



**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.*

# **Analyzing Developing Country Market Integration with Incomplete Price Data Using Cluster Analysis**

**Isaac Gershon Ansah, Cornelis Gardebroek, Rico Ihle, and Moti Jaleta**

*Selected Paper prepared for presentation at the International Agricultural Trade Research Consortium's (IATRC's) 2014 Annual Meeting: Food, Resources and Conflict, December 7-9, 2014, San Diego, CA.*

*Copyright 2014 by Isaac Gershon Ansah, Cornelis Gardebroek, Rico Ihle, and Moti Jaleta. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*

# Analyzing developing country market integration using incomplete price data and cluster analysis

*Paper presented at the 2014 IATRC Annual meeting, San Diego, December 7-9, 2014*

Isaac Gershon Ansah<sup>a</sup>, Cornelis Gardebroek<sup>b</sup>, Rico Ihle<sup>b</sup>, and Moti Jaleta<sup>c</sup>

<sup>a</sup> Dept. of Agric. & Res. Econ., Univ. for Development Studies (UDS), P. O. Box TL 1882, Tamale, Ghana

<sup>b</sup> Agric. Econ. and Rur. Pol., Wageningen University, Hollandseweg 1, 670KN Wageningen, The Netherlands

<sup>c</sup> CIMMYT, PO Box 5689, Addis Ababa, Ethiopia

## *Abstract*

*Recent global food price developments have spurred renewed interest in analyzing integration of local markets to global markets. A popular approach to quantify market integration is cointegration analysis. However, local market price data often has missing values, outliers, or short and incomplete series, making cointegration analysis impossible. Instead of imputing missing data, this paper proposes cluster analysis as an alternative methodological approach for analyzing market integration. In particular, we perform cluster analyses on various statistical indicators of eight Ethiopian local price series to analyze how they relate to world market prices. Moreover, recognizing several policy regimes in the period 2007-2010 we investigate how market clusters change over time. Results show that in periods with wheat imports via the private sector, several local markets are clustered with the world market. In periods with government controlled imports and exchange rate collapse, the characteristics of domestic prices were strongly dissimilar from those of world market prices.*

Keywords: Cluster analysis, Sub-Saharan Africa, market integration, missing data, wheat markets

JEL codes: C22, Q11, Q13

## 1. Introduction

Market integration is an important topic in agricultural economics. The degree in which local markets are integrated with wider regional or global markets signals the possibilities for (arbitrage) trade between different regions and whether price signals from international markets are transmitted to local markets. The latter is important for farmers in developing markets that may want to connect to larger international markets. However, knowledge of the extent of price transmission also provides better insight in how net food consumers in developing countries are affected by global food price shocks. If local markets are not integrated, global price spikes as experienced in recent years have limited or no effects on them.

A number of reviews on the theory and empirical analysis of market integration and price transmission are available (see e.g. Fackler and Goodwin, 2001; Rapsomanikis et al., 2006). A popular econometric approach to test for market integration is to test for cointegration. If price series of different markets are cointegrated this signals a stable long-run relation between these series, even if both series wander due to deterministic or stochastic trends. In this context, empirical analysis often focuses on analyzing the short and long term relationships between world market and local prices. If cointegration exists price series can be modeled together in a vector error correction model. Recently, extended versions of cointegration have been applied, including threshold cointegration (e.g. Cudjoe et al., 2010) and cointegration with different regimes (e.g. Brümmer et al., 2009; Busse et al., 2012).

However, these cointegration approaches can only be applied when extensive data series are available. Time series need to be sufficiently long and free of missing observations since lags and differences need to be taken. Whereas this may not be a problem for price series from developed countries, data from developing countries is often prone to missing observations, changing data frequencies, outliers, and/or measurement error. Moreover, missing values may be replaced by simple interpolations or last period's value. Finally, series are often too short for applying standard procedures or have too few observations per sub-period. Our contribution consists in suggesting a fresh perspective on empirical research. We propose adapting the method for analysis to the weaknesses of the data instead of "tuning" the data by interpolation and other techniques to a format which can be analyzed with established methods.

In this study we have the same focus as many of the price transmission and cointegration studies published in recent years in agricultural economics. We aim at obtaining evidence on the question to what extent domestic Ethiopian markets are integrated with the international market because "*the links between Ethiopia's domestic wheat market and the international market are by no means straightforward*" (cited from: Dorosh and Ahmed, 2009). In particular, we are interested in the

extent that integration with the world market price changed during the differing policy regimes that were implemented in Ethiopia in the period May 2006 to November 2009 as identified by Dorosh and Ahmed (2009). We analyze wheat prices from seven local markets in northern Ethiopia and prices in the capital Addis Ababa. This is very relevant because local food prices have a strong and direct impact on the welfare of the poor. Moreover, with rising food prices globally, concerns on increasing poverty in developing countries are often raised (e.g. Ivanic and Martin, 2008).

The available data shows many of the frequently encountered weaknesses and problems for time series of price data which are especially found in temporally and geographically disaggregated data from developing countries. The data we have available is bi-weekly but contains many missing values. For example, the data for Addis Ababa is only available on a monthly basis so that the observation of every second week is missing. In total, 111 observations are missing among the eight local price series of 86 periods each, so that 16% of observations of the total of 688 periods are missing for different markets at different points in time. Moreover, some of the prices are simply recorded constant in various periods. Although the Ethiopian government did regulate prices during certain months of our data range (Dorosh and Ahmed, 2009) it is striking to see that some markets kept prices constant whereas at neighboring markets prices fluctuated in the same period.

Since cointegration analysis is not feasible given such a dataset, we propose an alternative method in this paper to analyze interaction between the world wheat market and local Ethiopian wheat markets. We define a number of statistics describing the evolution of each series (e.g. mean, variance, skewness, kurtosis, number of spikes and price drops, average length of price spikes and drops, etc.) and then apply cluster analysis on subsets of these statistics. In this way we are able to classify the nine price series in different time periods in terms of their similarity in selected aspects of their statistical characteristics into clusters of markets. If the world market price is in one of these clusters we conclude that these markets are related.

Since we propose an alternative methodology for analyzing market integration there are no comparable studies available. There are however a number of studies that applied cointegration analysis to Ethiopian food markets. Studies on Ethiopian grain market integration have mainly focused on how the regional market prices are cointegrated with the central grain market price in Addis Ababa (Asfaw et al., 1998; Dercon, 2002; Getnet, 2007; Jaleta and Gebremedhin, 2012). However, many producers and traders in developing countries use the so-called world market prices as a yardstick upon which to base their pricing decisions (Minot, 2011). In this way it may turn out that local market prices are (partly) determined by the world market prices, which may lead to a long-run equilibrium relationship between world market prices and local market prices. Local conditions and shocks may however weaken this relationship. Rapsomanikis and Mugeru (2011)

reviewed a number of articles that assessed possible long-run relationship between world market prices and domestic market prices of commodities in selected countries. Their analysis revealed, particularly for the Ethiopian wheat market that there was cointegration with the world market. However, they mentioned that the speed of adjustment was rather slow (only about 0.06%).

The paper is built up as follows. After this introduction, section two discusses the problem on incomplete data in applied economic analysis. Section three provides some background about the Ethiopian wheat markets and the policy regimes. Section four describes the data whereas section five describes our methodology, including a detailed overview of the statistics used and the clustering methodology. Section six discusses the results of the cluster analysis and section seven provides conclusions and implications.

## **2. The challenge of incomplete data in applied economics analysis**

Empirical analysis in applied economics crucially depends on the quality and reliability of the dataset available for analysis. Especially types of data which have a natural ordering, such as time series or panel data, and are characterized by some kind of temporal dependence suffer frequently from various data problems. The less financial means and the lower the quality of infrastructure is available for the data gathering, the higher are the chances that repeated measurements of certain socio-economic variables are of low quality. The longer the time span and the higher the frequency of the measurement the higher this risk, especially if financial and infrastructural conditions are unstable in subsequent years. Continuous data gathering is at stake when governmental institutions have only limited financial sources for data gathering or face very high costs in doing so due to rain season, sparsely populated areas, low quality of infrastructure. Moreover, when the data collection institutions are weak and suffer from corruption or other issues these problems may be aggravated. Hence, for the applied researcher, one of the most substantial challenges for analysis consists in obtaining complete and reliable series of data. Very commonly this is not the case so that missing observations have to be dealt with in a very early step of analysis. Often, the more interesting the problems to be analyzed the more difficult is it to get complete and reliable data. Figure 1 illustrates some of the most frequently found data problems by displaying a subset of the price series that are used in this analysis:

- Missing data either as single or subsequent missing observations for one or more series in the same or different periods
- Non-corresponding data frequency between series
- Too short time series for applying standard time series approaches or too few observations per sub-period of interest for regime-dependent models

- Outliers
- Constant prices for (many) subsequent periods or other issues of measurement error or limited reliability

*Insert Figure 1 here*

Standard time series methods are more or less strongly influenced by such flaws in the data or can even not immediately be applied before the missing values have been imputed by some approach. The established approaches show a tendency towards increased sophistication (Hassouneh et al., 2012). However, the effort centers around making the data suitable for methods. On this way, many decisions and assumptions have to be made in order to obtain the final dataset. Examples are the choice of the imputation method, the decision whether and where to cut the data or whether to replace constant values or outliers by imputations. Many of these necessary choices are arbitrary and introduce uncertainty into the data because missing observations are in some way calculated. Time series models tend to become increasingly sophisticated and, therefore, depend increasingly on the quality and reliability of the underlying data. Nevertheless, estimation results are often displayed up to three or more decimal places, pretending accuracy and precision that does not exist due to the underlying data problems.

Based on these considerations, we pursue a different perspective in this research. We acknowledge the frequently existing data problems and suggest a correspondingly simple but potentially meaningful approach in order to make sense of the data and extract information from it in order to answer the questions of interest. Since the approach is based on a certain aggregation of the data, it is much less dependent on the completeness and reliability of the data and, therefore, less susceptible and largely robust to the above mentioned data problems. Our aim consists in reducing sources of uncertainty as far as possible or even excluding them. So the approach which we suggest and describe below in detail is easily applicable because it is based on a widespread and frequently used methodology. We suggest calculating a set of descriptive statistics either for the entire period of observation or for sub-periods which can either be irregular based on economic or other evidence predefined regimes or equally-spaced sub-periods of the data. These indicators can be grouped into categories measuring volatility or extreme price developments. The time series can then be analyzed via a multivariate cluster analysis by using several indicators or all indicators of several categories. The largest advantage is that this is a pragmatic approach which extracts most information possible from a dataset of time series with potentially many imperfections. It is easily implementable with almost any statistical software package and even partly with MS Excel. It is able to deal with the

data problems mentioned above which are frequently encountered in applied research. There are no assumptions made for the imputation of missing data. The analysis is very flexible because it can be run on a subset of indicators depending on the context of interest (e.g. only level or only volatility measures). Variations depending on the context are easily implementable. The approach can be run for the entire period of observations or repeatedly for equally spaced time windows or for unequally spaced regimes. The minimum requirement of observations can thereby be much lower than for classical time series analysis and would still produce meaningful results via the choice of suitable indicators for clustering.

Disadvantages of the approach are that it is not founded on established statistical theory as, e.g., cointegration. Also several decisions have to be made concerning the concrete specification of the cluster analysis (e.g. choice of algorithm and distance measure; choice of characteristics). However, since no uncertainty has to be introduced via the imputation of observations, and due to its easiness in implementation, it provides an interesting alternative to data-demanding time series approaches for analysing market integration.

### **3. Ethiopian wheat markets and the role of the world market**

Wheat is an important staple food in Ethiopia. Though a substantial amount of wheat is produced locally, Ethiopian farmers are not always able to meet the nation's demand. Physical-climatic factors such as drought and poor market access, among others contribute to this insufficient supply. Consequently, a substantial share of the nation's wheat supply comes from imports. Prices paid for imported wheat may affect local prices too. Thus, world market prices can play an important role in determining local wheat prices in importing countries. Moreover, many local traders and producers observe price movements in neighboring regions or other parts of the world and use this information to form price expectations for local trade. This way, local prices of goods and services are not only determined by the market conditions within a region or country, but also partly by the movements in prices elsewhere. In this way prices are said to be transmitted from one region or country to the other. Price transmission between two spatially separated regions means that price levels of commodities in one country influence price levels of similar or related commodities in other countries, especially when these countries are connected through trade.

Wheat is one of the six major staple grains produced in Ethiopia<sup>1</sup>. In terms of acreage wheat ranked fourth in 2012/2013 with 1.5 million hectares (Tefera, 2013) and in terms of average per capita calorie consumption wheat is the third most important food item in Ethiopia (Berhane et al., 2011). Variety in grain production and consumption are due to heterogeneity in agro-ecology and socio-

---

<sup>1</sup> Other major staple grains in Ethiopia are corn, teff, sorghum, barley and millet.



economic conditions. For example, wheat and teff are cool-weather crops whereas corn and sorghum are grown in warmer areas. Ethiopian wheat production and wheat markets are therefore mainly found in highland areas. Main wheat growing areas are in Tigray and around Addis Ababa (Tefera, 2013).

Among the Ethiopian staple grains, wheat is the only grain that is imported substantially and therefore potentially depends on price developments on world markets (Dorosh and Ahmed, 2009), allowing for integration of Ethiopian wheat markets with the global market. There are a number of caveats however, that may limit or block relations between global and Ethiopian wheat prices. First, part of the foreign wheat that enters the Ethiopian market is in the form of food aid, particularly in years with crop failures. The amount and the conditions under which food assistance imports enter the market vary. Tefera (2013) estimates that in recent years on average 20% of wheat imports were in the form of food aid. The effect of food aid is that it lowers local market prices by increasing supply. Second, after commodity prices rose globally in 2007, Ethiopia faced a shortage in foreign currency. Since March 2008 the government therefore rationed foreign currency, limiting the possibilities for commercial traders to import wheat (Dorosh and Ahmed, 2009). After crop failures in 2008 the state owned Ethiopian Grain Trade Enterprise (EGTE) was the main importer that sold imported wheat on the Ethiopian market at reduced prices.

Dorosh and Ahmed (2009) distinguish different regimes in domestic Ethiopian wheat markets. Table 1 summarizes the relevant information from their paper. The connection between domestic and world market prices is therefore also likely to vary accordingly. During Regime 1 until most of Regime 3, domestic prices (at least those in Addis Ababa) are expected to be around or slightly above world market price levels. In regime 4, domestic price levels diverge from world market prices, so that the levels become more different. Volatility in the first two regimes is expected to resemble world market conditions while it should be lower in the third regime and substantially lower in Regime 4 because the strongly subsidized sale of wheat by the government during the second half of 2008 is likely to reduce price levels as proved by Dorosh and Ahmed (2009) and to calm price uncertainty.

*Insert Table 1 here*

#### **4. Wheat price data**

This analysis uses a unique dataset of local Ethiopian wheat retail prices from the Tigray region, Northern Ethiopia. These local retail prices are bi-weekly, collected at 15-day intervals, on the 1<sup>st</sup>

and 16<sup>th</sup> day of every month from May 2006 to November 2009. This sample period captures the global price spike of 2007-2008 and the subsequent price decline. The data are complemented by retail and wholesale prices in the capital city Addis Ababa and a world market price from FAO. The prices were measured in Ethiopian Birr per quintal but converted to Birr per ton. Local prices were collected by the Tigray Agricultural Marketing Promotion Agency (TAMPA) from retail markets in Alamata, Maychew, Abi-Adi, Mek'ele, Hawzen, Adigrat, and Aksum. Figure 2 is a map showing the approximate locations of the various markets from which wheat prices were measured for this study.

*Insert Figure 2 here*

Since Addis Ababa is the capital city and also hosts the central grain market, we also included the retail prices from this market. The world market price was obtained from the FAO database on wheat prices within the same period (i.e. 2006-2009) and at the same frequency.

The summary statistics of the various market prices are presented in table 2 below. Across all the markets the average prices of wheat exhibited a reasonable degree of variability within the span of our data. Among the local markets, even though the Addis Ababa retail price had a relatively small mean price, it is the market that showed the largest variation in wheat prices, as evidenced by the largest coefficient of variation (46%). The local market with the lowest spread of wheat retail market price is Axum (with a coefficient of variation value equals 33%). Contrary to the expectation that Addis Ababa retail prices would show less variability because it hosts the central grains market, we rather see that Axum (CV = 33%) and Hawzen (CV = 35%) markets show less variability. These markets, especially Hawzen, tend to be isolated; hence it is quite realistic that prices there vary much less. Again, the Hawzen market is moderately distant from the main trunk road, and therefore price transmission to such a market might delay. In most African countries, capital cities (such as Addis Ababa) tend to serve as the point of concentration of most commodities produced in the countryside or production centers. In the process of concentration, the prices of goods may respond rapidly to price changes depending on the conditions of supply and demand. As might be expected, the wheat price at the international market showed the least variability of 27%.

*Insert Table 2 here*

Figure 1, already presented in section 2, plots the world market price and the local Ethiopian market prices transformed into indices with the base period of the last week of 2006. At a glance, the figures reveal that the variability of the wheat price in all local markets as well as the world market is not constant over the entire range of the dataset. However, based on regimes (as indicated by R2

– R5), variability is quite distinct in each regime. Particularly in regime 2, which spans the period of first week in May, 2006 to the first week in March, 2007 the prices appear relatively stationary or stable. In regime 3, the prices of wheat in all local markets as well as the world market price showed a marked degree of upsurge. A remarkable feature of this regime is the observation that the price of wheat at the world market was higher and above all local market prices. This regime (spanning between March, 2007 to May, 2008) coincides with the onset of the global food shocks which led to escalating international market commodity prices. According to Dorosh and Ahmed (2009) during this regime, there were significant food aid inflows, which might partly have contributed to the relatively lower domestic prices compared to the world market. Within the second and third regimes, the general observation is that local and international wheat prices tend to move closely together, supporting the market integration hypothesis. However, in regime 4 (beginning late May, 2008 to early May, 2009 prices fluctuated considerably. Prices initially surged upward for some months and then reverted to a downward trend. In this regime, international wheat prices were lower than the domestic (local) wheat prices. The trend continued for the ensuing months until the end of November, 2009.

Comparatively, over the entire period under consideration (denoted as regime 1), prices of wheat in almost all local retail markets were generally lower than the corresponding world market prices from May 2006 to May 2008. However, the trend was reversed and local retail prices attained higher levels than world market prices after May 2008, and this trend was sustained until approximately November, 2009. Interestingly, the graph reveals that after May 2008, the local market prices continued to increase even when the world market prices were declining. This might have resulted from the crop failure caused by severe drought and pest infestation that occurred around the same time (2008) in most wheat growing regions of Ethiopia (Meijerink et al., 2009; Dorosh and Ahmed, 2009). Moreover, the same period marks the culmination of the global food crisis which caused instability in food markets of most countries, especially in the developing world. As mentioned above, policy measures such as foreign exchange rationing rather aggravated the already fragile situation and prices remained high after April 2008.

## **5. Methodological Approach**

To analyze relations between the markets described, this study uses a cluster analysis of a set of descriptive statistics that characterize the individual price series. These statistics are to be calculated for the time periods of interest, so either one set of statistics for each series in the entire period for which observations are available, or sets of statistics for all relevant sub-periods. Such sub-periods can be regularly-spaced intervals of the data or, as in our case, several periods of differing lengths

which are defined by policy regimes, economic theory etc.. Thus, the information available in the time series is aggregated into descriptive statistics each of which is measuring a certain aspect of the respective time series in the respective time interval. Missing observations do not play a role in the calculations, and other problems concerning single observations or sequences are mitigated. Because individual statistics might be stronger affected than others by such data problems, regarding a set of such indicators will again mitigated this problem.

The central idea of the approach we suggest consists in classifying or ordering the descriptive statistics of the time series in each relevant time interval by employing some multivariate classification method in the second step. This step might involve all indicators calculated or only certain categories of them which are of immediate interest for the research question or as robustness analysis. Each observation  $s_{mp} = (s_{mp}^1 s_{mp}^2 \cdots s_{mp}^J)'$  for the second step consists of statistics  $s_{mp}^j, j = 1, \dots, J (J = 14)$  for all time series  $m = 1, \dots, M (M = 9)$  and all periods  $p = 1, \dots, P (P = 5)$ . The classification is performed separately for each period  $p$ . This leads to some ordering in terms of the degree of similarity of the multivariate vectors of the indicators of the time series for each period. In this way, one could identify the groups of time series which show the most similar measures of volatility or the levels of the time series or which series are most dissimilar in these respects to the remaining series. The relevant indicators can be identified based on the focus of the analysis or other considerations.

In the context of the measurement of the extent and the speed of price pass-through between markets, measures of price level and, possibly, volatility appear to be of largest interest. Market integration is about assessing the short-run and the long-run responses of prices to demand or supply shocks. Perfectly integrated markets are expected to show very similar price developments because shocks are passed through immediately and completely between them. In other contexts, one might be interested in similarities of extreme price developments or various characteristics of positive price changes etc..

Based on these ideas, our final strategy for empirical analysis is the following. We transform the price data into indices as they are plotted in Figure 1. Subsequently, we calculate the indicators for the eight Ethiopian price series and the world market price in USD<sup>2</sup> for regimes 2 to 5 and for the entire range of observations from May 2006 until November 2011. In particular, we calculated  $J = 14$  indicators belonging to one of the five categories price levels, volatilities, extreme price developments, direction of price development and further moments (Table 3). Each indicator measures one aspect of the development of the dynamics of the price series as indicated in the last

---

<sup>2</sup> We use the world market price in USD instead of in Ethiopian Birr because the exchange rate between both currencies strongly changed and was subject of governmental manipulation. Because our focus lies on the pass-through of price signals from the world market to domestic markets, we prefer to exclude the effects of changes in the exchange rate.

column of the table. This set can be extended by any other indicators or categories relevant for the analysis of interest. We run the cluster analysis on different subsets of indicators:

- I) All 14 indicators (Table 4),
- II) Only the two price level indicators (Table 5),
- III) Only the price volatility indicator (Table 6),
- IV) All price level, volatility and moment indicators (Table 7) and
- V) All indicators measuring price levels, price volatilities and the direction of price development (Table 8).

Additionally, we run a sensitivity analysis for the entire period in order to assess the robustness of the results to leaving out one indicator or an entire category of indicators.

We used the agglomerative hierarchical cluster analysis (Hair et al., 1998: 476) due to its ability to create hierarchically related set of clusters, which is relevant to our study. Further, the complete linkage method was selected for comparing the degree of similarity between time series from the different markets. The complete linkage method uses the farthest pair of observations between the two groups to determine the similarity or dissimilarity of the groups. The dissimilarity measure used for calculating the dissimilarity matrix was the L2 (Euclidean distance), which is a common measure of distance (dissimilarity) between any two observations.

*Insert Table 3 here*

## 6. Results

Summarizing and presenting the results of the various cluster analyses is challenged by the fact that the main output of such analysis is graphical in the form of a dendrogram (Figure 3). We follow the usual approach by choosing a cut-off threshold in order to obtain a small number of clusters which allows meaningful interpretation. We aim at obtaining 2-4 clusters of the nine markets. We choose as threshold the 75%-quantile of all obtained dissimilarity distances for the given model specification. Since the objective of the analysis consists in assessing to what extent price shocks are passed through from world wheat markets to domestic Ethiopian markets, interpretation focuses on the dissimilarity of the vectors  $s_{mp}$  of the indicators relatively to the vector of indicators of the world market price. In particular, we classify the data of the eight domestic markets according to their similarity to the indicators of the world market price by quantifying their distance to the cluster containing the world market price.

Figure 3 is therefore to be interpreted in the following way as quantified in the last column of Table 4. The threshold at 26.1 is marked with the horizontal dotted line, so that there are three resulting

clusters, that is, two clusters which do not contain the world market price. All prices except of Addis Ababa and Abiadi are classified together with the world market price into one cluster. Therefore, they are assigned the value 0 in the last column of Table 4. Addis Ababa is the next closest cluster and Abiadi forms the most dissimilar cluster. They are assigned the values 1 and 2, respectively. In order to ensure comparability between models with differing numbers of clusters, we standardize these quantities by dividing them by the number of clusters which do not contain the world market price, so that Addis (Abiadi) is assigned the value 0.5 (1) in Table 4.

Table 4 displays the results for the cluster analysis for all nine time series using the entire set of 14 indicators in the five categories both for regimes 2 to 5 and for the entire period of observations. A zero denotes that the series is in the same cluster as the world market price WM and unity denotes that the series is located in the most dissimilar cluster. Values in between denote growing dissimilarity to the characteristics of the cluster containing the world market price. Thus, growing numbers of our dissimilarity measure indicate increasing differences of the series relatively to the characteristics of the world market price, unity indicating maximum dissimilarity.

*Insert Table 4 here*

The last column indicates that six of the eight Ethiopian markets belong to the same cluster as the world market price does when considering the entire observation period without policy regimes. This result is based on the most comprehensive vector of the price characteristics encompassing all 14 indicators mentioned in Table 3. The characteristics of the capital Addis Ababa are more dissimilar and the ones of Abiadi which is a small remote town in the north of the country are most dissimilar. This can be interpreted as most of the local wheat markets considered being generally well integrated with the world market because the characteristics of their prices are very close to the ones of the world wheat price in USD.

The separate cluster analyses for regimes 2 to 5 indicate somewhat significant differences across the observation sub-periods. In regime 2, when the domestic prices were mainly determined by domestic supply and demand and profitable and significant private sector imports, the characteristics of Addis prices are very close to world market prices since most of imports arrive there. Most remaining local markets form a joint cluster which is separate from the capital with Abiadi being an exception. In regime 3, when private sector imports were no longer profitable because domestic price levels were below the world market price and foreign exchange rationing started, the nine markets split into two major clusters: Four markets in the relatively remote very north east of the country still closely resemble the characteristics of the world market price, while

the remaining markets form a different cluster signaling that prices in Addis Ababa and other markets showed characteristics very different from the world market price. In regime 4, four domestic markets are again categorized together with the world market and the remaining four markets split into two clusters. This appears to be plausible because Hawzen is a small town of less than 6000 inhabitants in the north and the Ethiopian government massively sold strongly subsidized wheat in the capital. In regime 5 this development relaxed because only three markets form a different cluster.

In summary, the integration of domestic markets with the world market and among each other worsens markedly the less private sector imports take place and the stronger governmental interference takes place. This evolution is also indicated by the vertical sums of the standardized dissimilarity measures in the columns of Table 4 which increase from 4 in regime 2 to 5 in regime 3, indicating that the import restrictions introduced in this period and the results uncertainty led to fragmented local markets, which slightly eases afterwards. Across the regimes, Hawzen, Makelle and Adigrat show least dissimilarities with characteristics of the world wheat prices because they have the smallest horizontal sums. Axum and Alamata appear to be most dissimilar which is reasonable because the former town is remotely located in the mountainous north of the country while the latter is a small town so that prices are likely to be more impacted by local factors.

Tables 5 to 10 give more detailed insight by presenting the results of cluster specifications II to VII mentioned above. Table 5 illustrates that a more pronounced evolution is observed when exclusively looking at price levels. In regime 2 during which private sector imports integrate world and domestic prices, the levels of all Ethiopian price indices resemble international wheat prices. In regime 3, all markets except Axum form a cluster separated from the world market. In regimes 4 and 5, when governmental interference restricted free trade and the exchange rate of the Birr against USD collapsed (Dorosh and Ahmed, 2009), none of the Ethiopian prices shows characteristics similar to the world market.

*Insert Table 5 here*

The picture is different for price volatility (Table 6). The dissimilarity in terms of volatility does approximately change as much across the entire regimes - as visible in the row “sum” - as it does for the level indicators. There are strong differences among the dissimilarities of the local markets. The largest city in the dataset, Addis Ababa, with 3m inhabitants, and Maichew located in the central production region are closest to international price volatility across all regimes. This finding is plausible since these cities possess the cheapest and the most connections to possible supply

sources and storage facilities are best and largest. In contrast, remote markets such as Makelle, Axum and Abiadi appear to be most dissimilar.

*Insert Table 6 here*

The joint consideration of three categories level, volatility and moment indicators (see Table 3) in Table 7 illustrates that a multivariate view gives a more comprehensive vision: in regime 2 only the four markets lined up along the road to the north form a separate cluster. In regimes 3 and 4, when private imports are heavily impeded, the characteristics of local markets strongly disconnect from world market developments while this eases slightly in regime 5. Across all regimes, characteristics of prices in Axum, Addis Ababa, Hawzen and Abiadi are most similar to the ones of the international wheat price while the ones in Adigrat and Alamata are most dissimilar.

*Insert Table 7 here*

Results of Table 8 showing the clustering of markets based on level, volatility and rice development which most closely resembles the logic of cointegration analysis indicators show a very similar pattern: dissimilarity to world market prices is lowest (strongest) in regime 2 (regime 4) when private sector imports were high (prohibited).

*Insert Table 8 here*

Tables 9 and 10 contrast the clustering of the markets in regimes 2 and 4 based on various indicator sets. In regime 2, strongest differences of local markets relatively to the world market exist for price volatility and extreme price observations since the average dissimilarity amounts to 0.6 and 0.4, respectively. Price characteristics in Addis Ababa appear to be most similar to world markets while prices in the remote small town of Abiadi are clearly most dissimilar. Both results are plausible since the significant private sector imports which took place in this regime were mainly shipped to the large capital market while remote regions remain also in such phases comparatively isolated because traders' profits are very limited.

*Insert Table 9 here*

In regime 4, this picture changes strongly: Dissimilarity particularly in terms of price levels and levels, volatility and price directions combined markedly increases.



*Insert Table 10 here*

In order to illustrate the differences in price characteristics for a phase during which governmental interference aimed at disconnecting domestic from international markets, Table 11 shows several differences between the two clusters identified for regime 4 in Table 8 which become also evident in Figure 1. One cluster consists only of the world market price while the other cluster is formed of all eight local Ethiopian prices. While the index of the world wheat price in USD is decreasing throughout this phase amounts to the 1.3 fold of its base period, continue the domestic Ethiopian prices to rise and reach on average the 2.6 fold of their base levels (C1) although their relative volatility amounts only to 73% of the former (C2). World market prices show a clearly negative long-run development: The last observation of the world market price in this phase is more than 20% smaller than the starting observation, decrease the local prices barely (C3). In the short-run, the share of non-negative price changes from observation to observation in the total period is by more than 40% higher for the domestic prices (C4).

*Insert Table 11 here*

We checked the robustness of our suggested method regarding the number and type of indicators considered in the multivariate clustering. Table 12 indicates that the results of the varying specifications appear to be very stable. Only if dropping all five indicators of extreme price developments in M6 or the indicator of the direction of the price development in M13, the results change somewhat in comparison to the benchmark model M1. For all other model variations, results concerning the cluster structure are identical. This suggests that considering a large number of indicators yields fairly stable and reliable results.

*Insert Table 12 here*

## **7. Conclusions**

Our study aimed to investigate whether local Ethiopian wheat prices relate to world market wheat prices. To achieve this goal, we used wheat prices from several local markets in Ethiopia combined with the world market wheat price over the period May 2006 to November 2009. However, as often observed for price series from developing countries, the local market prices showed many missing

values, making it impossible to perform time series cointegration analysis. This leads to an important contribution of this paper, presenting multivariate cluster analysis on various statistics of price series as an alternative methodology for studying market integration. This analysis was done for the whole sample period, but also for different market regimes in the period 2006-2009 that are identified in the literature. In this way we were able to validate the proposed clustering methodology as a useful tool in market integration studies.

Our results show that over the whole sample period six of the eight Ethiopian markets are clustered with the world market, i.e., have in general similar statistical properties. However, when the analysis is broken down to distinct policy regimes, we find that in periods with relatively free markets and wheat imports via the private sector, almost all price series belong to the same cluster as the world market price does, indicating that local and global market were closely related to each other. Periods with government controlled imports and exchange rate collapse, however, showed that very few or no local markets were clustered with the world market. These results are strongest when considering only clusters based on price level statistics. For price volatility the results are mixed. In periods with the Ethiopian government interfering with wheat markets, domestic markets are first fragmented and more different from the world market. This suggests that domestic policies indeed decoupled domestic wheat price volatility from world markets to some extent.

Although the results from the cluster analysis are plausible and in line with the known situation in various regimes, there are still a number of remaining issues to be investigated. We would like to test the sensitivity of our results to various methodological choices in the clustering methodology, such as choice of clustering methodology, choice of distance measure, and alternative measures for determining the optimal number of clusters. Despite these planned additions, the current results provide some first confidence in the usefulness of cluster analysis as a tool in studying market integration.

## References

- Asfaw, N., Jayne, T. S., 1998. *Vertical and spatial integration of grain markets in Ethiopia : Implications for grain market and food security policies*. Grain Market Research Project, Ministry of Economic Development and Cooperation, Addis Ababa.
- Berhane, G., Paulos, Z., Tafere, K., Tamru, S., 2012. *Foodgrain consumption and calorie intake patterns in Ethiopia*. ESSP II working paper 23. Addis Abeba, IFPRI.
- Brümmer, B., von Cramon-Taubadel, S., Zorya, S., 2009. The impact of market and policy instability on price transmission between wheat and flour in Ukraine. *Eur. Rev. Agric. Econ.* 36, 203-230.
- Busse, S., Brümmer, B., Ihle, R., 2012. Price formation in the German biodiesel supply chain: a Markov-switching vector error-correction modeling approach. *Agric. Econ.* 43, 545-559.
- Cudjoe, G., Breisinger, C., Diao, X., 2010. Local impacts of a global crisis: Food price transmission, consumer welfare and poverty in Ghana. *Food Pol.* 35, 294-302.
- Dercon, S., 2002. *The impact of economic reforms on rural households in Ethiopia: a study from 1989 - 1995*. Washington, DC, World Bank.
- Dorosh, P., Ahmed, H., 2009. *Foreign exchange rationing, wheat markets and food security in Ethiopia*. ESSP 2 discussion paper 004. Addis Abeba, IFPRI.
- Fackler, P.L., Goodwin, B.K., 2001. Spatial price analysis. In: B.L. Gardner and G.C. Rausser, (eds.). *Handbook of Agricultural Economics*. Amsterdam: Elsevier Science, 971-1024.
- FAO (2014). FAO Statistical Database - *Commodity Prices, Trade and Markets*. Available at <http://www.fao.org/economic/est/statistical-data/est-cpd/en/>. Accessed on 06/10/2011.
- Getnet, K., 2007. Spatial equilibrium of wheat markets in Ethiopia. *Afr. Dev. Rev.* 19, 281-303.
- Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1998. *Multivariate Data Analysis*. 5th edition. Prentice Hall, 730p.
- Hassouneh, I., von Cramon-Taubadel, S., Serra, T., Gil, J.M., 2012. *Recent Developments in the Econometric Analysis of Price Transmission*. Working Paper No. 2, Transparency of Food Pricing (TRANSFOP). Seventh EU Framework Programme.
- Ivanic, M., Martin, W., 2008. Implications of higher global food prices for poverty in low-income countries. *Agric. Econ.* 39, 375-391.
- Jaleta, M., Gebremedhin, B., 2012. Price co-integration analyses of food crop markets: The case of wheat and teff commodities in Northern Ethiopia. *Afr. J. Agric. Res.* 7(25), 3643-3652.
- Meijerink, G. W., Roza, P., Berkum, S. van., 2009. *East African governments' responses to high cereal prices*. LEI Wageningen UR, The Hague.

- Minot, N., 2011. *Transmission of world food price changes to markets in sub-Saharan Africa*. IFPRI Discussion Papers, Washington DC.
- Rapsomanikis, G., Hallam, D. Conforti, P. , 2006. Market integration and price transmission in selected food and cash crop markets of developing countries: review and applications. In: Sarris, A., Hallam, D. (eds.). *Agricultural commodity markets and trade. New approaches to analyzing market structure and instability*, Commodities and Trade Division of FAO, Rome, 187-217.
- Rapsomanikis, G., Mugera, H., 2011. Price transmission and volatility spillovers in food markets of developing countries, in: Piot-Lepetit, I., M'Barek, R. (eds.), *Methods to Analyse Agricultural Commodity Price Volatility*. Springer, New York, 165-179.
- TAMPA (2013). *Tigray Agricultural Marketing Agency*. P. O Box 523 Mekelle, Tigray, Ethiopia.
- Tefera, A., 2013. *Ethiopia. Grain and feed annual*. GAIN Report ET 1301. Washington, USDA-FAS.
- Wikipedia (2014). *Map of Ethiopia*. Available at <https://en.wikipedia.org/wiki/Ethiopia>. Accessed in October 2014.

## Tables and Figures

*Table 1. The policy regimes defined in Dorosh and Ahmed (2009)*

	<b>Regime 1</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>
Start	01/2000	07/2005	04/2007	06/2008
End	06/2005	03/2007	05/2008	05/2009
<i>Number observations in dataset</i>	0	22	29	24
<i>Domestic wheat prices</i>	Between import and export parity levels (Addis Ababa), 24% below import parity	Addis Ababa wholesale at import parity levels	Below/ above import parity levels, but above export parity levels	Above import parity levels, decreasing due to state imports & subsidized sales (but smaller than expected) Divergence from WM continued
<i>Private sector imports</i>	Not profitable & no role	Profitable & significant magnitude	Not profitable & no role until end of 2007 Significant until 03/2008 Restricted after 03/2008 via foreign exchange rationing	Restricted via foreign exchange rationing although profitable
<i>Trade &amp; domestic policies regarding wheat</i>	Government imports Food aid inflows	Increased domestic credit Food aid inflows	Government imports Import restrictions via foreign exchange rationing Policy uncertainty due to state imports & domestic sales etc.	Government imports Subsidized domestic sale at constant price (ca. 50% of Addis wholesale price) Huge implicit subsidies to millers, traders and poor households (US\$90m) Food aid inflows
<i>Major price determinant</i>	Domestic supply (incl. official imports) and demand, food aid depressed prices	Domestic supply & demand, food aid, private sector imports – therefore world market	Independent of WM	
<i>Expected price transmission from world to market</i>	Domestic prices vary within the band between import and export parity, shocks only transmitted if large enough	Domestic prices follow WM, shocks fully transmitted	Domestic prices vary within the band between import and export parity, shocks only transmitted if large enough	Domestic prices independent from WM, shocks not transmitted
<i>Expected level of domestic prices</i>	Around WM, follow in tendency WM	Larger than WM by transport costs, follow in tendency WM	Larger than WM by transport costs, follow in tendency WM	Diverging from WM
<i>Expected volatility of domestic prices</i>	No difference between domestic markets and WM	No difference between domestic markets and WM	Lower in domestic markets than for WM	Substantially lower in domestic markets than for WM

Notes: “WM” denotes the world market price. All characteristics from Dorosh and Ahmed (2009) except of the last three lines which are the hypotheses regarding price transmission which follow from the regime characteristics. Dorosh and Ahmed (2009) only cover the period until 05/2009 which only partly corresponds to the dataset available to us. Since our data starts only in May 2006, we are not able to consider Regime 1. Because our data ranges until November 2009, we consider Regime 5 starting in June 2009 with 11 observations in the empirical analysis for which no information about policy regimes is available.

*Table 2. Summary statistics of the price series*

<b>Price series</b>	<b>Obs</b>	<b>% miss</b>	<b>Mean</b>	<b>Std</b>	<b>Min</b>	<b>Max</b>	<b>CV</b>
<b>Maichew</b>	77	10%	4928	2243	2448	10560	0.46
<b>Alamata</b>	76	12%	5263	2297	2599	9120	0.44
<b>Makelle</b>	81	6%	5346	2064	2660	9100	0.39
<b>Abi-Adi</b>	79	8%	5034	1915	2800	9790	0.38
<b>Hawzen</b>	72	16%	4443	1669	2350	7900	0.38
<b>Adigrat</b>	75	13%	5030	1765	2900	8400	0.35
<b>Axum</b>	75	13%	5396	1964	2428	9990	0.36
<b>Addis Ababa</b>	42	51%	4126	1351	2500	7330	0.33
<b>World market (USD/t)</b>	84	2%	271	74	192	494	0.27

Notes: “Obs” is the number of observations per series,” % miss” denotes the share of missing observations, “Std” the standard deviation and “CV” the coefficient of variation. All prices in Birr/t except the world market price.

Source: Authors based on data from TAMPA (2013).

*Table 3. Categories of descriptive statistics considered*

<b>Category</b>	<b>Indicator</b>	<b>Price characteristic measured</b>
Price levels	Mean (C1)	Level and direction of index development
	Median	Identical to before but excluding extreme observations
Price volatilities	Coefficient of variation (C2)	Relative volatility measure
Extreme price developments	Ratio btw. first and second extreme value divided by temporal difference between max and min	Magnitude of extremes relative to their temporal distance
	Ratio btw. first and second extreme value divided by ratio btw. first and last observation in period	Directional measure of extremes relative to long-run development
	Difference in number of periods btw. first period and period in which occurs the max	Location of max (and indirectly: speed at which max is reached)
	Difference in number of periods btw. max and min in period	Order and temporal distance of extremes
	Number of absolute price changes from observation to observation $\geq$ mean of changes $\pm 2$ std of price changes in the period	Measure of uncertainty
Direction of price development	Sign of difference between last and first observation in period	Direction of long-run development
	Absolute value of ratio between last and first observation in period (C3)	Magnitude of long-run development
	Number of absolute price changes from observation to observation $\geq 0$ as share of total observations of time series in period regarded (C4)	Short-run directional measure
	Median of absolute price changes from observation to observation	Short-run price development, excluding impact of extremes
	Median of price changes from observation to observation	Short-run price development, excluding impact of extremes
Further moments	Skewness	Indication if extremely positive or negative observations dominate the price distribution

Source: Authors.

*Table 4. Results for multivariate clustering using the entire set of 14 indicators*

<b>Price series</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>	<b>Regime 5</b>	<b>Sum</b>	<b>Entire period</b>
<b>Maichew</b>	0.5	1.0	0.5	0.0	<b>2.0</b>	0
<b>Alamata</b>	0.5	1.0	1.0	0.0	<b>2.5</b>	0
<b>Makelle</b>	0.5	0.3	0.0	0.0	<b>0.8</b>	0
<b>Adigrat</b>	0.5	0.3	0.5	0.0	<b>1.3</b>	0
<b>Addis</b>	0.0	0.7	0.0	1.0	<b>1.7</b>	0.5
<b>WM</b>	0.0	0.0	0.0	0.0	<b>0.0</b>	0
<b>Hawzen</b>	0.5	0.3	0.0	0.0	<b>0.8</b>	0
<b>Axum</b>	0.5	1.0	1.0	1.0	<b>3.5</b>	0
<b>Abiadi</b>	1.0	0.3	0.0	1.0	<b>2.3</b>	1
<b>Sum</b>	<b>4.0</b>	<b>5.0</b>	<b>3.0</b>	<b>3.0</b>		<b>1.5</b>
<b>Number of clusters without WM</b>	<b>2</b>	<b>3</b>	<b>2</b>	<b>1</b>		<b>2</b>

Notes: “WM” denotes the world market price. The column "Entire period" quantifies the dendrogram in Figure 3 by using the approach described at the beginning of the results section. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).



*Table 5. Results for multivariate clustering regarding the two price level indicators*

<b>Price series</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>	<b>Regime 5</b>	<b>Sum</b>
<b>Maichew</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Alamata</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Makelle</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Adigrat</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Addis</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>WM</b>	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Axum</b>	0.0	0.0	1.0	1.0	<b>2.0</b>
<b>Abiadi</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>Sum</b>	<b>0.0</b>	<b>7.0</b>	<b>8.0</b>	<b>8.0</b>	<b>3.0</b>
<b>Number of clusters without WM</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>1</b>	

Notes: “WM” denotes the world market price. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).

*Table 6. Results for multivariate clustering regarding the volatility indicator*

<b>Price series</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>	<b>Regime 5</b>	<b>Sum</b>
<b>Maichew</b>	1.0	0.0	0.0	0.5	<b>1.5</b>
<b>Alamata</b>	0.0	1.0	1.0	0.0	<b>2.0</b>
<b>Makelle</b>	1.0	0.3	1.0	1.0	<b>3.3</b>
<b>Adigrat</b>	0.0	1.0	1.0	0.0	<b>2.0</b>
<b>Addis</b>	0.0	0.7	1.0	0.0	<b>1.7</b>
<b>WM</b>	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	1.0	0.0	1.0	0.0	<b>2.0</b>
<b>Axum</b>	1.0	1.0	0.0	1.0	<b>3.0</b>
<b>Abiadi</b>	1.0	0.3	1.0	0.5	<b>2.8</b>
<b>Sum</b>	<b>5.0</b>	<b>4.3</b>	<b>6.0</b>	<b>3.0</b>	
<b>Number of clusters without WM</b>	<b>1</b>	<b>3</b>	<b>1</b>	<b>2</b>	

Notes: “WM” denotes the world market price. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).

*Table 7. Results for multivariate clustering regarding all level, volatility and moment indicators*

<b>Price series</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>	<b>Regime 5</b>	<b>Sum</b>
<b>Maichew</b>	1.0	1.0	1.0	0.5	<b>3.5</b>
<b>Alamata</b>	1.0	1.0	1.0	1.0	<b>4.0</b>
<b>Makelle</b>	1.0	1.0	1.0	0.5	<b>3.5</b>
<b>Adigrat</b>	1.0	1.0	1.0	1.0	<b>4.0</b>
<b>Addis</b>	0.0	1.0	0.5	0.5	<b>2.0</b>
<b>WM</b>	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Axum</b>	0.0	0.0	1.0	0.5	<b>1.5</b>
<b>Abiadi</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Sum</b>	<b>4.0</b>	<b>7.0</b>	<b>7.5</b>	<b>5.0</b>	
<b>Number of clusters without WM</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>2</b>	

Notes: “WM” denotes the world market price. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).

*Table 8. Results for multivariate clustering regarding all indicators measuring price levels, price volatilities and the direction of price development*

<b>Price series</b>	<b>Regime 2</b>	<b>Regime 3</b>	<b>Regime 4</b>	<b>Regime 5</b>	<b>Sum</b>
<b>Maichew</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Alamata</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Makelle</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Adigrat</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Addis</b>	0.0	1.0	1.0	1.0	<b>3.0</b>
<b>WM</b>	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	0.0	1.0	1.0	0.5	<b>2.5</b>
<b>Axum</b>	0.0	0.0	1.0	0.5	<b>1.5</b>
<b>Abiadi</b>	1.0	0.0	1.0	1.0	<b>3.0</b>
<b>Sum</b>	<b>1.0</b>	<b>6.0</b>	<b>8.0</b>	<b>5.0</b>	
<b>Number of clusters without WM</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>2</b>	

Notes: “WM” denotes the world market price. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).

Table 9. Clusters of the markets in regime 2 using alternative sets of indicators

	Set of indicators						
Price series	All	Level	Volatility	Extremes	Direction	Levels, vol. and direction	Sum
<b>Maichew</b>	0.5	0.0	1.0	0.5	0.5	0.0	<b>2.5</b>
<b>Alamata</b>	0.5	0.0	0.0	0.5	0.0	0.0	<b>1.0</b>
<b>Makelle</b>	0.5	0.0	1.0	0.5	0.0	0.0	<b>2.0</b>
<b>Adigrat</b>	0.5	0.0	0.0	0.5	0.0	0.0	<b>1.0</b>
<b>Addis</b>	0.0	0.0	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>WM</b>	0.0	0.0	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	0.5	0.0	1.0	0.5	0.0	0.0	<b>2.0</b>
<b>Axum</b>	0.5	0.0	1.0	0.5	0.0	0.0	<b>2.0</b>
<b>Abiadi</b>	1.0	0.0	1.0	1.0	1.0	1.0	<b>5.0</b>
<b>Average</b>	<b>0.4</b>	<b>0.0</b>	<b>0.6</b>	<b>0.4</b>	<b>0.2</b>	<b>0.1</b>	<b>1.7</b>
<b>Number of clusters without WM</b>	<b>2</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>1</b>	

Notes: “WM” denotes the world market price. The column (row) “Sum” is the sum of the dissimilarity measures to the left (top).

Table 10. Clusters of the markets in regime 4 using alternative sets of indicators

	Set of indicators						
Price series	All	Level	Volatility	Extremes	Direction	Levels, vol. and direction	Sum
<b>Maichew</b>	0.5	1.0	0.0	0.5	0.0	1.0	<b>3.0</b>
<b>Alamata</b>	1.0	1.0	1.0	1.0	1.0	1.0	<b>6.0</b>
<b>Makelle</b>	0.0	1.0	1.0	0.0	1.0	1.0	<b>4.0</b>
<b>Adigrat</b>	0.5	1.0	1.0	0.5	0.0	1.0	<b>4.0</b>
<b>Addis</b>	0.0	1.0	1.0	0.0	0.0	1.0	<b>3.0</b>
<b>WM</b>	0.0	0.0	0.0	0.0	0.0	0.0	<b>0.0</b>
<b>Hawzen</b>	0.0	1.0	1.0	0.0	0.0	1.0	<b>3.0</b>
<b>Axum</b>	1.0	1.0	0.0	1.0	1.0	1.0	<b>5.0</b>
<b>Abiadi</b>	0.0	1.0	1.0	0.0	0.0	1.0	<b>3.0</b>
<b>Average</b>	<b>0.3</b>	<b>0.9</b>	<b>0.7</b>	<b>0.3</b>	<b>0.3</b>	<b>0.9</b>	<b>3.4</b>
<b>Number of clusters without WM</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>1</b>	

Notes: “WM” denotes the world market price. The column (row) "Sum" is the sum of the dissimilarity measures to the left (top).

*Table 11.Characteristics of world market price index vs. the average of the characteristics of the local price indices in regime 4*

	<b>Characteristic</b>			
<b>Cluster</b>	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>C4</b>
<b>WM</b>	1.30	0.16	0.79	0.38
<b>Average all local price series</b>	2.55	0.12	0.97	0.53
<b>Ratio local/WM</b>	196%	73%	122%	141%

Notes: “WM” denotes the world market price. The explanations of the characteristics can be found in Table 3.

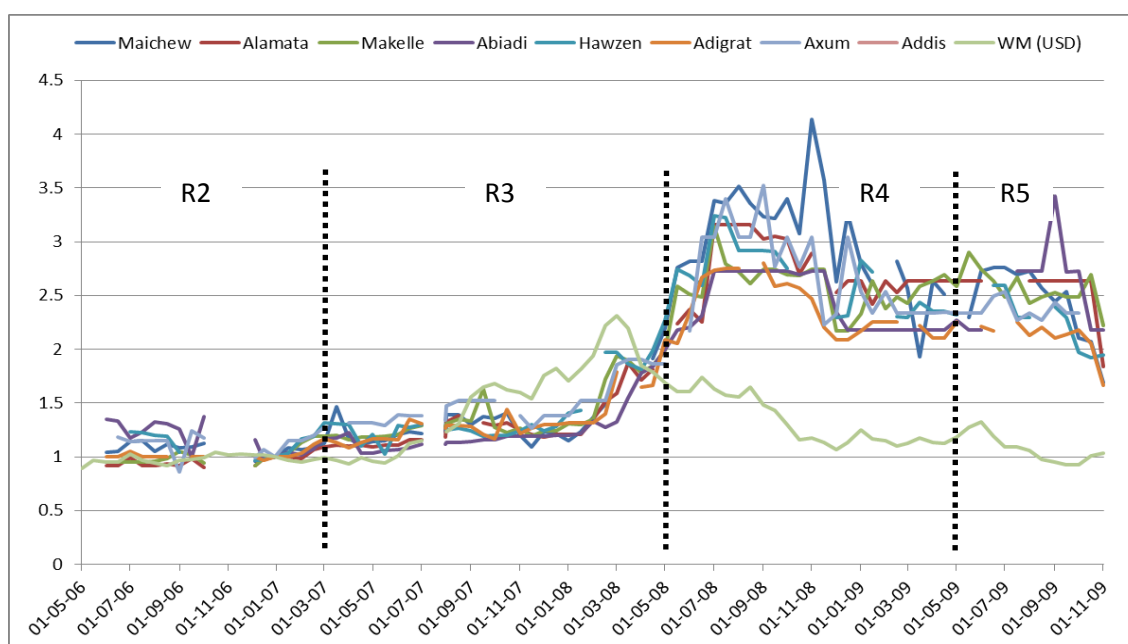
*Table 12. Results of the robustness analysis of the multivariate clustering*

<b>Model</b>	<b>Mai- chew</b>	<b>Ala- mata</b>	<b>Ma- kelle</b>	<b>Adi- grat</b>	<b>Addis</b>	<b>WM</b>	<b>Haw- zen</b>	<b>Axum</b>	<b>Abi- adi</b>	<b>Sum</b>	<b>No. of clus- ters</b>
M1	0	0	0	0	0.3	0	0	0	0.7	1	3
M2	0	0	0	0	0.3	0	0	0	0.7	1	3
M3	0	0	0	0	0.3	0	0	0	0.7	1	3
M4	0	0	0	0	0.3	0	0	0	0.7	1	3
M5	0	0	0	0	0.3	0	0	0	0.7	1	3
M6	0	0	0	0	0	0	0	0	0	0	1
M7	0	0	0	0	0.3	0	0	0	0.7	1	3
M8	0	0	0	0	0.3	0	0	0	0.7	1	3
M9	0	0	0	0	0.3	0	0	0	0.7	1	3
M10	0	0	0	0	0.3	0	0	0	0.7	1	3
M11	0	0	0	0	0.3	0	0	0	0.7	1	3
M12	0	0	0	0	0.3	0	0	0	0.7	1	3
M13	0.25	0	0	0	0.5	0	0	0	0.75	1.5	4
M14	0	0	0	0	0.3	0	0	0	0.7	1	3
M15	0	0	0	0	0.3	0	0	0	0.7	1	3
M16	0	0	0	0	0.3	0	0	0	0.7	1	3
M17	0	0	0	0	0.3	0	0	0	0.7	1	3
M18	0	0	0	0	0.3	0	0	0	0.7	1	3

Notes: “WM” denotes the world market price. The column “Sum” is the sum of the dissimilarity measures to the left.

Models M1 to M18 denote specifications of the cluster analysis for the entire period in which single indicators or entire categories are excluded from the set of the 14 indicators in the order as they appear in Table 3. M1 is the benchmark model considering all 14 indicators. In M2 all level indicators are dropped, in M3 and M4 only the mean and the median, respectively. In M5 all volatility indicators (only the coefficient of variation) are dropped, etc..





*Figure 1. Indexed local Ethiopian wheat prices and the world market price*

Notes: “R2” to “R4” denote the regimes as defined by Dorosh and Ahmed (2009). “R5” is an additional regime, see notes of Table 1. All prices in Birr/t except the world market price “WM”. The base week for indexing of the prices is the last week of 2006 starting on December 29, 2006.

Source: Authors based on TAMPA (2013) and FAO (2014).

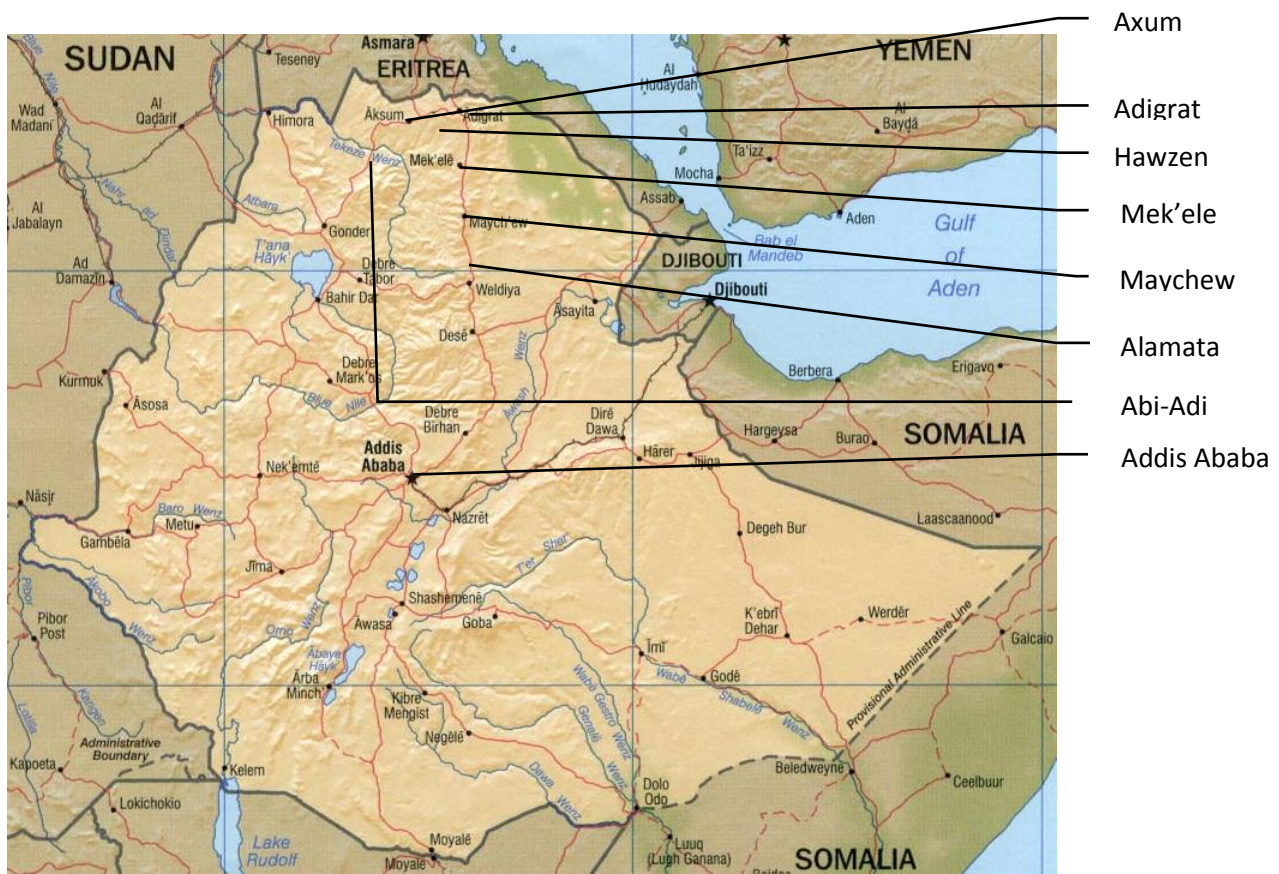
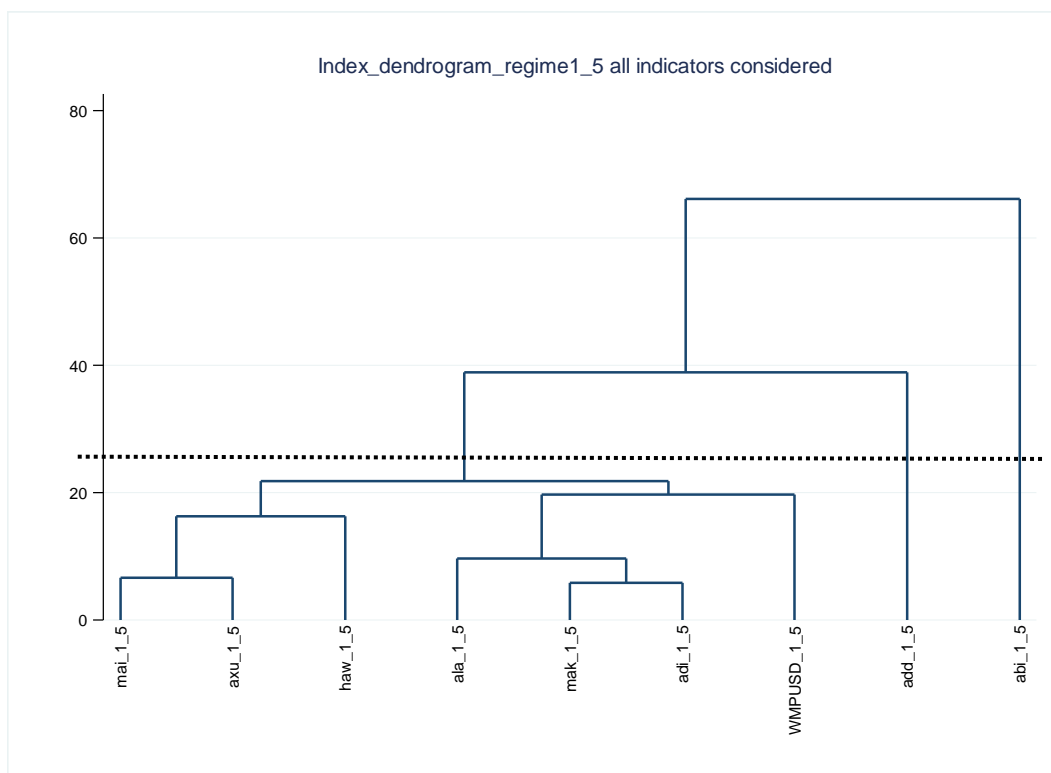


Figure 2: Map of Ethiopia indicating the local markets studied

Source: Authors based on Wikipedia (2014).



*Figure 3. Dendrogram for the multivariate clustering of all indicators for the entire period*

Source: Authors.

Notes: The names of the variables mean the following: For example “mai\_1\_5” denotes the 14 indicators of the Maichew series calculated based on all available observations in regimes 1 to 5.