which the bootstrapped moment vector  $m^b$  is calculated. This procedure is repeated for  $B = 10,000$  bootstrap samples to obtain  $m_1^b, \dots, m_B^b$ , from which we calculate the vector of moment means  $\bar{m}^b$  and estimate  $\Sigma$  and  $W$  as:

$$\Sigma = \frac{1}{B} \sum_{b=1}^B \left( m^b - \bar{m}^b \right) \left( m^b - \bar{m}^b \right)', \quad (21)$$

$$W = \Sigma^{-1}$$

### 4.3 Parameter estimation

Some parameters are fixed à priori and summarized in the upper part of Table 2. We follow Franke and Westerhoff (2012) and set the reaction coefficient of the market maker to  $\gamma_{MM} = 0.01$  and the intensity of choice to  $\beta = 1$  since both parameters are essentially “scaling parameters” for the price impact and the attractiveness index. The middle part of Table 2 shows the parameters to be estimated. The optimal parameter vector  $\theta^*$ , is derived by minimizing  $J$  subject to a specific simulation period ( $S$ ) and specific random number seeds ( $\nu_t^{CO,S}$ ) that are underlying the calculation of the stochastic volumes ( $\epsilon_t^{CO}, \epsilon_t^S$ ).  $S$  is chosen as  $10 \cdot N$  (Franke and Westerhoff 2012).<sup>7</sup> Prior to the estimation, the random number seed  $\nu_t^{CO,S}$  is drawn from a  $N(0,1)$  distribution and the stochastic volumes are calculated as  $\epsilon_t^{CO,S} = \nu_t^{CO,S} \sigma_{CO,S}$  (Franke 2009). Thus, we can ensure that any variation in the  $J$ -value is exclusively attributable to changes in the parameter vector  $\theta$  and not to the

---

<sup>7</sup>10% of the observations at the beginning of the simulation are discarded to ensure that no transient effects occur (cf. Franke and Westerhoff 2011, 2012).

Table 2: Parameter values

Fixed parameters		Value			
$\gamma_{MM}$	Reaction coefficient of market maker	0.01			
$\beta$	Intensity of choice coefficient	1			
$P_F$	Fundamental value of the commodity	$\ln 100$			
Estimated parameters		Bound	Model 1	Model 2	Model 3
$\gamma_{CO}$	Reaction coefficient of commercial traders	$> 0$	0.327	0.702	0.659
$\gamma_S$	Reaction coefficient of speculators	$> 0$	8.823	52.589	25.062
$\sigma_{CO}^2$	Variance of commercial traders' stochastic volume	$> 0$	19.904	69.459	76.074
$\sigma_S^2$	Variance of speculators' stochastic volume	$> 0$	247.88	2.745	15.770
$\alpha_p$	à priori preferences		-0.174	-0.279	-2.093
$\alpha_h$	Reaction to herding	$> 0$	2.176	1.746	3.884
$\alpha_m$	Reaction to price misalignment	$> 0$	0.834	9.390	12.372
$J$ -value	Value of the objective function		1.978	3.120	2.540

random number seed. Also, we select starting values for the price  $P_t$  ( $P_1 = \ln 99.75$ ,  $P_2 = \ln 100.25$ ) and the attractiveness index  $\alpha_t$  ( $\alpha_1 = 5$ ). The minimization problem is set up as:

$$\min_{\theta} J = (m^{sim}(\theta, S, \nu_t^{CO,S}) - m^{emp})' W (m^{sim}(\theta, S, \nu_t^{CO,S}) - m^{emp}) \quad (22)$$

The optimization implies considerable challenges including that  $m^{sim}$  cannot be expressed analytically as a function of model parameters and data (as in more standard estimation approaches), but requires the full simulation with the HAM and subsequent calculation of moments in each optimization step. The nature of the objective function also leads to multiple local minima (e.g. Gilli and Winker 2003; Franke and Westerhoff 2011). Frequently, a Nelder-Mead Simplex search algorithm (cf. Lagarias et al. 1998) is used to locate suitable minima. The error arising from failing to find the global minimum is considered relatively small and authors suggest to not put too much strain on its determination (Franke and Westerhoff 2011, pp.72-

74). Our focus during the optimization is thus to find a set of parameter values that constitute a local minimum and lead to a good match between  $m^{sim}$  and  $m^{emp}$ . We also choose a direct search approach but combine the Nelder-Mead and a Pattern Search algorithm (cf. Torczon 1997) where we can directly consider the parameter bounds shown in column 3 of Table 2.

First, we draw 15 random starting values for each of the estimable parameters from a uniform distribution within the bounds shown in Table 3. We then start the *patternsearch* solver in MATLAB and run the optimization for each of the 15 vectors of starting values. The solver allows to start the optimization with the built-in Nelder-Mead algorithm. Once we find a local minimum we use these parameter values as starting points for the subsequent pattern search with a large initial mesh (100), an expansion value of 8 and a contraction value of 0.5 and a full poll. For the 3 parameter combinations that lead to the lowest local minimum, we restart the procedure with an even larger initial mesh (10,000), an expansion value of 8, a contraction value of 0.5 and a full poll. We repeat this previous step and decrease the tolerances with respect to the mesh size, the improvement of the objective function and the distance between points chosen during the optimization until there is neither a significant reduction in the  $J$ -value nor a significant change in the parameter size.

The last columns in Table 2 show the three estimated parameters sets that led to the local minima with the smallest  $J$ -values. In model 1, speculators' stochastic volume is associated with the larger variance. In models 2 and 3, the opposite is true. Also, in models 2 and 3, speculators have higher reaction coefficients than in model 1 and there is a strong reaction to perceived price misalignments. The calculated  $J$ -value shown in the last row of Table 2 is contingent on the random seeds  $\nu_t^{CO,S}$  used during the optimization and provides insufficient information to select the best parameter set. We will therefore choose two alternative approaches to assess the quality of the model parameters.

#### 4.4 Evaluation of the model fit

The evaluation of the model fit is first based on how well the combined  $m^{sim}$  obtained with the parameter sets match the combined  $m^{emp}$ , which is accounted for by the  $J$ -value. Second, it checks how well each *single* moment is replicated. We compare  $m^{sim}$  and  $m^{emp}$  via their distributions. Thereby, we use the 10,000 moment vectors ( $m^b$ ) from the block-bootstrapped estimation of  $W$  to calculate a bootstrapped distribution for the loss functions  $J^b = (m^b - m^{emp})'W(m^b - m^{emp})$ , for  $b = 1, \dots, 10,000$

Table 3: Boundaries for sample of starting values

Parameter	Lower bound	Upper bound
$\gamma_{CO}$	1	10
$\gamma_S$	1	20
$\sigma_{CO}^2$	1	25
$\sigma_S^2$	1	25
$\alpha_p$	-5	5
$\alpha_h$	1	10
$\alpha_m$	1	10

and for each single moment within  $m^b$ . We interpret the distribution of  $J^b$  and the moments within  $m^b$  as an approximation of the true distribution of  $J$  and each empirical moment that would arise if the actual data generation process (DGP) behind the empirically observed return series were to be repeated 10,000 times to create 10,000 different return series (Franke and Westerhoff 2012). The comparison of the simulated and empirical (bootstrapped)  $J$  values and moments first uses the “p-values” from Franke and Westerhoff (2012), which we label “percentage coverage”<sup>8</sup> and second calculates the relative entropy, i.e. Kullback-Leibler (KL) divergence, for the respective distributions.

#### 4.4.1 Percentage coverages

We calculate a total and a moment-specific percentage coverage. If we use the distribution of  $J^b$  as an approximation of the true  $J$  distribution, then any simulation run using the optimal parameter vector  $\theta^*$  cannot be rejected as not being consistent with the actual DGP if the obtained value of  $J^{sim}$  falls within the 95% quantile of the distribution of  $J^b$  and thus below a critical value ( $J_{0.95}^b$ ) (Franke and Westerhoff 2012). The total percentage coverage calculates the percentage out of 10,000 simulation runs using  $\theta^*$  and a simulation period  $T=9,000$ <sup>9</sup> that lead to  $J^{sim} < J_{0.95}^b$ . We use the “p-value” from the “DCA-HPM” model in Franke and Westerhoff (2012) of 32.6% as a benchmark, as suggested by the authors. Results for the  $\theta^*$  parameter sets from the three different base scenario models are shown in Table 4. Models 1 and 2 both exceed the chosen benchmark value of 32.6%.

<sup>8</sup>We use this alternative term to illustrate that we actually calculate a percentage of observations that fit within a predetermined value bound rather than perform a formal statistical test.

<sup>9</sup>This corresponds to the length of the bootstrap return sample used to calculate  $m^b$  and ensures that the return series underlying the calculation of  $m^{sim}$  and  $m^b$  have the same length.

Table 4: Critical value, percentage coverages and benchmark value

$J_{0.95}^b$	Model 1	Model 2	Model 3	Benchmark
19.68	39.23	<b>42.74</b>	28.51	32.6

Table 5: Percentage coverages for single moments

	Empirical estimate	Critical values		Calculated p-values		
	$m^{emp}$	$m_{0.95}^b$	$m_{0.05}^b$	Model 1	Model 2	Model 3
$\bar{A}_t$	0.0102	0.0111	0.0095	76.27	<b>77.67</b>	74.81
$\rho_1^r$	0.0503	0.0696	0.0283	<b>72.01</b>	68.23	66.14
tail	2.5209	2.9897	2.3242	<b>90.70</b>	85.45	97.93
$\rho_1^{ac}$	0.2239	0.2839	0.1440	97.84	<b>97.33</b>	98.67
$\rho_5^{ac}$	0.2168	0.2693	0.1443	98.20	98.12	<b>97.52</b>
$\rho_{10}^{ac}$	0.1995	0.2603	0.1130	99.46	99.48	<b>98.81</b>
$\rho_{25}^{ac}$	0.1421	0.1934	0.0665	98.35	94.61	<b>91.53</b>
$\rho_{50}^{ac}$	0.0834	0.1207	0.0286	<b>86.22</b>	75.06	68.72
$\rho_{100}^{ac}$	0.0356	0.0658	-0.0170	<b>95.69</b>	68.98	63.41

While it permits an overall assessment of the model fit, the total percentage coverage does not measure how well the *single* moments are replicated. We calculate moment-specific percentage coverages by comparing the marginal moment distributions, which may have different left and right bounds. Thus we calculate both a critical 95th and 5th percentile moment value ( $m_{0.95}^b, m_{0.05}^b$ ). The moment-specific percentage coverage calculates the percentage out of 10,000 simulation runs using  $\theta^*$  and a simulation period  $T=9,000$  that leads to  $m_{0.05,i}^b < m_i^{sim} < m_{0.95,i}^b$ . We summarize the results in Table 5 where the best fit (value closest to 90%) is indicated in bold font. Model 1 is best at matching the fatness of the tail, the zero autocorrelation property of the raw returns ( $\rho_1^r$ ) and the long memory property of the data approximated with the autocorrelations of the absolute returns<sup>10</sup> while model 2 is best at replicating the overall volatility level ( $\bar{A}_t$ ). Before we decide on a parameter set, we first use the KL divergence to compare the distance between the sampling distributions of the simulated and bootstrapped  $J$  and moment values.

<sup>10</sup>The differences to model 2 and 3 coverages for  $\rho_1^{ac}, \rho_5^{ac}, \rho_{10}^{ac}$  are generally small but the coverage of  $\rho_{50}^{ac}, \rho_{100}^{ac}$  is best.

#### 4.4.2 Relative entropy

Compared to the percentage coverage, which only considers how many of the simulated values fit between the pre-defined bounds, the KL divergence can better account for differences in the distributional shape, e.g. with respect to skewness or kurtosis. It is calculated as:

$$\begin{aligned} d_{KL,J} &= \sum_{i=1}^d P_i(J) \ln \frac{P_i(J)}{Q_i(J)}, \\ d_{KL,m} &= \sum_{i=1}^d P_i(m) \ln \frac{P_i(m)}{Q_i(m)}, \end{aligned} \tag{23}$$

where  $P(J), P(m)$  are the probability density functions of the 10,000 bootstrapped  $J$ -values and moments and  $Q(J), Q(m)$  are the probability density functions of 10,000 simulated  $J$ -values and moments using the optimal parameter vector  $\theta^*$  and a simulation period  $T$ , as described above.  $d = 100$  is the number of bins in the histograms underlying the two probability distributions. In the case of two identical distributions,  $d_{KL} = 0$ , thus, the lower the calculated value, the lower the distance between the bootstrapped and simulated distributions. The results are shown in Table 6 where again the best distributional fit is indicated with bold font. The overall distance is lowest for model 2. For the single moments, model 1 best replicates the distribution of the fatness of the tail. Model 2 distances are lowest for  $\bar{A}_t, \rho_1^r, \rho_{10}^{ac}$  and thus the model performs best at replicating the volatility level and the zero autocorrelation property of the raw returns. Finally, model 3 shows the lowest distance for  $\rho^{ac}$  at all remaining lags and is thus best at matching the decaying ACF and the long memory effects in the return data. Nevertheless, distances for model 2 are in most cases not drastically different from model 3.

In summary, model 2 had the best overall percentage coverage and the lowest distributional distance between  $J^b$  and  $J^{sim}$ . Its single moment matching is comparable to model 3, which is why we choose model 2 as our main base scenario parameter set. Nevertheless, we found that the parameter values in model 1 were quite different concerning the reaction to price misalignments, the speculators' reaction coefficients and variances of the traders' stochastic volumes. Since its total percentage coverage still exceeds the benchmark of 32.6 and the individual moment matching is quite good, especially for the tail index, we will use model 1 as a comparative base scenario model. This may provide additional insights about the drivers

Table 6: KL divergence for moment and loss function distributions

	Model 1	Model 2	Model 3
$J^{sim}$	0.291	<b>0.105</b>	0.114
$\bar{A}_t$	0.546	<b>0.011</b>	0.053
$\rho_1^r$	0.177	<b>0.071</b>	0.223
tail	<b>0.687</b>	1.169	1.152
$\rho_1^{ac}$	0.251	0.189	<b>0.164</b>
$\rho_5^{ac}$	0.356	0.216	<b>0.193</b>
$\rho_{10}^{ac}$	0.299	<b>0.147</b>	0.234
$\rho_{25}^{ac}$	0.265	0.226	<b>0.166</b>
$\rho_{50}^{ac}$	0.257	0.197	<b>0.118</b>
$\rho_{100}^{ac}$	0.150	0.076	<b>0.056</b>

behind price dynamics and allows for some sensitivity analysis regarding the main results.

## 5 Scenario comparison

We simulate first the base scenario over a period of 9,085 trading days, equivalent to the number of observations in the base period data sample with the model 2 and 1 parameters from Table 2. For both parameter sets we use identical random number seeds ( $\nu_t^{CO,S}$ ) such that any result differences are solely attributable to the parameter values. In the later financialization scenario we fix the parameter values à priori in order to better understand sensitivities in the price dynamics with respect to parameter changes.

### 5.1 Base scenario results

A graphical summary of base scenario results for model 2 and model 1 is provided in Figure 3. In both models traders have a very small à priori preference for a chartist strategy. And, the reaction to herding incentives of the traders is of a similar magnitude ( $\alpha_{h,2} = 1.746$  vs.  $\alpha_{h,1} = 2.176$ ). But, the response of traders to a price misalignment between the current and the fundamental price is much stronger in model 2 ( $\alpha_{m,2} = 9.39$ ) than in model 1 ( $\alpha_{m,1} = 0.834$ ). Deviations from the fundamental price (horizontal line) increase the attractiveness of fundamentalism and decrease chartists' market weight. Due to the size of  $\alpha_m$ , the maximum level of  $\alpha_t$  in model 2 is higher than in model 1. Also, the stronger reaction to price misalign-



ment paired with a high reaction coefficient of speculators ( $\gamma_{S,2} = 52.589$  versus  $\gamma_{S,1} = 8.823$ ) in model 2 leads to higher fluctuations in the chartist weight (variance of 0.1 for model 2 and 0.05 for model 1) and deviations from the fundamental price are less persistent than in model 1. The mean price level is almost identical for both parameter sets with  $\bar{P}_1 = 4.67$  and  $\bar{P}_2 = 4.65$ .

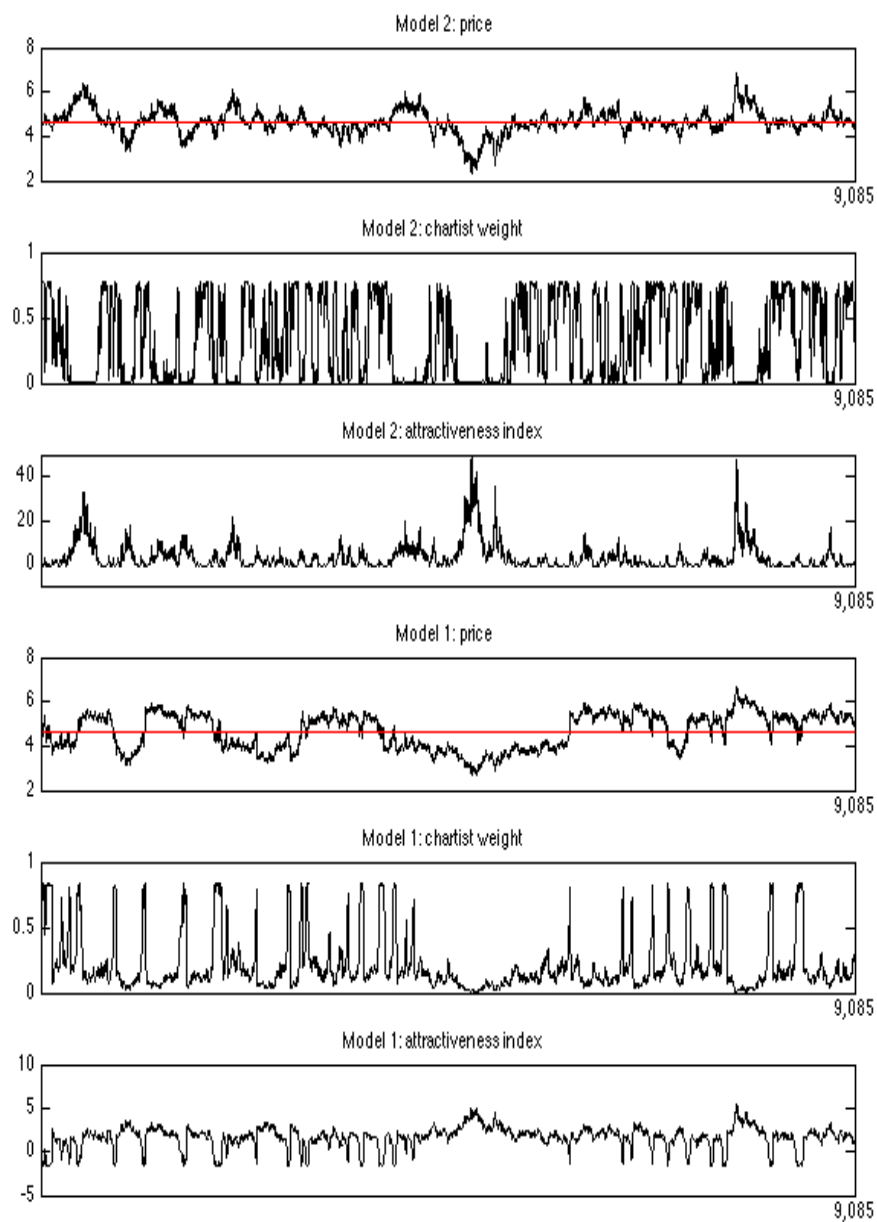
The price volatility effect, measured with  $R_t^2$ , is presented in Figure 4. In model 1 an increase in the chartist (speculator) weight and thus a higher chartist trading volume with a large variance in the stochastic component, leads to an increase in the level of  $R_t^2$ . In model 2 a higher fundamentalist (commercial trader) weight and thus larger fundamentalist trading volume (including the stochastic volume with its high variance) is associated with a high level of  $R_t^2$ . Thus, the trader group with the highest variance in the stochastic trading volume will carry the largest information shocks to the market and increase the short-term volatility. From a theoretical perspective, none of the two scenarios seems implausible. The forecast of the fundamental price of corn is associated with uncertainty, which could justify a fundamentalist-chartist variance relation as in model 2. On the other hand, the trading strategies within the group of speculators may be much more diverse than within the group of commercial traders, which could motivate the model 1 variance relation.

## 5.2 Financialization scenario results

The parameters for the financialization scenario that are set *in addition* to the base scenario parameters are summarized in Table 7. The fixed parameters, starting values and random number seeds are the same as in the base scenario and we use an identical simulation period length. We can think of it as a period of another 9,085 days that starts after the introduction of index funds and emergence of portfolio managers on the market. Base scenario reaction coefficients, variances of stochastic volumes and the coefficients in the attractiveness index are assumed to be unaffected by the market entry of portfolio managers.

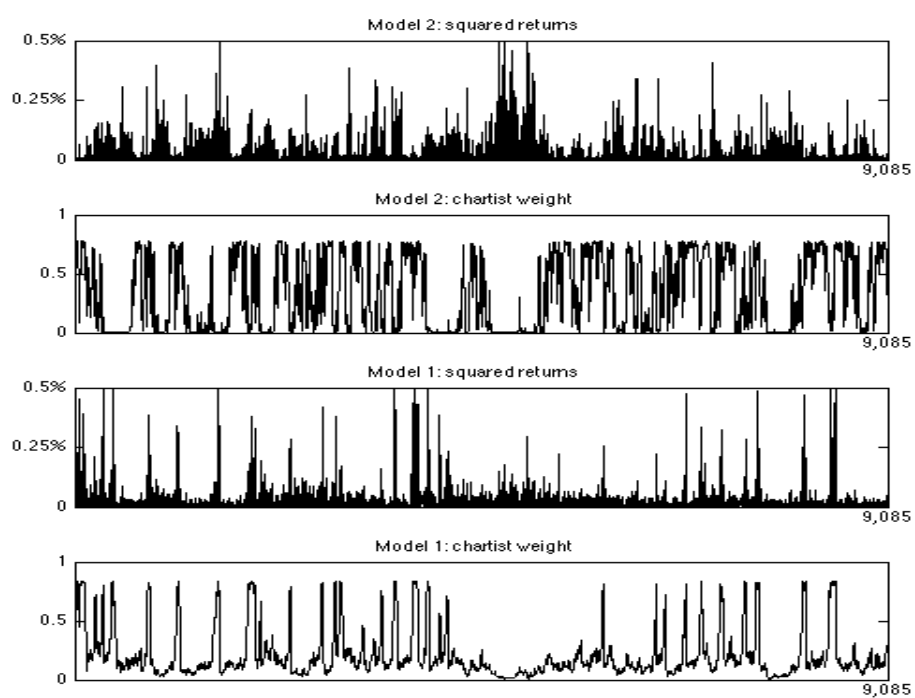
We define three parameter scenarios. Scenario 1 models a “high impact” situation with a relatively strong reaction coefficient for the portfolio managers and high variance in their stochastic volume. Scenario 2 is the “low impact” scenario where both the reaction coefficient and the variance of the stochastic volume are significantly reduced. Scenario 3 shows a situation of a “fast reaction” where portfolio

Figure 3: Base results overview



*Notes:* Horizontal (red) line in price charts represents the constant fundamental price, black lines represent base scenario results for model 1 and 2 parameter sets.

Figure 4: Base results volatility effect



*Notes:* Black lines represent base scenario results for model 1 and 2 parameter sets.

Table 7: Financialization scenarios

Parameter	Scenarios		
	(1) “High impact”	(2) “Low impact”	(3) “Fast reaction”
$\gamma_{PM}$	5	0.5	10
$\sigma_{PM}^2$	50	5	5
$\chi$	0.2	0.2	0.2

managers’ stochastic volume is still associated with a low variance but the reaction coefficient is twice as high as in scenario 1. Minimum position holdings ( $\chi$ ) are always set to 20% of the current total long position. We first show financialization scenario 1 results for a base scenario with model 2 parameters. There are only few differences for model 1 parameters and we only mention those that are significant or provide additional insights. The full set of results for the model 1 parameters are available from the authors upon request. In the following figures, black bars and lines will represent the base scenario and grey lines the financialization scenario.

### 5.2.1 Volume effect

The creation of index funds and the market entry of portfolio managers increases overall trading volume ( $V_t$ ) on the market, paralleling the empirically observed volume increase for CBOT corn futures (see Figure 2). The absolute position size associable with fundamentalist and chartist trading strategies is calculated as:

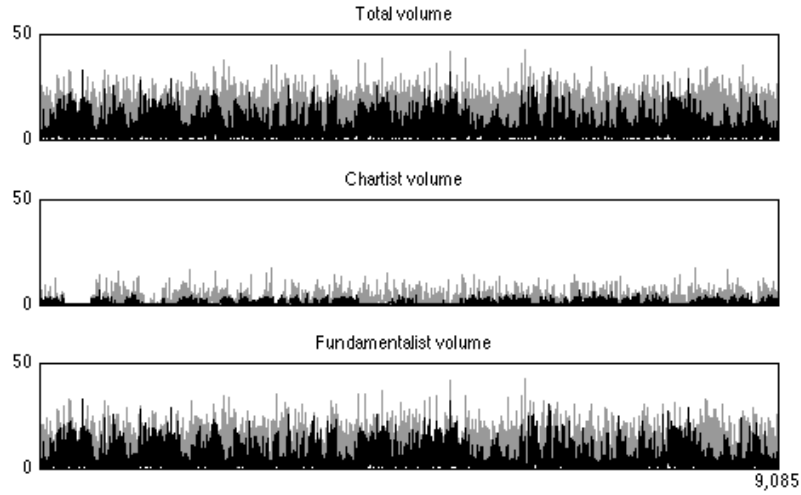
$$|V_t^{F,C}| = \phi_t^{F,C} |V_t^{CO,S}|, \quad (24)$$

$$|V_t| = |V_t^F| + |V_t^C|, \quad (25)$$

and shown in Figure 5. In the base scenario, the mean overall trading volume is 5.1 while in the financialization scenario it is 11.4. In the base scenario, the higher variance in the stochastic commercial trading volume contributes to a higher fundamentalist trading volume while the new portfolio managers’ fundamentalist and chartist volume is associated with the same stochastic variance ( $\sigma_{PM}^2 = 50$ ).

The total trading position for each trader group sums up the position holdings in each period  $t$  over the full simulation period. A positive (negative) total position holding equals a total net long (short) position. The total position development is shown in Figure 6. In the base scenario both commercial traders and speculators

Figure 5: Volume effect (Model 2, Scenario 1)



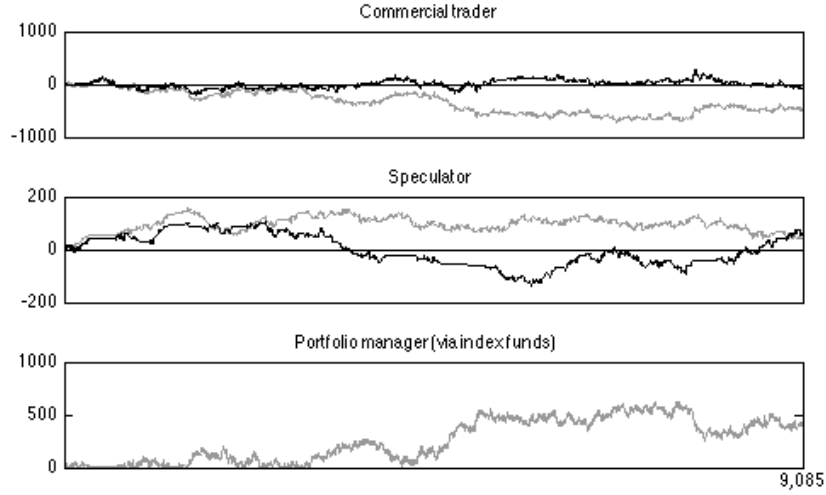
*Notes:* Black bars represent base scenario results with model 2 parameters, grey bars financialization scenario 1 results.

either take a net long or a net short position, which can change over the course of the simulation period, depending on the price dynamics. The new portfolio managers' volume is restricted to a net long position that can at maximum be reduced up to a percentage level determined by  $\chi$ . In the first few periods, the short-selling constraint is frequently binding but is later without effect. In the financialization scenario the commercial traders' position switches to net short while speculators are predominantly net long. In model 1, both commercial traders and speculators' positions are mostly net short. In any case the additional net long position of the portfolio managers leads to a change of net positions of the other trader groups. This is possible because the existing traders are as a group not limited in their possibilities to either take net short or long positions and switch between them as desired.

### 5.2.2 Price effect

The price dynamics from the base and financialization scenario are shown in Figure 7. The additional portfolio manager volume does not inflate the price level but rather has the opposite effect. Apart from the first few periods when portfolio managers trading volume is very low (and the short-selling constraint is binding), the price

Figure 6: Position holdings (Model 2, Scenario 1)



*Notes:* Black lines represent base scenario results with model 2 parameters, grey lines financialization scenario 1 results.

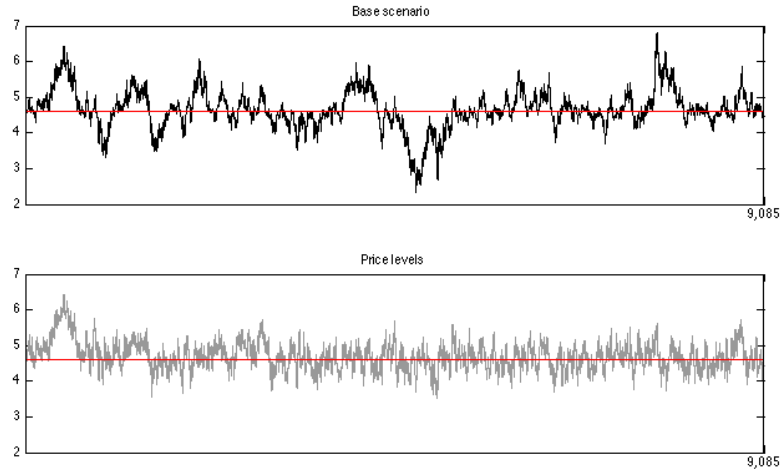
dynamics fluctuate closer around the fundamental value and there is less tendency for prices to misalign compared to the base scenario. For model 1 parameters, we obtain the same general results but due to the lower value of the price misalignment coefficient  $\alpha_m$ , prices can deviate further away from the fundamental value.

To investigate the volatility effect we use squared returns and the 30-day ( $Vol(30)$ ) and 90-day ( $Vol(90)$ ) return-based volatility, which are calculated as:

$$Vol(m) = \sqrt{\frac{1}{m-1} \sum_{n=1}^m (R_{t-n} - \bar{R}(m))^2}, \quad m = 30, 90 \quad (26)$$

The stochastic portfolio managers' volume in Scenario 1 is associated with a variance of  $\sigma_{PM}^2 = 50$ . The SSV model setup implies that the variance inflates the time-dependent variance of the market price (see equation (16)). We have interpreted the stochastic volume as a representing portfolio allocations due to correlations with other assets and as unrelated to stochastic volume from the other traders. If the variance of this volume is high then it significantly increases volatility levels in commodity markets by transmitting new information shocks. With model 1 parameters, the volatility increase is even stronger due to the above mentioned

Figure 7: Price level effect (Model 2, Scenario 1)



*Notes:* Horizontal (red) line in price charts represents the constant fundamental price, black lines represent base scenario results with model 2 parameters, grey lines financialization scenario 1 results.

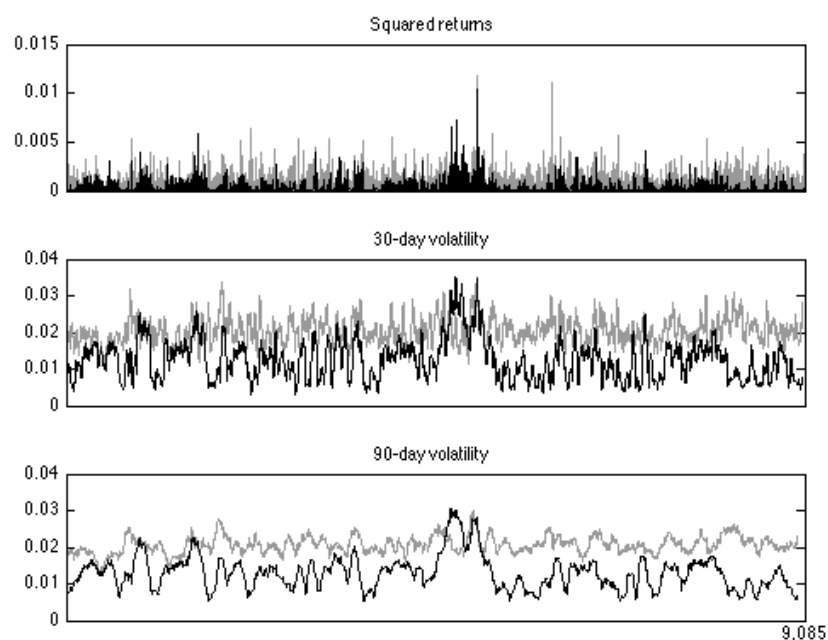
larger magnitude of price deviations away from fundamentals.

### 5.2.3 Effects of parameter changes

Figure 9 shows the price dynamics for the three different financialization scenarios. Comparing the outcome of the “high impact” scenario with the “low impact” and “fast reaction” scenarios, it becomes clear that the observed overall lower price levels and lower likelihood of a price misalignment (or bubble) in the financialization scenario are a result of the size of the portfolio managers reaction coefficient. Not only the commercial traders and speculators respond to price misalignment by entering or leaving the market but also the portfolio managers by readjusting the weights of their trading strategies. A high reaction coefficient entails a fast reaction to any perceived price misalignments and a market price that will fluctuate closer to the fundamental value. Unsurprisingly, for the model 1 parameter set, where the reaction to price misalignment is much lower, the effect is also visible but much less pronounced.

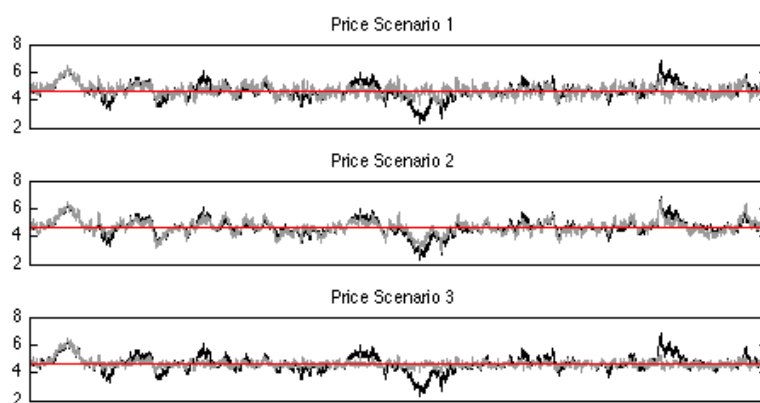
Figure 10 shows the short-term volatility effect. The strongest increase in overall price volatility is brought about by the “high impact” and the lowest increase by

Figure 8: Volatility effect (Model 2, Scenario 1)



*Notes:* Black lines represent base scenario results with model 2 parameters, grey lines financialization scenario 1 results.

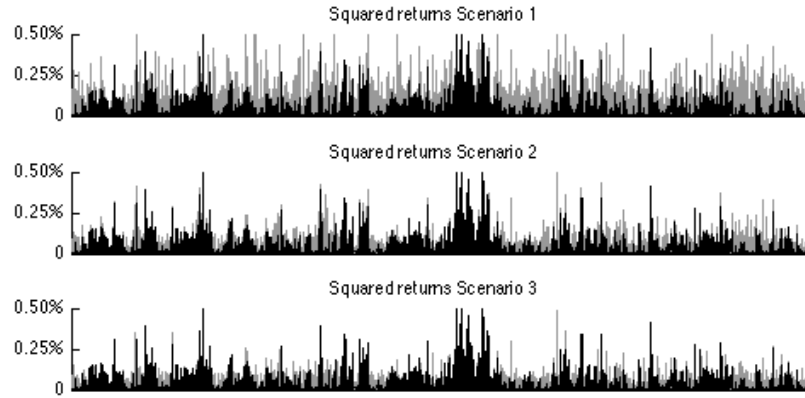
Figure 9: Price levels under different financialization scenarios



*Notes:* Horizontal (red) line represents the constant fundamental price, black lines represent base scenario results with model 2 parameters, grey lines financialization results for different scenarios.



Figure 10: Volatility levels under different financialization scenarios



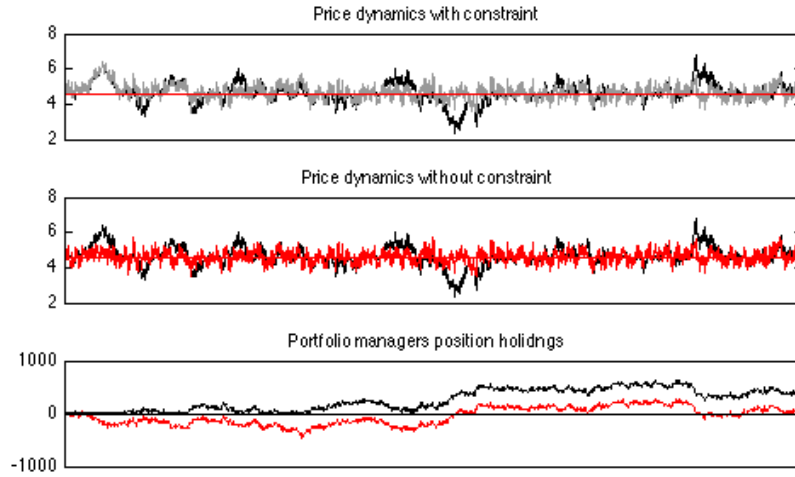
*Notes:* Black lines represent base scenario results with model 2 parameters, grey lines financialization results for different scenarios.

the “fast reaction” scenario. In the base scenario, the mean of  $R_t^2$  is 0.019% while in Scenario 1 it is 0.043%, in Scenario 2 0.026% and in Scenario 3 it is only 0.020% and thus relatively close to the base scenario mean. It is clear that a higher variance in stochastic portfolio managers’ volume entails a stronger increase in overall price volatility. Thus, in times of market crises that affect asset correlations, the volatility increase on the commodity market could be more pronounced than in tranquil periods. For a given level of stochastic variance, a higher reaction coefficient for the deterministic volume will decrease the volatility level and dampen spikes.

#### 5.2.4 Removal of the short-selling constraint

Finally we investigate the effect of a removed short-selling constraint within scenario 1. The unrestricted trading volume of portfolio managers is determined according to equation (11). The results are shown in Figure 11. In the first few periods of a binding constraint its removal leads to a net short position of the portfolio managers (third graph). While the short-selling constrained new trading volume did not lead to an inflation of prices above the base scenario levels, price levels are even lower once it is removed, which can be seen from the two price graphs in Figure 11.

Figure 11: Effects without short-selling constraint



*Notes:* Black lines represent base scenario results with model 2 parameters, grey lines financialization scenario 1 results, red lines are financialization scenario 1 results without implementation of the short-selling constraint.

## 6 Conclusions

To investigate price effects on agricultural commodity markets from portfolio inclusion of index funds, we employ a few-type HAM with a SSV approach (Franke and Westerhoff 2012) that depicts the price dynamics in the corn futures market populated by fundamentalist commercial traders and chartist speculators. We thereby extend the hitherto econometrics centered analysis on financialization effects with a simulation model approach that allows to directly consider price level and volatility effects of specific trading strategies. Our base scenario parameters are estimated from daily corn futures returns over the period 01/05/1970-12/31/2005 with the MSM. The selected moments capture the overall volatility level, zero autocorrelation of returns, long-memory effects and fat-tailed return distributions. Parameters are validated based on their performance in joint and single moment matching. We thereby extend previous approaches by looking at the whole moment distribution. In our financialization scenario, we model the situation after the year 2005. The increased availability of commodity index funds facilitates market entry of financial portfolio managers who use commodities as portfolio diversifiers and purchase index fund shares rather than single futures contracts. Thereby, portfolio man-

agers' demand depends on individual commodity returns, evaluated with a mixed fundamentalist-chartist strategy, and on return or volatility correlations with other portfolio assets, modeled as a stochastic demand component.

In the base scenario, we compare results from two parameter sets and demonstrate that the trader group with the highest variance in the stochastic volume carries the largest information shocks to the market and thus directly increases volatility levels. The price level, on the other hand, is most strongly affected by how fast traders respond to changes in the factors that affect their deterministic volume and by the traders' reaction to price misalignment on the market. Thereby, higher reaction coefficients decrease the persistence of price deviations and move prices closer to their fundamental value.

In the financialization scenario, portfolio managers' trading via index funds creates new long-only trading volume from the funds' index replication activities. But, price levels are not inflated but rather fluctuate more closely around the fundamental value when the deterministic demand of portfolio managers reacts to price misalignments and herding tendencies. Given these model assumptions, the Masters' hypothesis of index funds replication volume creating price bubbles on the market cannot be confirmed. A removal of the short-selling constraint would even further reduce the occurrence of price deviations. In contrast, the volatility effect is more pronounced. The information shocks created by the stochastic portfolio managers volume that are assumed to be linked to correlations with other asset markets directly increase volatility levels. The higher the variance of these demand or supply shocks, e.g. in times of financial crises, the larger the volatility increase. The transmission of information shocks affecting volatility is most closely related to the argument in Irwin et al. (2009) where new volume would affect prices if it transports new information to the market. In our model, index fund replication volume may thus increase price volatility but decreases price levels.

Future research could, on the one hand, focus on modifying the model design and addressing some current limitations. Liquidity constraints for the group of commercial traders and speculators or specific position requirements due to hedging of primary business activities could influence the price level effect. Time-varying correlations between commodities and other financial assets could be modeled more explicitly within a multiple market setup and also consider crisis effects. And, spot and futures markets could be linked via the fundamental value of the commodity.

Finally, the model could be used for the analysis of regulatory proposals such as transaction taxes or price limits. On the other hand, model estimation and validation also hold potential for future research, e.g. by considering different random number seeds already during the minimization rather than in an ex-post validation and extending the current methods used for parameter validation.

### ***Acknowledgements***

The authors thank the participants of the IATRC Annual General Meeting at Clearwater Beach, Florida, December 15-17, 2013 for helpful comments and discussions.

## References

- Alfarano, S., T. Lux, and F. Wagner (2005). Estimation of Agent-Based Models: The Case of an Asymmetric Herding Model. *Computational Economics* 26(1), 19–49.
- Ankrum, E. M. and C. R. Hensel (1993). Commodities in Asset Allocation: A Real-Asset Alternative to Real Estate? *Financial Analysts Journal* 49(3), 20–29.
- Anson, M. J. (1999). Maximizing Utility with Commodity Futures Diversification. *The Journal of Portfolio Management* 25(4), 86–94.
- Anufriev, M. and J. Tuinstra (2013). The impact of short-selling constraints on financial market stability in a heterogeneous agents model. *Journal of Economic Dynamics and Control* 37(8), 1523–1543.
- Bauer, C., P. De Grauwe, and S. Reitz (2009). Exchange rate dynamics in a target zone—A heterogeneous expectations approach. *Journal of Economic Dynamics and Control* 33(2), 329–344.
- Beja, A. and M. B. Goldman (1980). On The Dynamic Behavior of Prices in Disequilibrium. *The Journal of Finance* 35(2), 235–248.
- BlackRock (2011). ETF Landscape: Global Handbook. Q1 2011.
- Brock, W. A. and C. H. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple asset price model. *Journal of Economic Dynamics and Control* 22, 1235–1274.
- CFTC (2014a). Commitments of Traders (COT) Report – Explanatory Notes.
- CFTC (2014b). Disaggregated Commitments of Traders Report – Explanatory Notes.
- Chiarella, C. (1992). The dynamics of speculative behaviour. *Annals of Operations Research* 37(1), 101–123.
- Chiarella, C., R. Dieci, and L. Gardini (2002). Speculative behaviour and complex asset price dynamics: a global analysis. *Journal of Economic Behavior & Organization* 49(2), 173–197.
- Chiarella, C., R. Dieci, and L. Gardini (2005). The Dynamic Interaction of Speculation and Diversification. *Applied Mathematical Finance* 12(1), 17–52.
- Chiarella, C., R. Dieci, and X.-Z. He (2007). Heterogeneous expectations and speculative behavior in a dynamic multi-asset framework. *Journal of Economic Behavior & Organization* 62(3), 408–427.
- Diebold, F. X. and K. Yilmaz (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28(1), 57–66.

- Dieci, R. and F. Westerhoff (2010). Heterogeneous speculators, endogeneous fluctuations and interacting markets: A model of stock prices and exchange rates. *Journal of Economic Dynamics and Control* 34, 743–764.
- Duffie, D. and K. J. Singleton (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica* 61, 929–952.
- Ellen, S. and R. C. Zwinkels (2010). Oil price dynamics: A behavioral finance approach with heterogeneous agents. *Energy Economics* 32(6), 1427–1434.
- Etienne, X., S. H. Irwin, and P. Garcia (2014). Bubbles in food commodity markets: Four decades of evidence. *Journal of International Money and Finance* 42, 129–155.
- Franke, R. (2009). Applying the method of simulated moments to estimate a small agent-based asset pricing model. *Journal of Empirical Finance* 16(5), 804–815.
- Franke, R. and F. Westerhoff (2011). Estimation of a Structural Stochastic Volatility Model of Asset Pricing. *Computational Economics* 38, 53–83.
- Franke, R. and F. Westerhoff (2012). Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control* 36(8), 1193–1211.
- Frankel, J. A. and K. A. Froot (1990). Chartists, Fundamentalists, and Trading in the Foreign Exchange Market. *The American Economic Review* 80(2), 181–185.
- Gao, L. and L. Liu (2014). The Volatility Behavior and Dependence Structure of Commodity Futures and Stocks. *Journal of Futures Markets* 34(1), 93–101.
- Gilbert, C. L. (2010). How to Understand High Food Prices. *Journal of Agricultural Economics* 61(2), 398–425.
- Gilbert, C. L. and S. Pfuderer (2014). The Role of Index Trading in Price Formation in the Grains and Oilseeds Markets. *Journal of Agricultural Economics* 65(2), 303–322.
- Gilli, M. and P. Winker (2003). A global optimization heuristic for estimating agent based models. *Computational Statistics and Data Analysis* 42, 299–312.
- Gorton, G. and K. G. Rouwenhorst (2006). Facts and Fantasies about Commodity Futures. *Financial Analysts Journal* 62(2), 47–68.
- Grosche, S. C. (2014). What Does Granger Causality Prove? A Critical Examination of the Interpretation of Granger Causality Results on Price Effects of Index Trading in Agricultural Commodity Markets. *Journal of Agricultural Economics* 65(2), 279–302.
- Grosche, S. C. and T. Heckelee (2014). Directional Volatility Spillovers between Agricultural, Crude Oil, Real Estate and other Financial Markets. ILR Discussion Paper 2014:4.

- Gutierrez, L. (2013). Speculative bubbles in agricultural commodity markets. *European Review of Agricultural Economics* 40(2), 217–238.
- He, X.-Z. and F. H. Westerhoff (2005). Commodity markets, price limiters and speculative price dynamics. *Journal of Economic Dynamics and Control* 29, 1577–1596.
- Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics* 3(5), 1163–1174.
- Hommes, C. H. (2006). Heterogeneous agent models in economics and finance. In L. Tesfatsion and K. L. Judd (Eds.), *Agent-based computational economics*, Volume 13,2 of *Handbooks in Economics*, pp. 1110–1186. Amsterdam: Elsevier/North-Holland.
- Irwin, S. H. and D. R. Sanders (2012). Testing the Masters Hypothesis in commodity futures markets. *Energy Economics* 34(1), 256–269.
- Irwin, S. H., D. R. Sanders, and R. P. Merrin (2009). Devil or Angel? The role of speculation in the recent commodity price boom (and bust). *Journal of Agricultural and Applied Economics* 41(2), 377–391.
- Ji, Q. and Y. Fan (2012). How does oil price volatility affect non-energy commodity markets. *Applied Energy* 89, 273–280.
- Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright (1998). Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization* 9(1), 112–147.
- Lee, B.-S. and B. F. Ingram (1991). Simulation estimation of time-series models. *Journal of Econometrics* 47(2-3), 197–205.
- Liu, X., G. Filler, and M. Odening (2013). Testing for speculative bubbles in agricultural commodity prices: a regime-switching approach. *Agricultural Finance Review* 73, 179–200.
- Lux, T. (1997). Time variation of second moments from a noise trader/infection model. *Journal of Economic Dynamics and Control* 22(1), 1–38.
- Manzan, S. and F. Westerhoff (2005). Representativeness of news and exchange rate dynamics. *Journal of Economic Dynamics and Control* 29(4), 677–689.
- Mensi, W., M. Beljid, A. Boubaker, and S. Managi (2013). Correlations and volatility spillovers across commodity and stock markets: Linking energies, food and gold. *Economic Modelling* 32, 15–22.
- Redrado, M., J. Carera, D. Bastourre, and J. Ibarluća (2009). Financialization of Commodity Markets: Non-linear Consequences from Heterogenous Agent Behavior: Banco Central de la República Argentina. Investigaciones Económicas. Working Paper 2009/44.

- Reitz, S. and F. Westerhoff (2007). Commodity price cycles and heterogeneous speculators: a STAR–GARCH model. *Empirical Economics* 33(2), 231–244.
- Robles, M., M. Torero, and J. v. Braun (2009). When speculation matters: International Food Policy Research Institute, Issue Brief 57.
- Sanders, D. R. and S. H. Irwin (2011a). New Evidence on the Impact of Index Funds in U.S. Grain Futures Markets. *Canadian Journal of Agricultural Economics* 59(4), 519–532.
- Sanders, D. R. and S. H. Irwin (2011b). The Impact of Index Funds in Commodity Futures Markets: A Systems Approach. *The Journal of Alternative Investments* 14(1), 40–49.
- Silvennoinen, A. and S. Thorp (2013). Financialization, Crisis and Commodity Correlation Dynamics. *Journal of International Financial Markets, Institutions & Money* 24, 42–65.
- S&P Dow Jones Indices (2014). S&P GSCI Methodology.
- Stoll, H. R. and R. E. Whaley (2010). Commodity index investing and commodity futures prices. *Journal of Applied Finance* (1), 1–40.
- Torczon, V. (1997). On the Convergence of Pattern Search Algorithms. *SIAM Journal of Optimization* 7, 1–25.
- Westerhoff, F. (2003). Speculative markets and the effectiveness of price limits. *Journal of Economic Dynamics and Control* 28, 493–508.
- Westerhoff, F. (2012). Interactions between the Real Economy and the Stock Market: A Simple Agent-Based Approach. *Discrete Dynamics in Nature and Society* 2012, 1–21.
- Westerhoff, F. and S. Reitz (2005). Commodity price dynamics and the nonlinear market impact of technical traders: empirical evidence for the US corn market. *Physica A* 349, 641–648.
- Westerhoff, F. H. and R. Dieci (2006). The effectiveness of Keynes–Tobin transaction taxes when heterogeneous agents can trade in different markets: A behavioral finance approach. *Journal of Economic Dynamics and Control* 30(2), 293–322.
- Will, M. G., S. Prehn, I. Pies, and T. Glauben (2012). Is financial speculation with agricultural commodities harmful or helpful?—A literature review of current empirical research. Discussion Paper No. 2012-27 of the Chair in Economics, Ethics, Martin-Luther-University Halle-Wittenberg.
- Winker, P., M. Gilli, and V. Jeleskovic (2007). An objective function for simulation based inference on exchange rate data. *Journal of Economic Interaction and Coordination* 2(2), 125–145.



Winker, P. and V. Jeleskovic (2007). Dependence of—and Long Memory in—Exchange Rate Returns: Statistics, Robustness, Time Aggregation. University of Essex, Centre for Computational Finance and Economic Agents, Working Paper WP011-07.

Zeeman, E. C. (1974). On the unstable behaviour of stock exchanges. *Journal of Mathematical Economics* 1(1), 39–49.