Got data too poor for time series analysis? Can cluster analysis be a remedy? Studying wheat market integration in Ethiopia

Isaac Gershon K. Ansaha, Cornelis Gardebroekb, Rico Ihleb, and Moti Jaletac

Invited paper presented at the 5th International Conference of the African Association of Agricultural Economists, September 23-26, 2016, Addis Ababa, Ethiopia

Paper submitted to the 5th AAAE Conference in Addis Ababa, Ethiopia
"Transforming smallholder agriculture in Africa:
The role of policy and governance"

Got data too poor for time series analysis? Can cluster analysis be a remedy? Studying wheat market integration in Ethiopia

Isaac Gershon K. Ansah[a], Cornelis Gardebroek[b], Rico Ihle[b], and Moti Jaleta[c]
[a] Dept. of Agric. & Res. Econ., Univ. for Development Studies (UDS), Tamale, Ghana
[b] Agricultural Economics and Rural Policy Group, Wageningen University, Wageningen, The Netherlands
[c] CIMMYT, 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 a set of 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 form a common cluster with the world market. In periods with government controlled imports and exchange rate collapse, domestic prices measured by a comprehensive set of characteristics 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

The integration of agricultural and food markets at various levels of the food supply chain is a crucial determinant of the welfare of agricultural producers around the globe, of the food security of billions of consumers and of global food trade. This topic is of particular relevance in the context of developing economies. In these countries agricultural production plays a vital role for household income. Food expenses account for a larger share of household expenditures compared with industrialized countries. Processed food sold at retail markets is substantially lower than in industrialized countries. Therefore, shocks to household incomes due to changes in food prices tend to be more pronounced and affect many people more strongly. Therefore, stakeholders engaged in measuring and designing socio-economic development such as national governments (e.g., EMFA, 2015), international organizations (World Bank, 2009; AFDB, 2010), NGOs (Oxfam, 2010) or science, especially the fields of development as well as agricultural economics (Fackler and Goodwin, 2001), show vivid interest in this topic. For supporting decision making by these stakeholders, market integration needs to be measured empirically in order to infer needs for action or to monitor progress as result of initiatives or political decisions.

The degree to which local markets are integrated with regional or global markets signals how price signals from international markets are transmitted to regional and local markets. Knowledge of the extent of price transmission provides insight in how net food consuming households in developing countries are affected by global food price shocks. Therefore, increased market integration can benefit or harm producers and consumers depending on the volatility and dynamics of global prices. If markets are not integrated, global price spikes as experienced in recent years have limited or no effects.

A number of reviews on the theory and the empirical analysis of market integration and price transmission give an overview of the literature (Fackler and Goodwin, 2001; Rapsomanikis et al., 2006; Hassouneh et al., 2014). A popular econometric approach to assess market integration is to test for cointegration. The cointegration of price series of different markets is interpreted as signaling a stable long-run price relation between these markets. Empirical analysis often focuses on quantifying the short and long term relationships between world market and local prices. If cointegration is found to exist price series are often modeled in a multivariate system by means of some linear or non-linear specification of the vector error correction model. Recently, various nonlinear versions of this model have been developed and applied; examples are threshold cointegration (e.g. Goodwin and Piggot, 2001; Greb et al., 2013) or regime-dependent models (e.g. Brümmer et al., 2009; Würriehausen et al., 2014).

However, these cointegration approaches can only be applied when extensive data series are available. The series need to be sufficiently long and free of any kinds of data problems, e.g., missing observations, because the models estimate how observations in the current period are impacted by previous periods, so that lags and differences play a central role in modelling. Whereas this may be to a lesser extent a problem for price series gathered in industrialized countries, data from developing countries is often prone to a variety of such problems, such as missing observations, changing data frequencies, outliers, constant observations for a number of periods, substantial measurement error or other completeness and reliability issues. Finally, series are often too short for applying standard time series procedures in a meaningful way or have too few observations per sub-period not allowing the evaluation of sub-periods of interest. This poses a substantial challenge to applied food markets research in developing economies.

This article contributes a fresh perspective on empirical research in market integration. We propose adapting the method of analysis to the weaknesses of the data instead of 'tuning' or 'fixing' the data by some technique to a format which suits the requirements of established analytical methods. This paper illustrates how cluster analysis can be implemented in the context of practical market integration research and how the researcher can make sense of the results. We demonstrate the quantitative analysis of typically incomplete data qualitatively analyzed in the IFPRI working paper by Dorosh and Ahmed (2009) who study the effects of changing policy frameworks on Ethiopian wheat markets integration. 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*" (Dorosh and Ahmed, 2009). In particular, we are interested in the extent to which integration with the world market price changed during the differing policy regimes in Ethiopia between May 2006 and November 2009 as discussed in Dorosh and Ahmed (2009). If local markets are integrated with the world market in a policy regime, we hypothesize to see similar characteristics that cause the associated price series to cluster together. Characteristics of prices in markets less or not integrated with the world market can be expected to be more dissimilar.

A number of studies applied cointegration analysis to Ethiopian food markets (Asfaw et al., 1998; Dercon, 2002; Getnet, 2007; Jaleta and Gebremedhin, 2012). However, many producers and traders in developing countries use the world market price as a yardstick upon which their pricing decisions are based (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, national policies of economic isolation or demand-supply-driven market shocks may weaken or strengthen this relationship. Rapsomanikis

and Mugera (2011) find that the Ethiopian wheat market was cointegrated with the world market with a low speed of adjustment.

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

2. The Ethiopian wheat market

Wheat is one of the six major staple grains produced in Ethiopia. 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). Ethiopian wheat production and wheat markets are therefore mainly found in highland areas. Oromia, Amhara, Southern Nations Nationalities and Peoples Region, and Tigray are the four main wheat growing regions (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).

Dorosh and Ahmed (2009) distinguish regimes in domestic Ethiopian wheat markets which directly impact the connection between domestic and world market prices. Table 1 summarizes the relevant information from their paper. During Regime 1 until most of Regime 3, domestic prices 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.

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

	Regime 1	**Regime 2**	**Regime 3**	**Regime 4**
Start	01/2000	07/2005	04/2007	06/2008
End	06/2005	03/2007	05/2008	05/2009
Number observations in dataset	0	22	29	24
Domestic wheat prices	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
Private sector imports	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
Trade & domestic policies regarding wheat	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
Major price determinant	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	
Expected price transmission from world to market	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
Expected level of domestic prices	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
Expected volatility of domestic prices	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 rows 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.

3. Methodological approach

3.1 Cluster analysis for investigating market integration

We analyze wheat prices in seven local markets in northern Ethiopia and in the capital Addis Ababa. As already highlighted in the introduction, cointegration analysis is not feasible with the current dataset without generating artificial observations using some imputation method (i.e. filling the holes). 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 several static and dynamic aspects of the evolution of each series (e.g. mean, variance, temporal distance between extreme observations, etc.) and then apply cluster analysis on vectors containing subsets of these statistics. In this way we are able to classify the nine price series in each period with differing policy framework in terms of their similarity to each other regarding selected aspects of their statistical characteristics into clusters. This allows us to identify which price series have multivariate characteristics most similar to the ones of the world market price in a given period. The clusters are relatively homogeneous within, but differ markedly between each other. Local markets found in the world market price cluster are interpreted as being most similar to the world market price, that is, most closely related and best integrated, in a given policy regime. As with many statistical techniques, we cannot deduce statements of causality, but can flexibly assess multidimensional similarity in their characteristics.

The statistics are 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 measures a certain aspect of the particular 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 more strongly affected than others by such data problems, considering a set of such indicators will again mitigate this problem.

The central idea of the approach we suggest involves classifying or ordering the descriptive statistics of the time series in each relevant regime 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 $S_{mp} = \left(S_{mp}^{1} \; S_{mp}^{2} \; .....S_{mp}^{J}\right)'$ for the second step consists of statistics

$S_{mp}^{j}, j = 1, ..., J\ (J = 8)'$ for all time series $m = 1, ..., M\ (M = 9)'$ and all periods $p = 1, ..., P\ (P = 5)'$. 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 with respect to the remaining series. The relevant indicators can be identified based on the focus of the analysis or other considerations.

For measuring the extent and speed of price pass-through between markets, measures of price level, volatility as well as the direction of price development appear to be of largest interest. Market integration is about assessing the short-run and 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. Table 2 displays the indicators considered in our analysis and associates them to the price characteristics measured.

Table 2. Categories of descriptive statistics considered

Category	Indicator	Price characteristic measured
Price levels	Mean (C1)	Level and direction of price 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 >= mean of changes +/-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 >= 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.

3.2 Price data

We use 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 1st and 16th 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. 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 1 is a map showing the approximate locations of the various markets from which wheat prices were measured for this study.

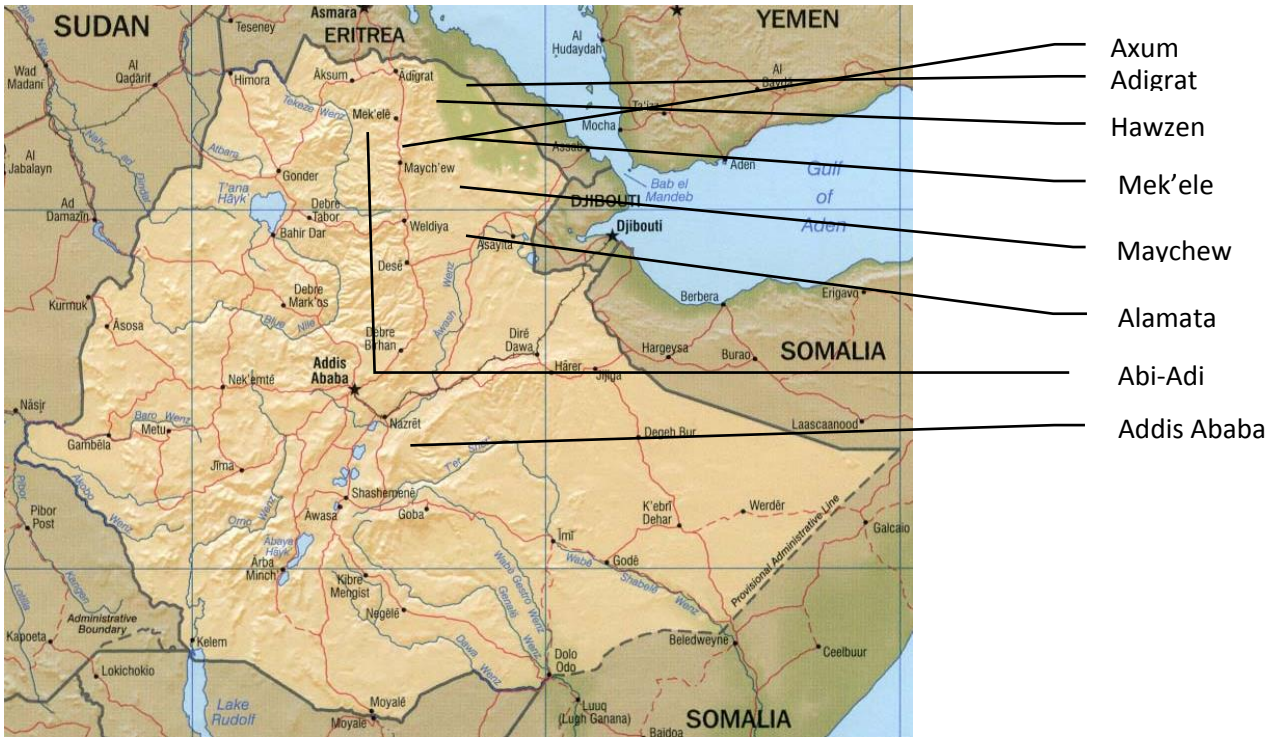

Figure 1: Map of Ethiopia indicating the local markets studied
Source: Authors based on Wikipedia (2014).

Since Addis Ababa is the capital city and also hosts the central grain market through which most traded grains pass from the point of surplus to deficit regions, 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 data available for this analysis shows many of the weaknesses and problems frequently encountered in temporally and/or geographically disaggregated time series of price data from developing countries. The data is bi-weekly and contains many missing values. Additionally, the

series of 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. Each of the eight local price series spans 86 periods. Therefore, 16% of observations of the total of 688 periods are missing.

The summary statistics of the various market prices are presented in Table 3 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 price variability, as evidenced by the largest coefficient of variation (46%). The local market with the lowest retail price spread is Axum (CV = 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%.

Table 3. Summary statistics of the price series

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

Figure 2 plots the world market price and the local Ethiopian market prices transformed into indices with the base period being 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 that the price of wheat at the world market was higher and above all local market prices. This regime (spanning from 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. In this regime, international wheat prices were lower than the domestic wheat prices. 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 also 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.

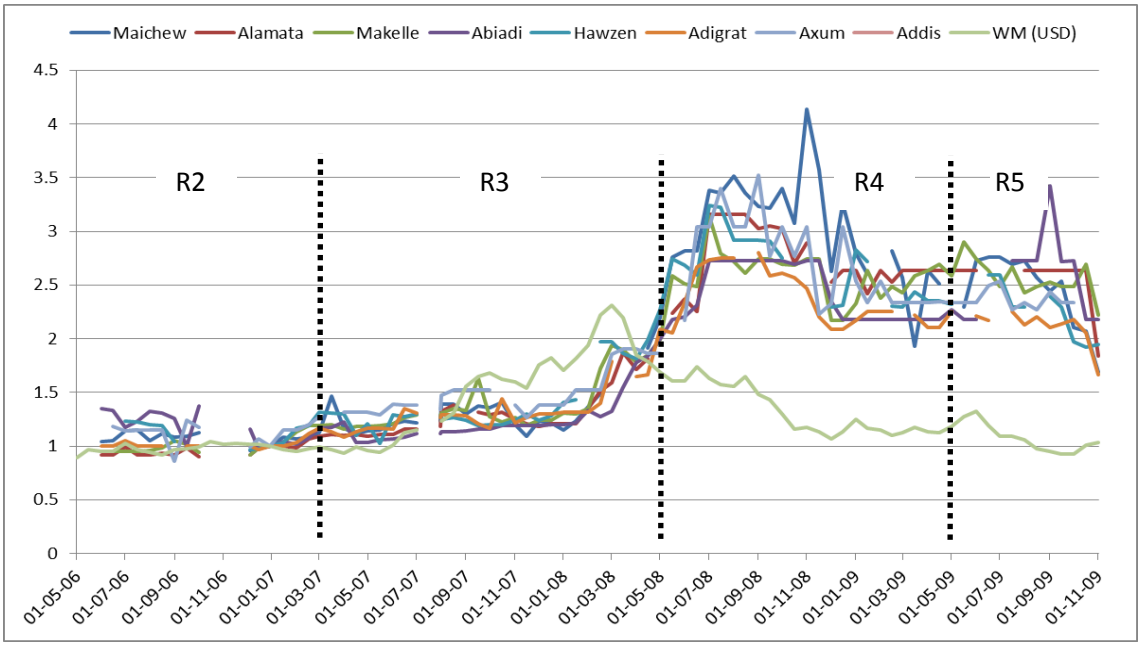

Figure 2. Indexed local Ethiopian wheat prices and the world market price

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).

4.3 Specification of the cluster analysis

Based on these ideas, our final strategy for empirical analysis is the following. We transform the price data into indices as plotted in Figure 2. Subsequently, we calculate the indicators for the eight Ethiopian price series and the world market price in USD[1] for regimes 2 to 5 and for the entire range of observations from May 2006 until November 2011. We use hierarchical agglomerative clustering (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.

[1] 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.

4.4. Results

The main output of cluster 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 relative to the vector of indicators of the world market price. In particular, we classify 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.

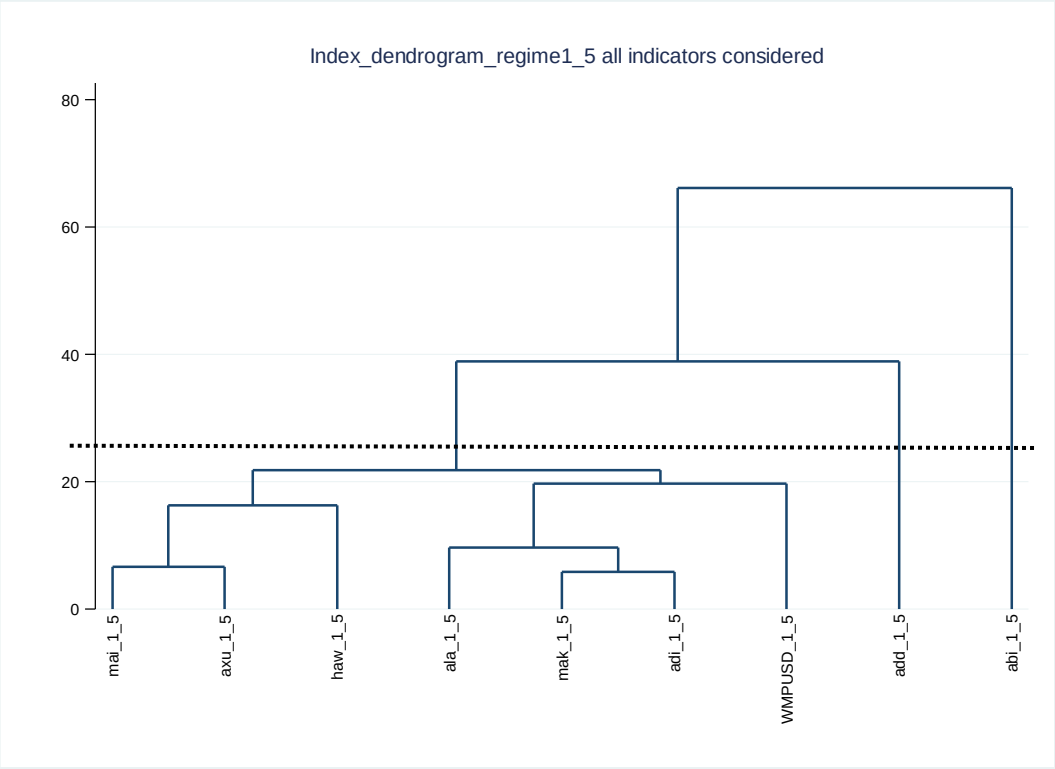

Figure 3. Dendrogram for the multivariate clustering of all indicators for the entire period

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.

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 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. Intermediate values 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 relative to the characteristics of the world market price, unity indicating maximum dissimilarity.

Table 4. Results for multivariate clustering using the entire set of 14 indicators and price level indicators

Price series	Regime 2		Regime 3		Regime 4		Regime 5		Sum		Entire period
	ALL	LEVEL	ALL	LEVEL	ALL	LEVEL	ALL	LEVEL	ALL	LEVEL	ALL
Maichew	0.5	0.0	1.0	1.0	0.5	1.0	0.0	0.0	**2.0**	**3.0**	0
Alamata	0.5	0.0	1.0	1.0	1.0	1.0	0.0	0.0	**2.5**	**3.0**	0
Makelle	0.5	0.0	0.3	1.0	0.0	1.0	0.0	0.0	**0.8**	**3.0**	0
Adigrat	0.5	0.0	0.3	1.0	0.5	1.0	0.0	0.0	**1.3**	**3.0**	0
Addis	0.0	0.0	0.7	1.0	0.0	1.0	1.0	1.0	**1.7**	**3.0**	0.5
WM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**	**0.0**	0
Hawzen	0.5	0.0	0.3	1.0	0.0	1.0	0.0	0.0	**0.8**	**3.0**	0
Axum	0.5	0.0	1.0	0.0	1.0	1.0	1.0	1.0	**3.5**	**2.0**	0
Abiadi	1.0	0.0	0.3	1.0	0.0	1.0	1.0	1.0	**2.3**	**3.0**	1
Sum	**4.0**	**0.0**	**5.0**	**7.0**	**3.0**	**8.0**	**3.0**	**3.0**			**1.5**
Number of clusters without WM	**2**	**0**	**3**	**1**	**2**	**1**	**1**	**1**			**2**

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).

The last column indicates that six of the eight Ethiopian markets belong to the same cluster as the world market price when considering the entire period without policy regimes. This result is based on the vector of price characteristics encompassing all 14 indicators mentioned in Table 2. 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. We interpret this to mean that most of the local wheat markets are generally well integrated with the world market because their price characteristics are very close to the characteristics of the world wheat price in USD.

The cluster analyses for regimes 2 to 5 separately indicate somewhat significant differences across the sub-periods defined by Dorosh and Ahmed (2009). In regime 2, when domestic prices were mainly determined by domestic supply and demand as well as profitable and significant private sector imports, the characteristics of Addis prices are very close to world market prices. This is plausible since most of imports arrive there (for trade flows see Ibidjola Agbahey et al., 2015). 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 is plausible because Hawzen is a small town of less than 6000 inhabitants in the north and the Ethiopian government massively sold highly 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 with less private sector imports and stronger governmental interference. 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 resulting uncertainty led to fragmented local markets, which slightly eased afterwards. Across the regimes, Hawzen, Mak'elle and Adigrat show least dissimilarities with characteristics of the world wheat prices because they have the smallest validate the proposed clustering methodology as a useful tool in market integration studies.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.

5. Conclusions

Our study aims at investigating whether cluster analysis can be an alternative to time series approaches when data is of poor quality. As a case study, we analyze 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. As often observed in practice for price series from developing countries, prices show many missing values and further data problems, making it impossible to perform cointegration analysis which required complete time series. Since the researcher cannot apply time series models due to the data weaknesses, she has the choice either not to perform any analysis or to look for an alternative approach which is capable to analyse the given data. This paper suggests multivariate cluster analysis performed on descriptive statistics of price series as an alternative methodology for studying market integration. The cluster analysis is conducted for the entire sample period as well as for separate market regimes in the period 2006-2009 qualitatively identified by Dorosh and Ahmed (2009).

Our results show that over the whole sample period six of the eight Ethiopian markets are clustered with the world market, i.e., have statistical properties which are very similar to the of the world market price. When the analysis is broken down into 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. This indicates that local and global markets were closely related to each other. For periods with government controlled imports and exchange rate collapse very few or no local markets were clustered with the world market. In periods with the Ethiopian government interfering with wheat markets, domestic markets are fragmented and more different from the world market. This suggests that domestic policies succeeded in decoupling domestic wheat price volatility from world markets to some extent. These results indicate that some clustering analysis can be useful tool in market integration analysis.

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

AFDB, 2010. *African Development Report 2010: Ports, Logistics, and Trade in Africa*. African Development Bank, Oxford University Press.

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.

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 Ababa, IFPRI.

EMFA (2015). *Ethiopia: Center of Regional Integration in the Horn of Africa*. Ethiopian Ministry of Foreign Affairs. Available at: http://www.mfa.gov.et/pressMore.php?pg=55. Accessed in June 2015.

Everitt, B.S., Landau, S., Leese, M., Stahl, D., 2011, *Cluster Analysis*, 5th edition, Wiley.

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.

Goodwin, B.K., Piggott, N., 2001. Spatial Market Integration in the Presence of Threshold Effects. *Am. J. Agric. Econ.* 83, 302-317.

Greb, F., von Cramon-Taubadel, S., Krivobokova, T., Munk, A., 2013. The Estimation of Threshold Models in Price Transmission. *Am. J. Agric. Econ.* 95, 900-916.

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.

Ibidjola Agbahey, J.U., Luckmann, J., Grethe, H. and B. Alamirew Alemu (2015): How do domestic policies affect the integration of Ethiopian fertiliser markets with world markets? Journal of Agriculture and Rural Development in the Tropics and Subtropics 116: 213-226.

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.

Oxfam, 2010. *The Emergency Market Mapping and Analysis Toolkit*. Practical Action Publishing.

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.

Rencher, A.C., Christensen, W.F., 2012, *Methods of Multivariate Analysis*. Wiley.

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.

World Bank, 2009. Eastern Africa: A Study of the Regional Maize Market and Marketing Costs. Report No. 49831 - Agriculture and Rural Development Unit, Washington DC.

Würriehausen, N., Ihle, R., Lakner, S., 2015. Price relationships between qualitatively differentiated agricultural products: organic and conventional wheat in Germany. *Ag. Econ.* DOI: 10.1111/agec.12137.