Academic year 2019/2020

SPATIAL MARKET INTEGRATION IN IMPROVING RURAL LIVELIHOODS: A CASE OF AVOCADO FARMING IN KENYA.

Bor, Baroness Chelangat

Promotor: Prof. Dr.Ir. Stijn Speelman

Thesis submitted in partial fulfilment of the requirements

for the joint academic degree of International Master of Science in Rural Development from Ghent University (Belgium), Agrocampus Ouest (France), Humboldt University of Berlin (Germany), Slovak University of Agriculture in Nitra (Slovakia), University of Pisa (Italy) and University of Córdoba (Spain) in collaboration with Can Tho University (Vietnam), China Agricultural University (China), Escuela Superior Politécnica del Litoral (Ecuador), Nanjing Agricultural University (China), University of Agricultural Science Bengaluru (India), University of Pretoria (South-Africa) and University of Arkansas (United States of America)

This thesis was elaborated and defended at Ghent University within the framework of the European Erasmus Mundus Joint master's degree "International Master of Science in Rural Development " (Course N° 2015 - 1700 / 001 - 001)

Certification

Gent, 4 juni 2020.

The Promoter,	The Author,
Prof. Dr.Ir. Stijn Speelman	Baroness Bor
………………………	……*B. B*……

Thesis online access release

I hereby authorize the IMRD secretariat to make this thesis available online on the IMRD website

The Author: Baroness Bor

………*B. B*………

DEDICATION

This thesis is an output of my master's degree in International Rural Development. I dedicate it with much love and lots of thanks to my family for their love, constant guidance and support throughout my education. I also hope that it will inspire my younger siblings to pursue their dreams.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude first and foremost to The Almighty God, who has seen me through my stay in Europe and South Africa while pursuing my studies.

My gratitude to VLIROUS for the award of their highly prestigious scholarship that funded this master's programme. It has been a privilege to work through the completion of this research under the supervision of Prof. Dr. ir. Stijn Speelman. Thank you very much for the timely feedback, useful insights and guidance. I would also like to thank the entire staff at the faculty of Bio-Science Engineering and the international relations' office. I cannot forget to thank the program coordinators at partner universities- The University of Pisa and The University of Pretoria for their support in my smooth transition into the second and third semesters. The objectives of this research would not have been achieved without the contribution of the Ministry of Agriculture, Livestock and Fisheries in Kenya. I specifically thank the Agribusiness Directorate for providing the data needed.

Special thanks to my family members for their love, mentorship, support and encouragement. Thank you for inspiring me to work hard. I also take this opportunity to thank my fellow classmates and the friends I made during this program; their unrelenting positive critic and assistance has enhanced the quality of this research project. I look forward to future research cooperation.

May God Bless You All!

ASANTENI SANA!

ABSTRACT

Avocado is steadily gaining popularity globally due to its health attributes. This has subsequently expanded the production area. Kenya is among the leading producers and exporters of avocado. While its exports are significantly large, its domestic market is larger. This study investigates spatial market integration within the domestic avocado markets in Kenya. The selected markets for this study are; Nairobi, Mombasa, Nakuru and Eldoret representing the urban markets while Kisii is selected as the producer and rural market. The study applies Engle and Granger cointegration technique in the analysis. Monthly time series price data in the last ten years was obtained from the Ministry of Agriculture, Livestock and Fisheries in Kenya. The Data transformation and analysis is done using Excel, RStudio and EViews. The results from the study reveal that the price series exhibit a downward trend which is likely to persist until 2022. The Engle and Granger test result did not establish a long run equilibrium to which the market prices respond. This implies that the domestic avocado markets are not cointegrated; hence segmented. In addition, there are no causal relationships found in any of the selected markets. The lowest prices in the study period are recorded in Kisii while the highest prices are recorded in Mombasa. Market segmentation implies low accessibility to markets by the rural producers. Based on these results, we conclude that market integration would go a long way in improving rural livelihoods.

Key words: Avocado, Cointegration, Rural Livelihoods, Spatial Market Integration

ABBREVIATIONS AND ACRONYMS

ADF	Augmented Dickey Fuller test
CPI	Consumer price index
FAO	United Nations Food and Agriculture Organization
KShs/KES	Kenyan Shillings
LOOP	Law of one price
Ltd	Limited
Max	Maximum
Min	Minimum
PBM	Parity Bound Model
SD	Standard deviation
TAR	Threshold autoregressive Approach
USA	United States of America
USD	United States Dollar
VAR	Vector Autoregressive
+ve	Positive
-ve	Negative
VECM	Vector error correction model
WFP	World Food Programme

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER ONE

INTRODUCTION

Agricultural markets play a crucial role in both urban and rural economies. Both producers and consumers rely on markets for sale or purchase of commodities. However, for maximum benefits markets should operate efficiently. Market efficiency has been closely linked to market integration and quite often these two terms are used interchangeably. Market efficiency implies reduced price spread between different markets. Price spread is the difference between the price that the seller receives and the price that the consumer pays at a given point in time. Price influences affordability of a commodity and impacts income hence has a major influence on poverty and food security at household level. Price is also a key indication of market functioning therefore, used as a measure of integration. Understanding market integration is key in identifying market inefficiencies and designing strategies for better functioning.

Spatial market integration is also used as a measure of market efficiency. Two markets are spatially integrated if the price difference between them is explained only by the transaction costs involved (Faminow et al., 1990). Many studies on market integration within Eastern Africa focus on staple foods, they include Gitau et al. (2019); Mose (2007) and Rashid (2004) who focus on maize, Kabbiri et al. (2016) who focuses on milk and Waluube (2009) who focused on rice . In the horticultural field, only limited studies were found: Zewdie (2017) who focused on papaya and Worako (2015) who focused on fruits and vegetables.

While staple foods remain crucial in food security, avocado stands out as a nutritional healthy fruit. Globally, there is an increasing demand for avocado due to heightened consciousness for healthy foods. In 2018, avocado imports in Europe increased by about 25% from the previous year's imports. The capacity to import more remains wide (CBI, 2020b). Likewise, in the United States avocado imports have experienced an increasing trend. Its import volume rose from about 1260 million pounds in 2013, to 2300 in 2019, implying an 82% increase over the 7-year period (Statista, 2020c).

Avocado (*Persea americana*) is believed to have originated in Mexico over 12000 years ago (FAO, 2004a). It was introduced in Kenya by the Portuguese in the 1930s (Griesbach, 2005). It is a highly nutritious fruit that is widely consumed. It contains vitamin B, C, K and E. It also contains minerals and oil with no cholesterol (Chaudhary et al., 2015). It is listed among the main fruits grown in Kenya. In recent years, there has been a growing number of farmers more interested in planting avocado as opposed to other cash crops. Some farmers for example, are shifting from coffee to avocado (Thomson Reuters Foundation, 2018).

Using time series market price data, this study will investigate market integration in avocado trade in different domestic markets in Kenya. The markets selected for the study are Nairobi, Kisii, Eldoret, Nakuru and Mombasa. They are selected due to their important role in avocado demand and supply. Kisii is a producer zone while Nairobi, Eldoret, Nakuru and Mombasa are consumer zones. The markets are also geographically separated making them ideal for this study.

This paper is organized as follows: this first chapter gives an overview of the agricultural sector paying key attention to the avocado sub-sector. It expands on the problem statement, research questions, hypothesis and justification of the study. Chapter two will review relevant studies in the field of market integration. Chapter three will discuss the methodology of this research, the theoretical framework and the conceptual framework upon which this study is built. Chapter four will consist of an analysis of the data, discussions as well as appropriate policy recommendations. The final chapter will be the conclusion.

1.1 Background Information

Kenya like other developing countries, relies heavily on the agricultural sector. Agriculture contributes to employment creation, food security and is a major source of foreign exchange income. It is estimated that about 33% of Kenya's Gross Domestic Product, 60% of informal employment and 60% of export volume comes from the agricultural sector (Government of Kenya, 2019). Given its significant role, it means that the performance of the agricultural sector is directly related to the performance of Kenya's economy. It is for this reason that the government has continuously prioritized the sector through several development strategies; for example, "Agricultural sector development strategy" (2010-2020), "Kenya's Vision 2030" (2015-2030). Recently it has embarked on a new game plan: "Agricultural Sector Transformation and Growth Strategy" (2019-2029) whose main aim is to increase small holder farmers' income, increase agricultural output, value addition and improve household food resilience (Government of Kenya, 2019).

The agricultural sector is diverse and can be classified into various categories; "Industrial crops, food crops, Horticulture, livestock, fisheries and forestry" (Oluoch-kosura, 2017). This research focuses on the horticultural sector since this sector is dominant in Kenya, it consists of many small-scale producers and offers a potential venture of wealth creation in rural areas if fully exploited; thus, improving rural livelihoods. Horticulture farming involves the production of fruits, vegetables, flowers, root crops and herbs (Ongeri, 2014). The major fruits grown in Kenya are; banana, mango, pineapple, avocado, water-melon and pawpaw listed in order of importance (Horticultural Crops Directorate, 2019).

The horticultural sub-sector in Kenya has evolved from minimal exports in the 1970s to high exports most recently (Steglich et al., 2009). It is also a major source of foreign income and ranks third after tourism and tea (Embassy of the Kingdom of Netherlands, 2017). Kenya exports its avocado mainly to the European Union and this supply continues to increase (see figure 1 below). High demand in the export markets, has expanded production of avocado in Kenya. The cultivated area increased by 7,5% from 15353 ha in 2017 to 16501 ha in 2018 (Horticultural Crops Directorate, 2019). In 2018, total income accrued from avocado was USD 59,7 million. Its importance in contributing to livelihoods and boosting economic development cannot be overemphasized. While the export market is lucrative in income generation, the domestic market remains of key importance. In 2017 for example, about 24% of Kenya's avocado production was exported while the rest was consumed by the domestic markets. (FAOSTAT, 2018).

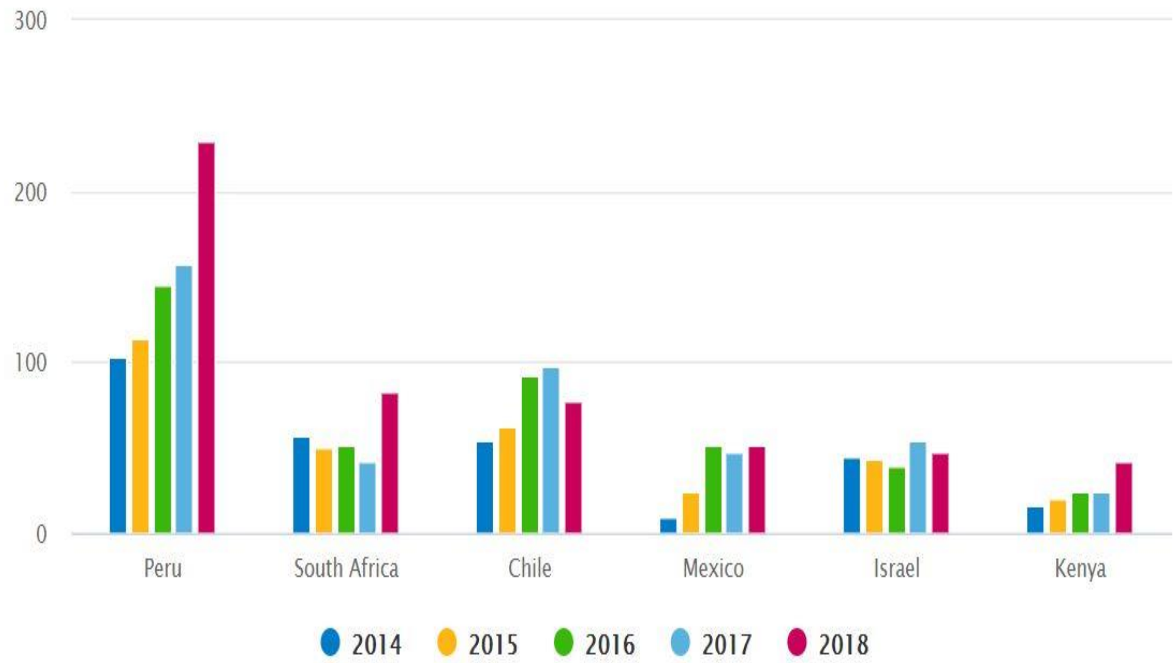


Figure 1: Lead Avocado suppliers to the EU; Source (CBI, 2020a)

Avocado farming is carried out in high altitude areas ranging between 1000-2000 metres above sea level. The optimum temperature range is between 20^{o}C to 24^{o}C with soil pH between 5-7. Kenyan farmers mainly plant the grafted varieties which take about two years after planting to start harvesting (Griesbach, 2005). The trees require minimal attention, offering the farmers an opportunity to engage in other economic activities to improve their wellbeing. The costs are negligible, since farmers prefer using manure instead of chemical fertilizers (Omolo et al., 2011). Kenya has about 16501 ha under Avocado. This represents 9.36% of total land under fruits in Kenya (Horticultural Crops Directorate, 2019).

Avocado production is done by both small-scale and large-scale farmers. The small-scale farmers account for about 85% of the total production (Wasilwa et al., 2017). There are different varieties grown; hass, fuerte, and pinkerton for export markets while puebla, duke, and g6 are produced for the local market (Horticultural Crops Directorate, 2017). The local market varieties account for over 70% of the total production (Amare et al., 2019). Production over the years has however been fluctuating (see table 1 below). This was because farmers grew the fuerte variety and local varieties which are susceptible to diseases. In a bid to address this, some farmers cut down their trees and replaced them with the hass variety (Jones et al., 2010). Other factors that affect production are weather shocks and pests (Amare et al., 2019).

Table 1: Area under Avocado and production volumes in Kenya

	Unit	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**[1]
Area Harvested	**Ha**	10320	9960	11021	11439	11583	8486	10305	16460	14497
Production	**Tonnes**	202294	149241	166948	177799	218692	136420	176045	217688	233933

Source: (FAOSTAT, 2020)

[1] The production levels in Kenya for the year 2018 were obtained from two sources (results in tables 1 and 2), there exists a slight difference in the figures obtained, this difference is however not significant in this research.

As high-altitude fruits, avocados are grown in different counties in Kenya. The lead producer counties are Murang'a, Kiambu and Kisii (see Table 2).

Table 2: Avocado Production in Kenyan Counties, 2018

County	**Volume (in Tonnes)**	**Area under Avocado (In Ha)**	**Value (in USD)**
Murang'a	123555	4321	25438736,6
Kiambu	37964	1819	6820310
Kisii	28830	1532	4295300
Nyamira	29280	1482	3093800
Bomet	10590	474	2178000
Embu	14543	709	2165250
Meru	8553	755	2099666,67
Bungoma	6028	299	2013200
Kirinyaga	5892	367	1470400
Nyeri	5784	584	1127020
Makueni	3078	335	1001875
Taita Taveta	9183	180	851299,4
Vihiga	4554	389	837050
Elgeyo Marakwet	3493	371	809500,35
Homabay	2061	299	710700
Migori	3284	315	67684,2
Nandi	2073	127	565810
Baringo	2760	202	560000
Kericho	1554	93	458850
Nakuru	1664	371	421400
Narok	1519	155	341300,09
Machakos	2280	298	339250
Others	6280	713	769641,53
Total	**318087**	**16501**	**59721044,28**

Source: (Horticultural Crops Directorate, 2019)

In the global scene Mexico is the lead producer and exporter. Kenya ranks seventh in the global production and eight in exports value (see figure 2 and 3).

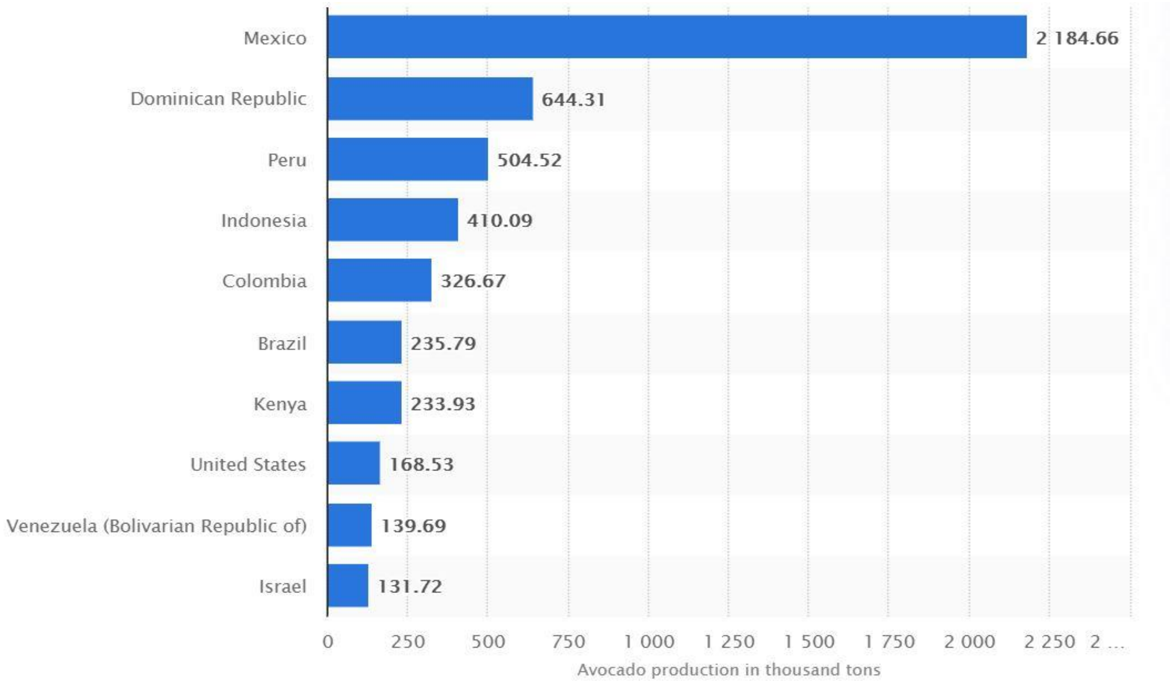


Figure 2: Avocado global production in 2018; Source (Statista, 2020a)

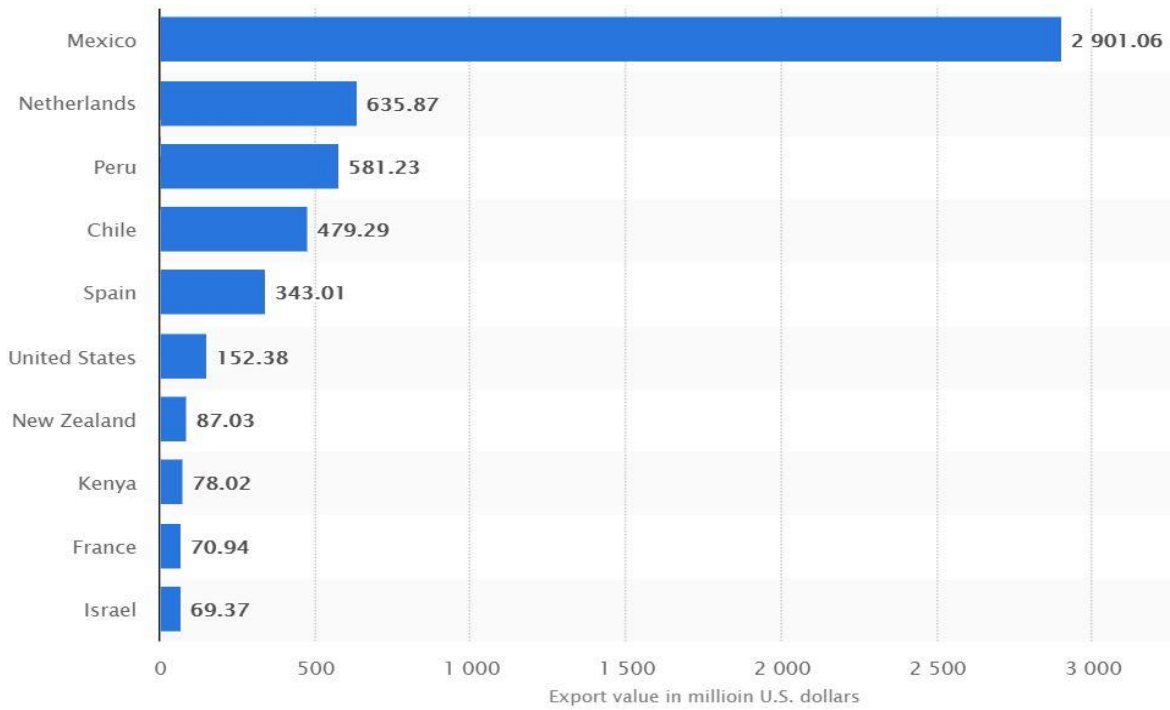


Figure 3: Global Export Value in 2017; Source (Statista, 2020b)

Locally, the avocado marketing system is complex. The small scale producers sell their avocado is sold in both formal and informal settings. The formal setting consists of large retail stores, including supermarkets that sign contracts with the farmers while the informal setting consists of farm gate sales, open air markets and intermediaries. The local marketing system reveals that middlemen have more control in the chain; supplying wholesalers, retails and other markets[2] (see figure 4).

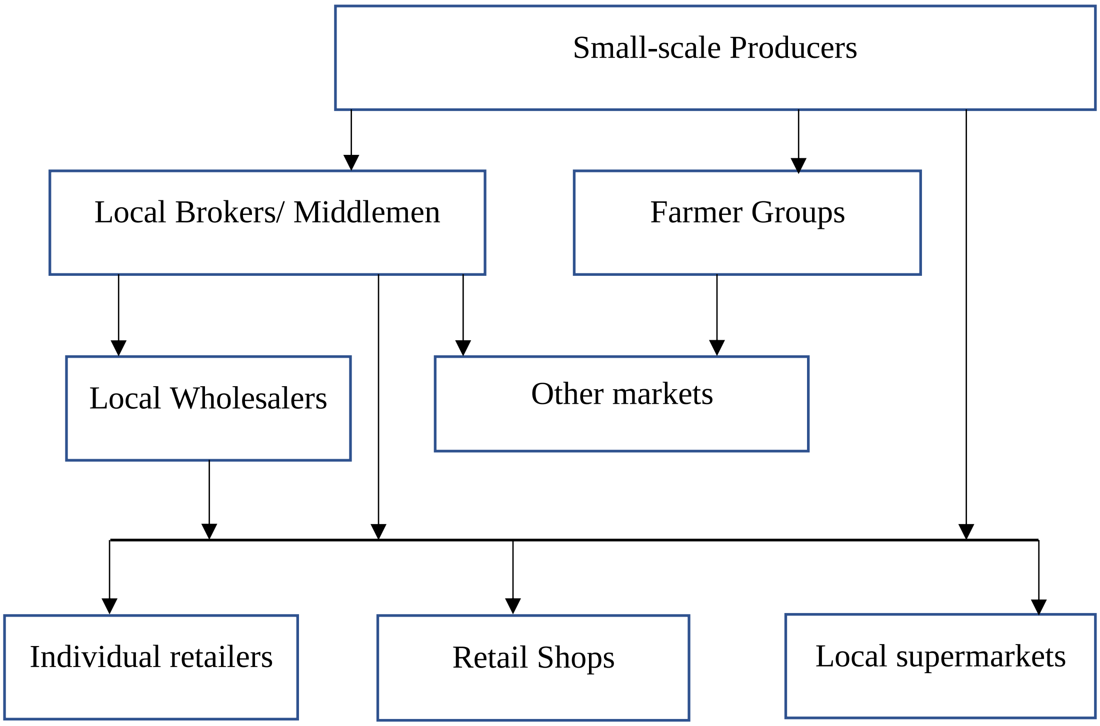


Figure 4: Smallholder Avocado marketing channels in Kenya

Source: Own illustration based on information from (Machoka et al., 2011)

[2] Other markets refer to oil processors and exporters

1.2 Problem Statement

Markets are crucial in improving livelihoods of rural people but in Sub Sahara Africa, agricultural commodity markets are often either missing or disorganized (Ashraf et al., 2008). In Kenya, the local horticultural produce market is not efficient. Access to market information remains a real challenge and hampers its development. The National Horticulture Market Information System Baseline Survey reveals that only 19% of Kenyan horticultural farmers have access to information regarding production, market intelligence, business support and legal support (Horticultural Crop Directorate, 2014). Another challenge facing this sector is the poor state of road networks which hinders market accessibility (Ongeri, 2014).

In the 1980s, Kenya liberalised its agricultural sector widening international trade. Despite this, smallholder horticultural farmers still face challenges, including quality and quantity requirements that hinder their penetration into the export markets (Mwambi et al., 2016). This makes them dependent on the domestic markets. However, the local avocado prices are not stable and face extreme price fluctuations (see figure 5). In addition, considering the perishable nature of avocado, farmers try to prevent risk of postharvest losses by selling the produce as soon as it is harvested to the readily available middlemen. The middlemen often buy at low prices.

In the recent two years, severe shortages and rising avocado prices in the domestic market have led to the imposition of an avocado export ban. In November 2019 for example, the horticultural directorate placed an export ban effective for four months (Business Daily, 2020). Likewise in 2018, a similar ban had been put in place (Business Daily, 2018). Given these efforts by policy makers to enhance the efficiency of local avocado markets, there arises a critical concern; measures put in place are short term and the problem may still recur in the long run. It therefore is needed to design long term policy measures. Since the local avocado market exceeds the export market, it is important to understand whether the local markets are integrated.

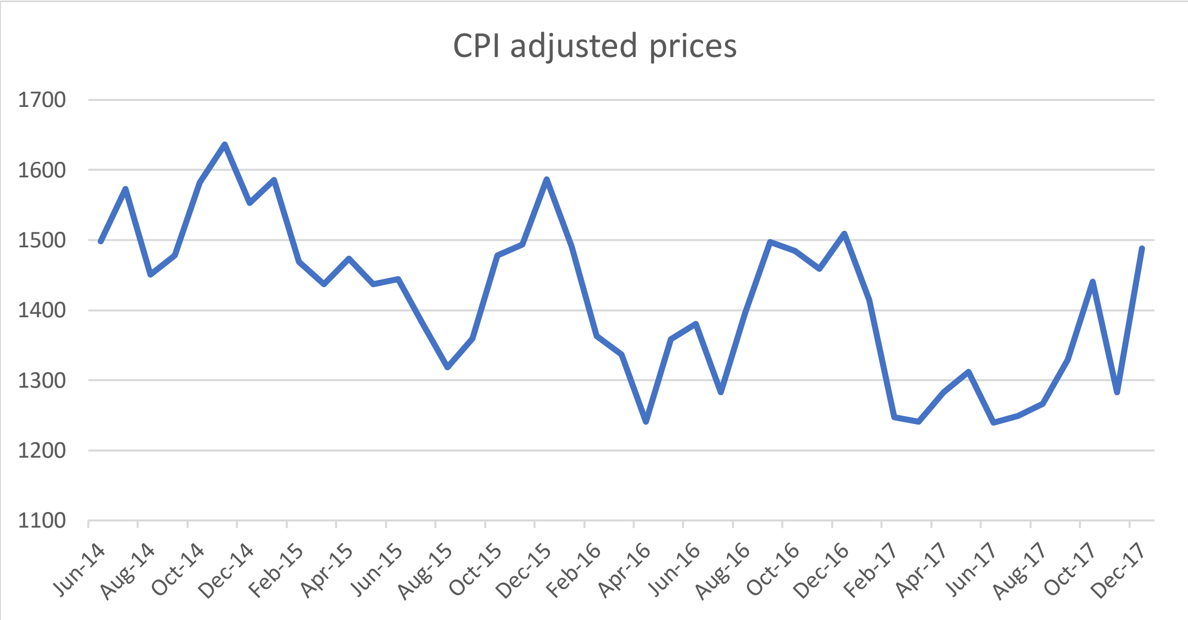


Figure 5: Trends in local Avocado prices[3]*; Source (Business Daily, 2018)*

1.3 Research question

Is the local avocado market well integrated and efficient?

1.3.1 Objectives

1. To analyse spatial market integration in the domestic avocado markets
2. To analyse the trends in local avocado market prices.
3. To determine the speed of adjustment to long run price equilibrium.

1.3.2 Hypothesis

The avocado market in Kenya is integrated

[3] These are prices for a 90Kg bag of Avocado expressed in Kenyan Shillings.

1.4 Justification for this research

Previous studies on avocado in Kenya have focused on production; Wasilwa et al. (2017), Mwambi et al. (2016) studied contract farming while Omolo et al. (2011) did an analysis of the local market channels. This therefore prompts a need for the study on market integration. This study examines spatial market integration in the Kenyan domestic avocado markets. Both rural and urban markets are selected. The focus is on how the rural market integrates with the urban markets. Avocado presents huge opportunities for income generation and food security (Hakizimana et al., 2018). Previous studies also confirm avocado's economic impact in stimulating rural development Gyau et al. (2016); Mwambi et al. (2016). Avocado is highly perishable compared to staple foods hence it is important to understand how market dynamics may influence its trade. It is for this reason that it is chosen for this research.

The study will add on to the existing literature on market integration and development of the avocado value chain. Analysis of spatial market integration is crucial since it has direct linkages to economic growth, societal wellbeing, market functioning and resource allocation (Ke, 2015). The findings of this study will provide an overview of market functioning which will be useful in designing policies and strategies for price stabilization. It will also be relevant to farmer organizations, county governments and policy analysts since it will provide critical insights in designing programs meant to improve income levels of avocado farmers.

1.5 Limitations of this Research

The focus of this study is one product (avocado) within the diverse horticultural sector. The theoretical aspect of this study is limited to spatial market integration and the law of one price. For purposes of this analysis, the study assumes that:

- Avocados that are sold locally are homogenous. This is due to the nature of the price data used in the study. The data does not specify the price for each variety but rather cumulative wholesale prices.
- Supply and demand are the main influencers of price. There are many other factors that influence price, but these are held constant in this analysis. We assume that the law of demand and supply applies meaning all other factors that affect the demand or supply are held constant. This makes price a function of demand and supply. The price thereafter determines the income that the rural farmers obtain from sale of produce. The price differences in the markets is explained only by the transaction costs.
- Avocado farmers get a substantial income from the sale of produce. We assume that the farmers produce enough avocado that has a significant influence on the household income. Low prices would mean low income levels while high prices imply a better income.
- Nairobi is the main avocado market. This is due to the fact it is the most populated city in Kenya. We also assume that many households in Nairobi are avocado consumers and their demand is significant.
- Households in Nairobi, Eldoret, Nakuru and Mombasa are net consumers of Avocado while households in Kisii are net producers of Avocados. This research assumes Kisii is an avocado surplus area supplying to the other regional markets. This is based on the production data in table 2 above.

CHAPTER TWO

LITERATURE REVIEW

2.1 Market Integration

In a perfectly competitive market, agricultural commodities move from surplus regions to deficit areas. When markets are integrated, this movement is smooth. Several authors define market integration in different ways. Earlier definition was based on price interdependence between different markets (Faminow et al., 1990). Later Barrett (2001) defined it as the smooth flow of commodities from one place to another. Further, Negassa et al. (2003) defined it based on the transmission of demand and supply levels from one market to the other. The most recent definition by Pan (2019) explains it in terms of the flow of both goods and price information between markets. There is therefore no clear definition of the term market integration. However, consensus relates it to the movement of commodities and information across space, time and form (González et al., 2001). Some factors that influence market integration include transport and transaction costs, market power, exchange rates, border and domestic policies (FAO, 2004).

Market integration can occur at two levels vertically or spatially. Vertical integration occurs mainly between international markets and domestic markets while spatial integration occurs within domestic markets. Another form of market integration is cross commodity integration which refers to integration between different commodities. The focus of this study is spatial market integration. Theoretically, spatial integration models imply that if two markets are integrated, excess demand or price shocks will have the same effect on prices in both markets (Jena, 2016). The analysis of spatial market integration addresses concerns such as causality patterns, long run equilibrium attainments and the dynamic interaction between markets that are geographically separated (Zewdie, 2017). It has been linked to the law of one price (Barrett, 2001). This means that price differential between two markets is explained by transaction costs and any changes in demand or supply in one market will affect prices in the other market until equilibrium is regained through spatial arbitrage (Rapsomanikis et al., 2006).

Arbitrage is the situation of buying and selling commodities in different markets to take advantage of price differences between them (Fackler et al., 2001). The law of one price forecasts that arbitrage restores equilibrium prices in spatial markets. Consider two markets; market A and market B, both producers and consumers of avocado. If market A, encounters calamities such as crop disease, which reduces the production levels, then prices automatically increase in this market. Traders from market B are inclined to sell in market A provided that the price in market A is greater than the transaction costs incurred. As a result, the supply in market B will reduce, hence increasing the price. This continues until prices in both markets are at equilibrium making the trader profit zero. In this situation, inter-market arbitrage exists making traders indifferent to trade in either markets (Barrett et al., 2002). When markets are not integrated, they are said to be segmented (Jena, 2016). Market segmentation may send wrong price signals and negatively influence production and marketing decisions (Haji, 2014).

2.2 Market Integration Analysis

Market integration as discussed above can be used to indicate how price changes are transmitted from one market to the other. A reduced time lag indicates better arbitrage and is a proof of better functioning markets. Market integration analysis helps diagnose some challenges that face agricultural marketing (Rashid et al., 2010). There are various methods that are used in measuring market integration. The first is the static method, the second consists of dynamic methods such as granger causality tests, Ravallion, Timmer models and co-integration technique. The dynamic models account for the dynamic nature of prices. The final method consists of switching regime regression models. They include; the error correction model, Parity bound model (PBM) and threshold autoregressive model.

2.2.1 Static models

Static models are based on the existence of a correlation between time series price data. One such test looks at the bivariate correlation which measures the co-movement of prices in two different markets. It assumes fixed transaction costs and calculates the correlation coefficient. In perfectly integrated markets, the coefficient should be equal to one and in segmented markets tends towards zero. This technique is relatively simple, only requiring price data series for two markets. This method was later critiqued as being biased as results may be influenced by inflation, common trends and seasonality in agricultural marketing.

Further, the model does not capture the dynamics of a marketing system meaning it cannot account for trade reversals (Negassa et al., 2003). It is also ineffective because only pair-wise comparisons were possible and still it was not possible to account for the direction of price transmission. Another backlash is that it does not factor in the non-stationarity nature of price series data (Goletti et al., 1995). These shortcomings are addressed by the dynamic models.

2.2.2 Dynamic models

Due to the numerous short comings of static models, in the 1980s economists developed dynamic models. These models account for the dynamic nature of prices and transaction costs. An example is the Ravallion model which assumes there are several regional markets, but a central market exists. A major critique on the model is the fact that it fails to acknowledge transaction costs. An elaborated version of the Ravallion model is the Timmer model which assumes that the central market is predetermined. Quite often the central market is an urban market which then connects to rural markets. The central market influences trade and price formation making it the price leader (Ravallion, 1986). There is however no clear methodology for the selection of the main market. Another dynamic model is the cointegration technique which assumes the existence of a long run equilibrium. It is extensively discussed in the methodology section below.

2.2.3 Switching Regime Regression models

The switching Regime Regression Models are based on the realisation that price relationships may be non-linear. The PBM analyses market integration at three regimes; at the parity bound where price differentials equal transactions costs, inside the parity bound where parity bound is less than transaction costs and outside the parity bound where parity bound is greater than transaction costs (Sunga, 2017). These bounds determine the price efficient zone. The probability estimates for the three regimes are then calculated. These three regimes are represented by equation 1, 2 and 3 respectively.

$$P_{it}-P_{jt}=TC_t \text{………… equation 1}$$

$$P_{it}-P_{jt} < TC_t \text{……………equation 2}$$

$$P_{it}-P_{jt} > TC_t \text{……………. equation 3}$$

Where;

P_{it} *is Price in market i at time t*

P_{jt} *is Price in market j at time t*

TC_t *is Transaction costs at time t*

In the third regime, arbitrage is not fully exploited implying inefficiencies which could be a result of trade restrictions, price support or non-competitive pricing mechanism (Hillen, 2019). When production and consumption are specialised, only regime one is consistent with market integration. On the other hand, when it is not specialized both regime 2 and 3 can imply market integration. The non-specialisation is more common in developing countries (Penzhorn et al., 2002). The model factors in market dynamics which include transaction costs, trade reversal and autarky situations (Negassa et al., 2003).

The PBM can be used to analyse the effects of different market policy regimes. In order to develop the model, data on transaction costs are needed. In the case of developing countries, these data may be difficult to obtain. Some studies estimated the transaction costs by calculating the transport costs between two markets and inflating it to cater for other transactional costs such as information costs, negotiation costs and monitoring costs. A major weakness of this analysis is the fact that it handles only limited markets and short run deviations may be interpreted as inefficiencies whereas they could be as a result of rational trader behaviour.

Threshold autoregressive Approach (TAR) is another switching regime model which was developed as an improvement of the PBM model. The TAR model tests for market integration within a framework of switching regime. This means that the relationship between variables is non-linear and that equilibrium price is not constant. It does not rely on transaction costs data, but factors in their effects by creating a threshold band within which market prices are not linked (Negassa et al., 2003). This threshold band is as a result of transaction costs. The wider the band, the greater the price volatility.

Results obtained from a TAR model display the probability of being outside the threshold. This probability is the measure of market integration. In addition, the period needed to eliminate the violations is obtained (Abdulai, 2000). However, its crucial weakness is the fact that it assumes fixed transaction costs (Fackler et al., 2001). The argument for this however is that it is best to assume constant costs rather than to ignore them. Another challenge in its application is the cumbersome computational requirements (Sunga, 2017).

2.3 Market efficiency

Market efficiency has been associated with market integration, prompting the need to elaborate their association. Spatial market efficiency occurs in situations where arbitrage condition is fully exploited resulting in equilibrium while spatial market integration refers to the transmission of price variations between markets (Negassa et al., 2003). Another distinctive aspect is the fact that market integration is associated with physical trade of commodities while market efficiency is associated with the trade partners (Hillen, 2019). However, it is important to note that in the presence of a full arbitrage situation, the markets are said to be integrated. This therefore means that integration and efficiency are related but not equivalent and tests on integration are not an accurate reference for market efficiency (Barrett et al., 2002). Rather efficiency is just a pre-condition for integration (Federico, 2007). Spatial market integration can therefore be considered as an indirect measure of market efficiency (Faminow et al., 1990).

Market efficiency can be defined in three different ways. The first type relates it to maximizing output per unit input. The second relates it to a competitive or effective market structure while the last relates it to reduced marketing margin/ price spread. The third view relates closely to market integration. A reduced-price spread is a precondition for the law of one price. A low-price spread implies that both consumers and producers gain a reasonable price or profit. The law of one price is one of the existing conditions for an efficient market, and its violation implies an instance of inefficiency (Federico, 2007). The law of one price is also a testing framework for market integration.

2.4 Rural Livelihoods

About 80% of Kenya's population reside in rural areas. Amongst them, 70% rely on agriculture as the main livelihood source (Government of Kenya, 2019). Many rural households are engaged in small-scale agriculture with little entrepreneurial dominance. Predominantly, the rural agribusiness sector is dominated by cash crops and staple foods. However, horticultural production has huge economic potentials that could revitalise rural economies and contribute to poverty eradication (Hakizimana et al., 2018). Many rural households who engage in horticultural farming depend on it for food security and income. Horticultural Crop Directorate Authority (2014), report states that the sector is rapidly expanding but despite this, there still exists high poverty prevalence in rural areas.

Poverty is multidimensional in nature cutting across different aspects of human lives. However, the most common definition of poverty is based on income level demarcated by a poverty line threshold which currently stands at USD 1,9 per capita per day. Poverty eradication is a top priority and many governments are keen on it. Economic development is known to assist in poverty alleviation. Market integration is one of the ways to boost economic development (Varela et al., 2016). It has spill over effects which benefit the non-farm sector paving way for economic growth (Söderbom et al., 2009). This is especially useful for developing countries which heavily rely on the agricultural sector.

Market integration particularly plays a crucial role in improving rural livelihoods in the sense that if markets are integrated, farmers incur more profits by selling their produce where it is much more expensive (Jena, 2016). This situation continues until prices level up making both consumers and sellers better off. Given the seasonality and secular trends of agricultural marketing, farmers do not always have a choice on where to sell, integration gives them the opportunity to have options.

The horticultural sector has specifically been of interest in poverty eradication due the numerous income opportunities that it presents. Previous studies that linked livelihood improvement to the horticultural sector include Gyau et al. (2016), who investigated the factors that influenced participation of avocado farmers in group action. The results revealed that age, education, gender and farmer perceptions in knowledge and technology had a huge influence in their participation. Age and education positively influenced participation. They also found that female participation was lower compared to that of male counterparts. Farmers who perceived groups as a channel of getting information had a higher participation rate. They conclude that education of farmers is crucial in enhancing group participation.

The majority of horticultural produce in Kenya comes from small-scale farmers implying its high ability to revitalize rural areas through jobs and wealth creation. Ongeri (2014), investigated the impact of input prices on the livelihood of small holder horticultural crops in Kenya. His results indicate that input prices were high and significantly impacted the economic wellbeing of the farmers. He recommends implementation of policies to cushion against this and improvement of the road network. Another study is by Mwambi et al. (2016), who investigated the role of contract farming in improving farmers' income. Their results indicate that the participation of avocado farmers in contract farming was insufficient in improving household income. The low participation rate was attributed to mismanagement of the farmer groups leading to dismantlement.

Varela et al. (2016) in their study "Market integration and poverty"; estimated the degree of integration in seven food markets and two energy markets in South Sudan. The foods selected were sesame, millet, maize, sorghum, wheat flour, beans and groundnuts; while gasoline and diesel represented energy. The study utilized monthly price data for the period 2009 to 2013. Results indicated that the markets were highly segmented, this situation was worsened by trade restrictions. The study also revealed that prices were highly volatile and transport costs explained about 33% of food price variations. They also noted a positive relation between fuel prices and food prices, a 1% increase in fuel price caused a 0.4% increase in food price. Finally, they conclude that an improved road network would increase market integration, welfare gains and reduce poverty rates by up to 4%.

2.5 Empirical Studies

There are several studies carried out in the field of market integration. They are discussed below within the framework of the methodology used in the analysis.

2.5.1 Static models

This was the earliest methodology of testing for market integration widely used in the 1970s and early 1980s. It is based on price correlation among time series data. Although it has been numerously critiqued, the model can be applied as a preliminary test in application of other methods (World Food Programme, 2007). In addition, it has been used independently to assess market power. A study by Morlacco (2017), analysed market power in input markets in the French manufacturing sector. She estimated this using production data and the firms' buying power. Results indicated positive correlation existing between the firms' production volumes and input volumes. The power was greatest in markets for food, rubber and machinery.

2.5.2 Dynamic Models

Zewdie (2017), studied "Spatial market integration and price transmission for papaya in Ethiopia" in the period 2002 to 2014. He applied the cointegration technique to analyse five local markets in Ethiopia. His results indicate that four of the markets were cointegrated. The rate of price adjustments among papaya prices was slow and differed from market to market. The Arbamnch market removed about half of price disturbances improving disequilibrium situations. A second study within the Ethiopian context was by Worako (2015), who analysed fruits and vegetable markets in the period 2008 to 2015. The analysis involved retail prices from twenty-one markets within Ethiopia. His results indicated that cointegration existed in some markets. Price transmission from the central market (Addis Ababa) took between three to seven months to be transmitted to other markets.

Another study is by Oktarina (2015), who analysed five rice markets in Indonesia using the cointegration technique. He used monthly retail prices in the period 2004 to 2009. The results show that the markets are integrated and three of the selected markets responded quickly to price variations. One more study that applied the same technique was Rahman et al. (2018), who did an integration analysis of fed cattle in the USA in the setting of a mandatory price reporting era. They used weekly prices for steer and heifers in the period 2001 to 2015. The sample consisted of five major markets in the USA. Results indicate that all markets were cointegrated. They also found that a causal relationship existed among most steer markets, but this was not the case for heifer markets.

A different study by Zakari et al. (2014), also applied the cointegration technique to analyse market integration and price transmission within a vertical framework including both international and regional grain markets. International markets selected for this study were Nigeria, Mali, Vietnam, Burkina- Faso and Togo while the regional market selected is Niger. The grains selected were millet, sorghum, maize and rice. These are staple foods in Niger and the local demand is often supplemented by imports. Results indicated that the different markets responded differently to the long-run equilibrium. The speeds of adjustment were 30%, 35%, 48% and 40% respectively and reveal the degree of integration. All international markets revealed significant price transmission to the local Niger markets, meaning price variations in the international markets had a significant impact on the local markets.

Ahmed et al. (2017), also applied co-integration technique in their study of market integration and price transmission in the onion markets of India. They used monthly wholesale price data from six major markets in India in the period 2006 to 2014. The results indicated that four of the six markets were co-integrated, both unidirectional and bidirectional causality existed between prices in the different markets. They also found that price variations were quickly transmitted to all the other markets except Mumbai and Kozhikode.

Further, Acquah et al. (2012), applied the cointegration technique in their analysis of three plantain markets in Ghana in the period 2008 to 2010. Ghana is a major producer and consumer of plantations; therefore, the crop is crucial in the country's food security. The results from their study indicate that the plantain markets are integrated. They also found that about 7% of price disequilibrium was corrected within a week and about 16% disequilibrium corrected within two weeks.

In addition to the above studies; Ghosh (2011), also applied the cointegration technique in spatial analysis of food grain markets in India. His study focused on two periods; pre-liberalisation and post liberalisation. In the post liberalisation phase, there was a reduction of government intervention in agricultural commodity marketing. He used extensive data of rice and wheat prices stretching across the periods 1984 to 2006. His results revealed that the rice markets were not integrated during the pre-liberalisation period. However, thereafter they are strongly integrated. Likewise, for rice; the markets during the pre-reform period were segmented, but this improved remarkably well post reform. He concludes that this improvement results from the agricultural reforms introduced post liberalisation.

2.5.3 Switching Regime Regression models

Mtumbuka et al. (2014), applied the Threshold autoregressive Approach (TAR) in the analysis of nine bean markets in Malawi. Their findings revealed that the bean markets were cointegrated, but price transmission was uneven. Poor infrastructure was deemed as the main cause. They concluded that the bean markets in Malawi are integrated. Tsiboe et al. (2016), applied TAR to analyse eight rice markets in Liberia in the period 2009 to 2014. About half of rice consumed in Liberia is imported. Their study investigated markets in the ports of entry. The results indicate that local rice markets are integrated. Positive and negative price changes were transmitted symmetrically in all markets within a period of five months. Both studies conclude that improving transport and market infrastructure would go a long way in improving market integration.

Gitau et al. (2019), applied the vector error correction model (VECM) to investigate integration among nine domestic maize markets in Kenya under four regimes. The first regime was liberalization of the agricultural sector. The second was the fertiliser subsidy regime, the third was import bans for food commodities with genetically modified organisms while the fourth was zero rating of import tariffs. He used wholesale monthly prices cutting across all regimes. The period selected in his study was 2000 to 2016. His results indicate that price spread was highest in regime one and that all selected markets were cointegrated. He also notes that regime one had little policy interventions consequently the forthcoming policies distorted the maize markets. He therefore concluding that policy formulation and implementation requires more consultation and coordination to achieve the desired results.

Some studies applied more than one technique. One such is Kabbiri (2018), who applied both Ravallion and TAR techniques to analyse dairy markets in Uganda from 2000 to 2011. His analysis was limited to sixteen local markets in Uganda. He performed exogeneity tests to identify a central market for milk trade in Uganda. To assess the degree of price transmission, he developed a TAR model using weekly prices. The results indicated that Mbarara is the central milk market of Uganda, but not Kampala which is the country's economic hub. He found that price differentials between the main market and other markets stabilized within 2-3 days.

Another study that applied both techniques is Chalmers et al. (2019), who analysed market integration of two dairy products powdered milk and liquid milk in Malawi. They used monthly data in the period 2006 to 2011. Their results indicate that there exists a long run price relationship in the selected markets in northern and southern Malawi. Although most of the dairy processing units are in the South of Malawi, they adequately supply the northern part and transaction costs between the markets was not a concern at all; hence, there was no need to construct processing units in the north. He concluded that prices for the selected dairy products in Malawi are highly similar implying an integrated market.

Negassa et al. (2003), proposed an advancement of the parity bound model to understand price dynamics within different regimes of marketing policy. This contribution is significant because it shows whether the changes resulting after implementation of a policy are statistically significant and calculates the duration needed for a marketing policy to take effect in spatial markets. One such study that applied the methodology is Hillen (2019), in her study 'market integration and market efficiency under seasonal tariff rate quota'. Her study is based in Switzerland following the imposition of a tariff rate quota on fruits and vegetable imports during summer.

Hillen used wholesale weekly data from March 2011 to May 2015; to analyse trade between Italy and Switzerland. The transaction costs were calculated from estimation of transport and marketing costs. The results indicate that when the quota is in place, market inefficiency increases and hence market integration. However, traders benefit from large rents. The rents are allocated based on historical purchase hence of convenience to large existing traders and not new players in the market. She concludes that the quota maintains status quo for both importers and retailers in Switzerland. However, buyers end up paying a higher price.

2.6 Summary of Literature Review

From the above illustrations, there exist several studies within the field of market integration. Both wholesale and retail prices are used in these analyses, this primarily depended on data availability. One main necessity for market integration analysis is the availability of time series price data. Most studies used monthly price data. This is although Amikuzuno (2010), observed that the use of high frequency data such as daily or weekly prices would generate more precise results, especially in estimating the price adjustment parameter. However, in developing countries such detailed data is often unavailable. The most frequently used market integration analysis method is the cointegration technique. Kabbiri et al. (2016), made an overview of market integration studies globally and found that it was the most populous technique.

This research applies the same technique and builds on the above-mentioned studies in the following ways. Firstly, it will investigate the degree of market integration between Kisii- a purposely selected rural market with high poverty rate and four other urban markets. To the best of my knowledge, it is the only existing study on market integration analysis of avocado. It also applies the most recent data highlighting the situation of market integration in the Kenyan avocado sector as it is currently.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This section describes the methodology used to achieve the objectives of this research, the theoretical framework, the conceptual framework, the data needed and the source. It also describes the targeted study area.

3.2 Study Area

The study is limited to five domestic markets in Kenya that is; Nairobi, Eldoret, Mombasa, Nakuru and Kisii. These markets are selected due to data availability and their importance in avocado demand and supply. They are also geographically separated (see figure 6 below).

Figure 6: Map of Kenya showing the study areas; Adapted from (Elimu.org, 2020)

Nairobi is the most populous county in Kenya hosting about 4,3 million people, **Nakuru** hosts 2,1 million, **Mombasa** hosts 1,2 million, while Uasin Gishu county where **Eldoret** is located hosts about 1,1 million individuals (Kenya National Bureau of Statistics, 2019). These four regions are selected for the study because of their high population densities implying a huge market for agricultural commodities. In addition, the data available for these markets was also relatively complete, with only few missing values.

Kisii county is in the South West of Kenya. It hosts a population of 1,1 million people and has high attitude climate ranging between 1000 to 1800 m above sea level. It is generally cool and wet all year round. It has two rainy seasons; the long rains which come between April and June and the short rains which come between September and November. Agriculture is the main economic activity in the county; therefore, many households engage in it. About 78% of the county's land is arable with 58% being under crops (Omiti et al., 2009).

Kisii is purposely selected because of its high avocado production. In 2019, the county produced about 28000 tonnes of avocado ranking it fourth in Kenya after Murang'a, Kiambu and Nyamira counties (Horticultural Crops Directorate, 2019). At the same time, the poverty prevalence in the county is high. It is estimated that about 41% of the county's population live below the poverty line (Kenya National Bureau of Statistics, 2018). Its selection for this study is ideal since it has high production levels and poverty rates; both typical characteristic of rural areas. This research assumes that Kisii is the main producing area supplying to the other four selected domestic markets. The distance between these markets is illustrated in table 3 below.

Table 3: Distance between producer and consumer domestic markets

Market 1	**Market 2**	**Distance in Km**
Kisii	Mombasa	803,5
Kisii	Nairobi	321
Kisii	Nakuru	196,7
Kisii	Eldoret	195,1

3.3. The Theoretical framework

This research is based on the law of one price (LOOP). The law states that, if markets are operating efficiently, a homogenous product sold in different locations sells at the same price in the same currency, the price difference if any should be explained by transport costs (Rapsomanikis et al., 2006). This is illustrated by the general formula;

$$P_{it} = P_{jt} + T_t \text{ …………………} \textit{equation 4}$$

Where;

P_{it} is the price at location i at time t

P_{jt} is the price at location j at time t

T_t is the transport and other transfer costs

The law holds if there are many sellers to ensure competition, if the commodity in question is standardized and the transportation costs are low (Rashid, 2007). However, the law of one price does not hold in all situations. This is due to short- run market dynamics which cause deviations from equilibrium. These deviations indicate market inefficiencies which could possibly be resolved in the long run. Market integration analysis draws upon this theory of LOOP, in the sense that although price deviations exists, in the long run the prices converge. These price movements in the spatial markets are in the same direction. This co-movement and long run relationship is tested by co-integration tests.

In normal situations, the law of one price indicates that the price difference between the two spatial markets should not exceed T_t (Baltzer, 2015). This situation represents a weak form of the LOOP (Fackler et al., 2001) and is illustrated by equation 5 below

$$P_{it} - P_{jt} \leq T_t \text{………} \textit{Equation 5}$$

Where;

P_{it} is the price at location i at time t

P_{jt} is the price at location j at time t

T_t is the transport and other transfer costs

If equation 5 holds true, then domestic markets are said to be integrated. The LOOP has argumentatively been challenged; Barzel (2007), argues that the law assumes that information is costless. He then proposes its replacement with the price convergence law which factors in information costs. However, for this analysis, the LOOP still applies since the focus is in the long-run and the data on information costs was unavailable. The co-integration test is therefore applied as the measure of market integration. Another test for market integration is price transmission which draws inference from LOOP. Price transmission occurs when price disturbances in one market are automatically transmitted to other markets (Rapsomanikis et al., 2006). Just like LOOP, price transmission can fully be tested in the long-run period and not the short-run.

3.4 The Conceptual framework

This research applies cointegration technique to test for integration. It involves three processes; the first being a test for stationarity, the second is the test for a long-run relationship while the last is the estimation of the error correction model. Whether markets are integrated or segmented, forms a crucial basis for policy formulation. The framework used in this research was developed based on ideologies from the World Food Programme (2007) and Rapsomanikis et al. (2006). It represents a systematic flow of this analysis to address all the objectives of this research. The first step of the framework is analysing the price series data to identify common trends. Thereafter, a cointegration analysis is applied. Cointegration enables the identification of short-run relationships between the variables and the estimation of the long-run equilibrium (Vargova et al., 2018). The framework concludes with policy implications which is based on the results from the study. It is illustrated in figure 7 below:

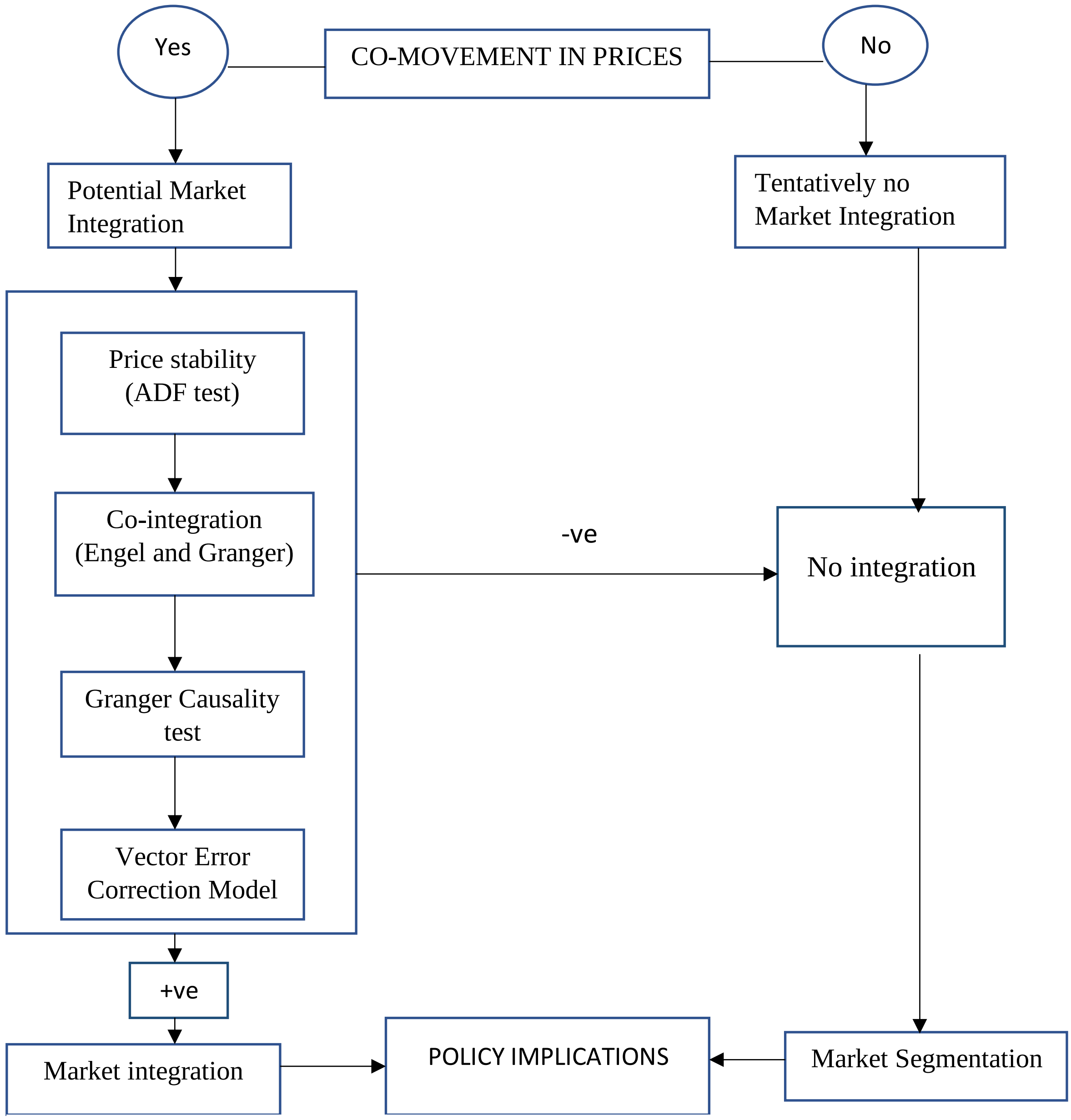


Figure 7: The conceptual framework

3.5 Co-Movement in Prices

Co-movement shows the relationship between the price series data. To verify it, a correlation analysis is performed on all the variables. The analysis is conducted in RStudio using the package *Hmisc*. The guiding hypotheses for this test are:

H_o .... There is no correlation in the price series data

H_1....... There is correlation in the price series data

3.6 Analytical methods

3.6.1 Unit Root Tests

Unit root tests are used to test for stationarity and non-stationarity. A stationary series is one in which price disturbances generated in the series are time independent. This implies that the variance and the mean and do not vary over time. Stationarity tests are vital since they determine the order of integration. This is the number of times that a series is differenced to make it stationary. These tests form an important precondition in analysis of time series data since these data exhibit trend components (Acquah et al., 2012). When a series is non-stationary it requires transformation to avoid errors in estimation. Tests that are used to check for stationarity include the Augmented Dickey-Fuller test (ADF) and the Phillips Perron test. However, the Phillips-Perron test is not effective when the error series follows a negative moving average process, therefore it is recommended to perform the ADF test (Jena, 2016).

Augmented Dickey-Fuller test

The Augmented Dickey-Fuller test was developed as an advancement of the previous Dickey-Fuller test. It aimed at addressing the endemic serial correlation of the error term in the later. It includes lagged differences as additional regressors in the model to clean up existing correlation if any (Neusser, 2016). It is presented by the general formula:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \Sigma^{m}_{i=1} \alpha i \Delta Y_{t-i} + \varepsilon_t \text{.........} \textit{equation 6}$$

Where;

Y_t is the price series of a given market

t is the length of time series

β_1 is constant

β_2 is the coefficient of linear time trend t

δ is the parameter of interest

Y_{t-1} is the price of a given market in the previous year

m is the number of lags to be included in the model,

α_i is the coefficient of the ΔY_{t-i}

ε_t is a random error term

The variable m should be large enough to ensure that the error term is a white noise. The error term should also be uncorrelated to any regressor in the model. If the Y_t values follow an unpredictable trend it means that β_1 and β_2 values are 0, reducing equation 6 to;

$$\Delta Y_t = \delta Y_{t-1} + \Sigma^{m}_{i=1} \alpha i \Delta Y_{t-i} + \varepsilon_t \text{………..} \textit{ equation 7}$$

The guiding hypotheses for this test are:

H_o ……. Price series data are non-stationary

H_1……. Price series data are stationary

When ‘Y’ is statistically indifferent from zero, the series has a unit root thus is non-stationary. Stationarity means the price series pairs move in the same direction in the long run. The next step is then to check the direction of this movement. The ADF test is performed using EViews version 9. The test is performed at three levels; intercept, at intercept and trend and no trend or intercept. They are represented by equation 8, 9 and 10 respectively.

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \alpha i + \varepsilon_t \text{………..} \textit{ equation 8}$$

$$\Delta Y_t = \beta_1 + \beta_2 + \delta Y_{t-1} + \alpha i + \varepsilon_t \text{………..} \textit{ equation 9}$$

$$\Delta Y_t = \delta Y_{t-1} + \alpha i + \varepsilon_t \text{………..} \textit{ equation 10}$$

3.6.2 Co-integration tests

The co-integration tests are done to verify the law of one price. It hypothesizes that although price disturbances may occur in the short run, a long run equilibrium exists. This means that any dynamics of the price relationships in co-integrated markets converge in the long run towards the LOOP. Approaches used in co-integration analysis are the Engel and Granger

(1987) and Johansen (1990) vector autoregressive (VAR) approach. The former is used in bivariate analyses, the later in multivariate analyses (Haji, 2014). These tests are a follow-up of the Unit root tests. The Engel and Granger test has an advantage over the other tests in that it is less sensitive to distribution assumptions (Yoon, 2017). It is also the simplest cointegration technique. This research applies Engel and Granger test since the focus is on how the rural market chosen (Kisii) connects to the urban markets.

Engel and Granger Test

The Engel and Granger test uses stationary data for testing the long-run relationship. This technique enables the researcher to understand the extent and direction of market integration. The technique analyses spatial market integration in a context of short run price deviations that level up in the long run (Castillo et al., 2015). The guiding hypotheses for this test are:

H_o ……. There is no co-integration

H_1……. There is co-integration

The co-integration regression equation to be estimated is (Abdulai, 2006);

$$P_t^{1} = \alpha_0 + \alpha_1 P_t^2 + U_t \quad \text{……..Equation 11}$$

Where

P_t^1 *is the price at regional market at time t*

P_t^2 *is the price at the central market at time t*

α_0 *is a constant term*

α_1 *is the co-integration parameter*

U_t *is the random error term with constant variances/ price spread*

Granger Causality

The granger causality test is used to show the association between time series variables. It also determines how current and past prices in a market influences future prices in another market (Jena, 2016). It analyses all possible pairwise combinations in the selected markets to identify the existing relationships. There are two directions which causality can follow that is unidirectional and bidirectional. Unidirectional occurs when one variable influences the other while bidirectional occurs when both variables influence each other (Zewdie, 2017). If both situations lack, then there is no causality relationship.

Prices in one market are regressed on the lagged values of the other market, this means that if two markets are integrated, the prices in one market influence the prices in the other market. It represents a lead and lag situation, but does not explain the nature of price relationships (Fackler et al., 2001). It handles pairs of time series variables and determines the leader and the follower role of the markets (Chalil, 2016). This assessment is done within the Vector Autoregressive (VAR) framework testing the hypothesis P_t^2 causes P_t^1. This tests the price transmission within the selected markets (Rapsomanikis et al., 2006). Causality tests are more applicable in the analysis of domestic markets since causality from domestic to international market is a rare occurrence (Minot, 2011). The hypotheses to be tested are;

H_o .... There are no causal relationships in the domestic market.

H_1....... There exists at least one causal relationship

3.6.3 Vector Error Correction Model

The Vector Error Correction Model is based on the principle that a long run relationship exists between co-integrated variables. It aims at correcting short run equilibrium disturbances. Within the model, any disruptions that occur in one period are corrected in the next period. The long run model is combined with short run dynamics to give an estimate of the speed of adjustments. It is interpreted as a long run steady situation within which model dynamics fluctuate (Neusser, 2016). If two time series values of X and Y are integrated, then an error correction model exists. The VECM is applicable when certain conditions are met. The first condition is that the variables are co-integrated while the second is that the variables follow a random walk (Minot, 2011). It estimates the speed at which prices adjust to the long run equilibrium (Chalil, 2016). This speed is essential since it enables the researcher to understand the extent of market integration in the short run (Rapsomanikis et al., 2006). The general formula of VECM is;

$$\Delta y_{it} = \alpha_i + \sum_{j-1}^{n} \theta\, \Delta y_{jt-1} + Y_i \varepsilon_{it} + \epsilon_t \ldots\ldots\ldots\ldots \text{equation 12}$$

Where

Δy_{it} *is the first difference of price series*

y_{jt} *is the price series in the second market*

ε_{it} *is the residual series*

ϵ_t *is the error term of the residuals*

γ_i *is the speed of adjustment*

θ_i *is the short run dynamics*

The VECM assesses the dynamics of price relations and the speed of adjustment over time. (Rapsomanikis et al., 2006). The hypotheses to be tested by the model are;

H_o ……. *Error correction term is equal to zero*

H_1 ……. *Error correction term is not equal to zero*

3.7 Data Sources

This study utilises secondary data from the Ministry of Agriculture, Livestock and Fisheries in Kenya. Price is a key indication of the functioning of markets. It is for this reason that it is used in market integration studies. The data needed include;

Avocado price data- This study uses time series data for its analysis. The Ministry of Agriculture in Kenya computes early morning data prices of agricultural commodities daily, after which weekly and monthly averages are calculated. The data used in this analysis are monthly average prices in domestic markets in the period 2010 to 2019. This period is adequate to give insights into the functioning of regional avocado markets in Kenya. The prices are wholesale prices (price of a 90Kg bag). Wholesale prices are used due to limited availability of retail prices. Monthly data is used since it has fewer missing values compared to the weekly and daily data.

3.8 Data Processing

Data processing involves organizing and evaluating the data collected using analytical and logical reasoning. The data collected is usually in raw form, analysis therefore tries to obtain useful information from this. Descriptive statistics were used to quantitatively analyse the price series data sets for mean, minimum, maximum and trends. The tools used in data processing are Microsoft Excel spreadsheet which was used for sorting and reorganizing the relevant data, RStudio (version 3.5.1) was used for data preparing and cleaning and initial analysis while E-views (version 9) is used in performing the ADF test and the cointegration analysis. Data cleaning is an essential step to prepare for analysis and avoid wrong inferences. Graphical analysis is used to show the trend in prices. The RStudio script used is included in the annex section.

The prices obtained are in Kenyan shillings and remain as so in the analysis. In order to deal with currency fluctuations, all prices are deflated using consumer price index to cater for inflation. CPI compares the current price values to the previous year's price values at a particular point in time (FAO, 2010). The real prices obtained thereafter, are then used for analysis. The CPI values for the years 2010 to 2018 are obtained from World Bank (2020) and that of 2019 is obtained from Kenya National Bureau of Statistics (2020). The year 2010 is taken as the base year. The execution is done in an excel spreadsheet. The formula applied is;

$$\text{Real price in current year} = \text{CPI}_{\text{base year}}/\text{CPI}_{\text{current year}} * \text{Nominal price}_{\text{current year}}$$

3.9 Limitations in the analysis

There exist several techniques to assess market integration. This work is limited within the co-integration technique which requires extensive data availability. The technique also assumes that a long run equilibrium exists, which does not hold true in all situations. The existence of fixed adjustment costs may act as a hindrance (Abdulai, 2000). Barrett et al. (2002) argues that the co-integration technique is insufficient in the analysis of spatial markets since it assumes stationary transaction costs

Another limitation of this analysis is the fact that it is limited to a few domestic markets in Kenya and not all. The markets were specifically selected due to data availability. In the selected data set, there are a few missing values, out of 600 observations, 40 were missing. This represents 6.6% of the data. The missing values are as a result of non-reporting. For analysis purposes, the missing values are estimated using *imputeTS* package (Moritz et al., 2017) of RStudio and analysis proceeds thereafter using the complete dataset.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Introduction

This chapter describes the findings of this research. They are displayed and discussed in accordance with the conceptual framework developed in chapter three above.

4.2 Descriptive statistics

The number of observations made in each market is 120. Prices in the markets varied in the ten-year period (2010 to 2019). This variability is crucial in understanding the price relations between the markets. Descriptive statistics are performed on the data and the results are displayed in table 4 below. We observe that prices in the urban markets surpass the prices in the rural and producer market. Kisii which is the rural market; recorded the lowest prices in the entire period. Mombasa recorded the highest prices. The lowest recorded price is KES[4] 284, while the highest price is KES 3235. The high prices in Mombasa could be due to the fact it is the furthest from the producer market. As a result, it is affected by high transport costs involved in receiving produce especially perishable commodities. In addition, it is likely to experience adverse shortages due to any irregularities that may occur in the supply chain. All these factors influence trade in distant markets.

The degree of skewness in the dataset ranges from positive values to negative values. Nairobi, Eldoret and Kisii are skewed to the right while Nakuru and Mombasa are skewed to the left. The level of skewness in both directions is moderate. The Kurtosis values range from -0,86 to 1,73 lying within the acceptable range of -2 and +2 (Sudha , 2017). The level of Kurtosis depends on the outliers available in the data. As depicted in Annex 3, it was only in Nairobi

[4] 1 Euro=115. 988 KES (Kenyan Shilling). Exchange rate on 27/5/2020, from www.oanda.com

that the price series did not exhibit outliers. Our data therefore has minimal deviations from a normal distribution. We also observe that Mombasa has the highest standard deviation implying a higher price volatility in Mombasa. As discussed above, it is most distant market and its price levels are subject to other external influences. On the other hand, Nakuru has the lowest standard deviation implying low variability in price levels. This could be attributed to the fact that Nakuru does not entirely depend on other markets for supply of avocado as it also produces its own on a limited scale.

Table 4: Characteristics of the distribution of prices in Kenyan Shillings

	Nairobi	**Kisii**	**Eldoret**	**Nakuru**	**Mombasa**
Mean	1717	705	1370	1514	2134
SD	350	347	340	284	396
Min	1152	284	700	448	950
Max	2480	2033	2431	2240	3235
Median	1685	609	1295	1576	2072
Skew	0.31	1.48	0.94	-0.55	-0.04
Kurtosis	-1,06	1,96	0,56	0,85	1,77
Observations	120	120	120	120	120

4.3 Co-movement in Prices

Time series data can either display a deterministic character subject to trend, seasonality and cyclical components or be irregular implying a stationary process (Buteikis, 2018). To understand price relations in the five markets in the ten-year period, time series plots are developed using *ggplot2* package in RStudio. The result is illustrated in figure 8 below. An interaction graph of all price series is included in annex 4. We observe seasonal price variations in all the markets (see annex 5). Nairobi and Kisii experienced seasonality and a downward trend. In Eldoret and Nakuru, the prices exhibit both seasonality and cyclical components. They also exhibit a downward trend, but it is at a lower magnitude compared to that of Nairobi and Kisii. The prices in Mombasa do not display any specific deterministic feature, but rather high price variability over the years. A 2-year forecast graph of the market prices (in annex 6) indicates a 95% chance that the prices in Nairobi and Kisii will continue to decrease; but at a slower rate. On the other hand, the prices in Eldoret, Nakuru and Mombasa will incur higher price variations. When prices of an agricultural commodity are decreasing, farmers are worse affected. From the observation, the avocado farmers have continuously incurred low prices in the last decade. This is despite avocado being a high value crop.

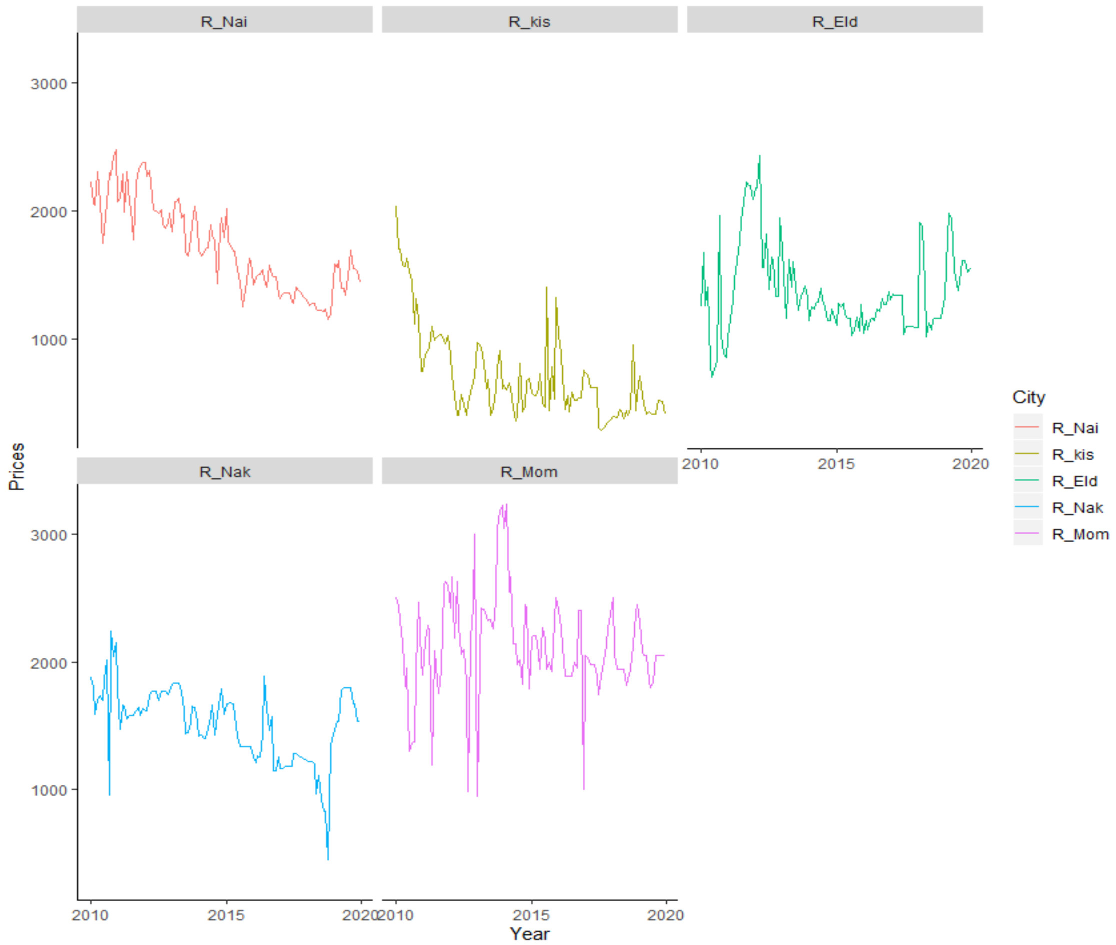


Figure 8: Graph showing price series[5] trends

[5] R_Nai, R_Kis, R_Eld, R_Nak and R_Mom refer to CPI adjusted prices for Nairobi, Kisii, Eldoret, Nakuru and Mombasa respectively.

To understand the linear relationship of the prices in the different markets, a bivariate correlation analysis is performed on the lagged values of the price series. The results are displayed in table 5 below. We observe significant positive correlations in the markets. The strongest correlation exists between price series in Nakuru and Nairobi at 0,7; followed closely by those of Nairobi and Kisii at 0,5. The weakest correlations are between price series in Eldoret and Nakuru; Eldoret and Mombasa each at 0,1. There is no significant correlation in price series of three pairs of markets that is Mombasa and Nakuru, Eldoret and Kisii, Kisii and Mombasa.

The price series of Mombasa has little or no association with the price series in the other markets, this is due to its high price volatility that is not experienced in the other markets. There is also positive correlation between price levels in Kisii and lagged prices of the other markets. The exception is between Kisii and the lagged price series of Eldoret and Mombasa. Since most correlation coefficients are positive, we can presume that the series have a common movement. The coefficients range from high to very low. A high coefficient would imply integration. However, correlation among price series could possibly be due to other factors and not necessarily market integration. For this reason, further analysis is undertaken to determine integration.

Table 5: Correlation Analysis

Covariance Analysis: Ordinary
Date: 05/28/20 Time: 13:30
Sample: 2011M01 2019M12
Included observations: 108
Balanced sample (listwise missing value deletion)

Correlation	KISII	LAG_ELDORET	LAG_KISII	LAG_MOMBASA	LAG_NAIROBI	LAG_NAKURU
KISII	1.000000					
LAG_ELDORET	-0.067337	1.000000				
LAG_KISII	0.436973	0.020782	1.000000			
LAG_MOMBASA	0.027323	0.104935	-0.047657	1.000000		
LAG_NAIROBI	0.495684	0.423504	0.495201	0.164461	1.000000	
LAG_NAKURU	0.407666	0.111359	0.354178	0.068943	0.750359	1.000000

4.4 Stationarity tests

The ADF test is performed on all observations at 1%, 5% and 10% level of significance. Schwarz Info Criterion is used in the selection of the number of lags; it yields 12 lags. This is appropriate for monthly time series data (Wooldridge, 2012). The ADF tests are done at level and first differencing of prices in the five markets. This analysis is performed in three different models of ADF; that is at the intercept, trend and intercept and with no trend or intercept.

The models of interest in this study are the trend with intercept and no trend or intercept models. These two are chosen since the plots of the series (in figure 8 above) reveal either a downward trend or no specific trend. When the series exhibits a clear trend and the model underrepresents it, results obtained are biased in favour of the null hypothesis. Likewise, if the series has no trend but the ADF regression contains a trend, the power of the test is reduced (Neusser, 2016). Annex 8 shows the results at level, we observe that for the trend and intercept model, the null hypothesis (price series is non-stationary) is rejected in Nairobi, Kisii, Nakuru and Mombasa. This means that these series are stationary while Eldoret is non-stationary. On the other hand, from the results of the no trend or intercept model, the null hypothesis is rejected in Kisii. This means that only price series for Kisii are stationary.

The results from the two models imply there is evidence of stationarity, but it is not conclusive for all the markets. Series with different order of integration imply that the markets are not integrated (Chalil, 2016). Annex 9 shows the results of the first differencing of the price series. Differencing eliminates trends in the series. Since all p-values resulting are zero, we reject the null hypothesis (price series are non-stationary). The results indicate that at first differencing the price series are stationary and integrated of order 1(1). The resulting graph is displayed in figure 10 below. Stationarity implies that the mean and variance remain static over time and the effects of seasonality are minimal. Markets must be integrated in the same order to allow cointegration analysis (Baiyegunhi et al., 2018). The ADF results at 1st differencing is satisfactory to proceed with the co-integration analysis. If the price series variables have a unit root, and this is not addressed; any further analysis result is spurious and not dependable (Hatemi-J, 2018).

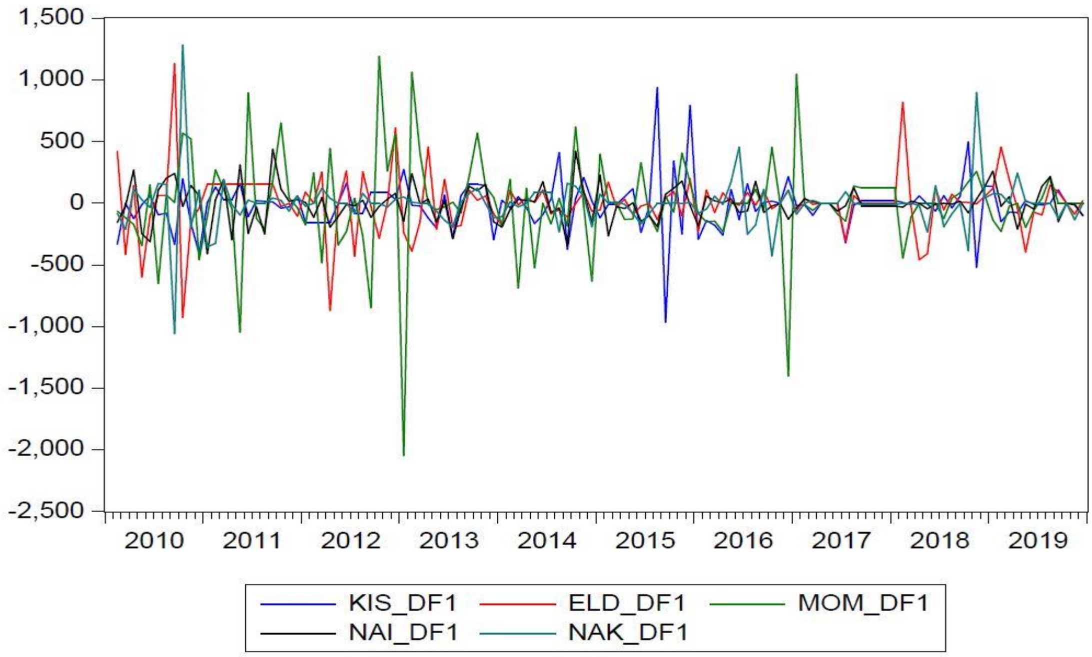


Figure 9: Graph of stationary price series[6] data

4.5 Engel and Granger Test

In the second part of the cointegration technique, the Engel and Granger test is performed. This test is a bivariate cointegration test and therefore analyses one relationship at a time. It involves two steps. The first being the estimation of the regression equation consisting of the independent and dependent variables; the second is testing stationarity of residuals from the model (Engel et al., 1987). The variables are cointegrated only when the residuals are stationary. This pairwise analysis is done on the differenced price series. Differencing is

[6] KIS_DF1, ELD_DF1, NAI_DF1, NAK_DF1 and MOM_DF1 refer to the 1st differencing of price series in Kisii, Eldoret, Nairobi, Nakuru and Mombasa respectively.

beneficial in that it removes any linear time trend from the series. This is a strict requirement for cointegration analysis (Wooldridge, 2012).

In the first step an OLS regression equation is estimated, Kisii is selected as the dependent variable while the independent variables are urban markets (Nairobi, Mombasa, Nakuru and Eldoret). In the first model price variables from all the markets are included, these results are displayed in annex 10. Subsequently, the other four modes are estimated; Kisii being the dependent variable, and each of the four urban markets representing the independent variable (see annexes 12, 14, 16 and 18). All the estimated models are valid since all R-squared values obtained are lower than the Durbin Watson statistic.

Thereafter, an ADF test is performed on the resulting residuals. The results from all the models are similar, they are represented in annexes 11, 13, 15,17 and 19. The t-statistic values from these ADF tests are compared against the cointegration critical values at 10%, 5% and 1% significance levels (see table 5 below). The critical values are obtained from MacKinnon, (2010). Given that the values of the t-statistic are greater than all the critical values, there is not enough evidence to reject the null hypothesis (there is no Cointegration of the price series) and conclusion is made that there is no interdependence in the local domestic markets, implying segmentation.

Table 6: Critical Values vs ADF test statistic for model with 2 variables

Significance level	**Observations**	**Critical values for model with no trend**	**ADF test statistic**
1%	<600	-3,90	-2,58
5%	<600	-3,30	-1,94
10%	<600	-3,04	-1,61

4.6 Granger Causality test

The causality test is an additional evidence to the results of the Engle and Granger test. The results obtained are displayed in annex 20. The observation is that no causal relationships exists, meaning that price changes in one market do not influence changes in any of the other markets. This observation validates the result of the Engle and Granger test that reveal no cointegration in the domestic avocado markets in Kenya. When markets are integrated; there exists at least one relationship where price series in one market influence the other (Rapsomanikis et al., 2006).

There are several reasons which could explain this situation. For example; high transport costs between the markets, imperfect substitution, inadequate information on prices and poor market infrastructure (Goundan et al., 2016). Since there are no causal relationships in the domestic markets it means there is asymmetric information flow. The farmers' production decisions are entirely independent and not based on market intelligence. This creates an overreliance on middlemen who in most cases would buy the produce at low costs and sell at high costs. While consumers complain of high food prices, the farmers complain of low prices. The benefits are mostly accrued by the middlemen.

4.7 Vector Error Correction Model

When two variables are not cointegrated, it implies there is no long run relationship and any regression model formed thereafter is spurious and the results meaningless (Wooldridge, 2012). However, this model can be useful in explaining the differences between the variables but not their relationship. Since from the cointegration analysis, we found no integration existing in the local markets, conducting the VECM is not necessary. The absence of integration means that the markets are not linked by competitive arbitrage.

4.8 Policy Implication

This analysis was done on a step by step basis. The general observation is that there is price co-movement in price series. Correlation also existed; tentatively informing that market integration exists. However, on further analysis using the cointegration and granger causality tests, these results were not confirmed. It is noted that although co-movement and correlation exists, there is no long run equilibrium to which prices respond. In designing policies this is a crucial consideration since existence of correlation in the markets means that policy changes in one market would have a similar effect on all the other markets (FAO, 2010).

The absence of causal relations means that demand and supply in each of the markets are unique. Policy should therefore be market specific addressing the needs of each market. However, on the downside this could mean extra costs in implementation. Given that each of the markets are in different counties and the government is devolved, then each county could work on its own implementation plans all aimed at achieving specific goals. The goals could include improved market access by both producers and consumers. An important consideration in designing the policies would be specifying a goal which can be achieved through the implementation of several individual projects.

One policy option is the improvement of existing markets functions. Since market efficiency can be improved through market integration; policy should address the issues that improve market integration. They include enhanced access to market information, the establishment of market infrastructure including physical buildings, road networks and storage facilities. Improved access to market information will go a long way in helping the farmers in sourcing for the buyers once the produce is ready. It will also mean that the buyers can communicate their needs to the producer. The marketing costs often hinder main market access, these could be reduced as well.

Geographic distance may influence trade and arbitrage (Goundan et al., 2016). This means that food miles are a key policy consideration. Once an agricultural commodity is produced, it needs to be distributed to the consumers. In some instances, the consumer is far from the production zones; therefore, a lot of time and distances are covered to get food to the end consumer. This movement of the produce from one place to another, involves costs which increases the food prices to the end consumer. Given the perishable and bulky nature of agricultural commodities, it means special attention is needed in transportation. These could include handling and packaging requirements, which further increase the costs. Policy should aim at making the food distribution system efficient. This will ensure availability and accessibility of agricultural commodities. It will also significantly reduce postharvest losses.

Locally traded commodities experience more price volatility compared to internationally traded goods (Goundan et al., 2016). Given that the larger avocado market in Kenya is the domestic market, then volatility is unavoidable. Price stabilization policies could play a crucial role in cushioning the producers against adverse price changes. Farmers could need support in distribution mechanisms to access distant markets. Important considerations in designing food distribution programs are food safety and perishability.

Since production volumes are known, it would be crucial to understand consumer behaviour in designing the market development strategies. Demand is influenced by household income, at the same time, the supply from a household will influence its income level. Considerations in examining consumer behaviour include; dietary needs, food preferences, urbanization, income levels, market prices, among others. Some factors that influence the supply include; production levels, postharvest losses, demand for substitute commodities among others. Changing consumer patterns and urbanization also influence food systems and rural economies. All these must be considerations for policy.

Policy should also drive production by supporting farmers in acquisition of high-quality tree seedlings and education on crop management practices. In addition to economic incentives for promoting production, a well-organised support system is crucial. Policy should aim at strengthening social capital to enhance the success in the implementation of projects. This could be through partnerships, strengthening civil societies and farmer education through farmer field schools. In order to ensure successful partnerships, both private and public stakeholders should be involved. Strengthening social capital also involves working with the farmers on follow up activities to oversee policy implementation. A strong social system could be useful in the strengthening or establishment of farmer cooperatives. These would benefit from economies of scale through; increased bargaining power in fetching higher prices for their produce and reduced informational and other transactional costs.

CHAPTER FIVE

CONCLUSION

This research is a spatial market integration analysis done in five avocado markets in Kenya. The technique applied is the Engle and Granger cointegration. The results obtained at the initial stages of the analysis implied existence of integration, but further analysis revealed that this was void. One evident conclusion is that correlation alone cannot detect integration and further analysis is always needed to validate or disapprove the results of correlation analysis. We can conclude that correlation and cointegration are different and the fact that market prices are correlated does not mean that the markets are integrated. A general observation from the results of this study is that a successful policy should be based on evidence from research.

Another conclusion that we can draw from the analysis is that tradability between the domestic markets is not fully exploited and market integration would serve a long part in improving it.

The domestic markets are not cointegrated and the producers from Kisii don't easily access the other markets. The persistence of this situation means that the farmers do not have adequate market access and will continue to incur low prices from the sale of avocado produce. The prices will remain low provided that their access to other markets is limited. One of the crucial goals in improving income among farmers is by improving market access which can be attained by integration.

Seasonal variations are experienced in the markets with huge price variations experienced in the months of July to December. This is in line with the annual production cycle. The seasonal patterns vary from market to market. The farthest market (Mombasa) experienced high seasonal price variability. We also conclude that prices in the short run move in the same direction, but a long run relationship is not established. In addition, the lack of a long run equilibrium in the avocado markets implies future instability in the price levels. However, this study did not establish the speed of adjustment to the long run equilibrium since no cointegration relationships were present between all market pairs. The granger causality test was used to test for price transmission. These results showed that none of the selected markets had influence over prices in the other markets.

REFERENCES

Abdulai A. (2000). Spatial price transmission and asymmetry in the Ghanaian maize market. *Journal of Development Economics*, *63*(2), 327–349. https://doi.org/10.1016/S0304-3878(00)00115-2

Abdulai A. (2006). Spatial Integration and Price Transmission in agricultural commodity markets in sub-Saharan Africa. In Agricultural Commodity Markets and Trade (pp 163-186). Food and Agriculture Organisation. Rome, Italy

Acquah H.D. & Owuso R. (2012). Spatial Market Integration and Price Transmission of Selected Plantain Markets in Ghana. *Journal of Sustainable Development in Africa*, *14*(5), 208–217. Retrieved from http://jsd-africa.com/Jsda/Vol14N5-Fall2012A/PDF/Spatial Market Intergration.HenryAcquah.pdf

Ahmed M. & Naresh S. (2017). Market Integration and Price Transmission in Major Onion Markets of India. *Journal of Economic Affairs*, *62*(3). https://doi.org/10.5958/0976-4666.2017.00051.1

Amare M., Mariara J., Oostendorp R., & Pradhan M. (2019). The impact of smallholder farmers' participation in avocado export markets on the labor market, farm yields, sales prices, and incomes in Kenya. *Land Use Policy*, *88*(June 2019), 104168. https://doi.org/10.1016/j.landusepol.2019.104168

Amikuzuno J. (2010). Spatial Price Transmission Analysis in Agricultural Markets : Does the Data Frequency Improve our Estimation ? In *Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19-23, 2010.*

Ashraf N., Gine X., & Karlan D. (2008). Finding Missing Markets (and a disturbing epilogue): Evidence from an Export Crop Adoption and Marketing Intervention in Kenya. Working paper No. 08-066.

Baiyegunhi L., Sharaunga S., Dlangisa S., & Ndaba N. (2018). Tomato Market Integration: a Case Study of the Durban and Johannesburg Fresh Produce Markets in South Africa. *Journal of Agribusiness and Rural Development*, *49*(3), 239–249. https://doi.org/10.17306/j.jard.2018.00419

Baltzer K. (2015). International to Domestic Price Transmission in Fourteen Developing Countries during the 2007–8 Food Crisis. In *Food Price Policy in an Era of Market Instability* (pp. 21–50). https://doi.org/10.1093/acprof:oso/9780198718574.003.0002

Barrett C. B. (2001). Measuring Integration and Efficiency in International Agricultural Markets. *Applied Economic Perspectives and Policy*, *23*(1), 19-32. https://doi.org/10.1111/1058-7195.00043

Barrett C. B., & Li J. R. (2002). Distinguishing between equilibrium and integration in spatial price analysis. *American Journal of Agricultural Economics*, *84*(2), 292–307. https://doi.org/10.1111/1467-8276.00298.

Barzel Y. (2007). *Price Theory: Replacing the Law of One Price with the Price Convergence Law*. Department of Economics, University of Washington.

Business Daily. (2018, January 21). Avocado Exports banned as price rises. Retrieved January 12, 2020, from https://www.businessdailyafrica.com/news/Avocado-exports-banned-as-price-rises/539546-4272600-9gthm2z/index.html

Business Daily. (2020, February 17). Avocado exports resume after ban. Retrieved April 5, 2020, from https://www.businessdailyafrica.com/markets/marketnews/Avocado-exports-resume-after-ban/3815534-5457852-to68q2z/index.html

Buteikis Andrius. (2018). Time series with trend and seasonality components. Retrieved May 4, 2020, from http://web.vu.lt/mif/a.buteikis/

CBI. (2020a). Entering the European market for avocados. Retrieved April 16, 2020, from https://www.cbi.eu/node/3070/pdf/

CBI. (2020b). The European market potential for avocados. Retrieved April 16, 2020, from https://www.cbi.eu/node/3069/pdf/

Chalil D. (2016). Sustainability of Sustainable Palm Oil: A Market Integration Analysis. *Jurnal Manajemen Dan Agribisnis*, *13*(2), 157–167. https://doi.org/10.17358/jma.13.2.157

Chalmers N., Revoredo-giha C., & Jumbe C. (2019). Measuring the Degree of Integration in the Dairy Products Market in Malawi. *Social Science*, *8*(66). https://doi.org/10.3390/socsci8020066

Chaudhary P., Khamar J. & Sen D. (2015). Avocado: the holistic source as a natural doctor! *World Journal of Pharmaceutical Research*, *4*(8), 748–761. Shri Sarvajanik Pharmacy College, Gujarat Technological University, Arvind Baug, Mehsana-384001, Gujarat, India.

Elimu.org. (2020). Major Towns in Kenya. Retrieved June 5, 2020, from http://learn.e-limu.org/topic/view/?c=48&t=287

Embassy of the Kingdom of Netherlands. (2017). Horticulture study : Mapping of production of fruits and vegetables in Kenya, (March).

Engel R., & Yoo S.B. (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics*, *35*, 143–159.

Fackler P.L. & Goodwin B. K. (2001). Spatial price analysis. In *Handbook of agricultural economics.* (pp. 971–1024). Elsevier.

Faminow M.D., & Benson L. B. (1990). Integration of Spatial Markets. *American Journal of Agricultural Economics*, *72*(1), 49–62.

FAO. (2004a). *Avocado Post-harvest Operations*. Rome. Retrieved February 15, 2020, from http://www.fao.org/fileadmin/user_upload/inpho/docs/Post_Harvest_Compendium_-_Avocado.pdf

FAO. (2004b). *Price Transmission in Selected Agricultural Markets*. Commodity and Trade Policy Research Working Paper No. 7. Rome, Italy

FAO. (2010). *Markets & Price Monitoring Training Manual*. *Food Security Technical Secretariat of the Ministry of Agriculture*. https://doi.org/10.1111/j.1365-2753.2006.00669.x

FAOSTAT. (2018). Kenyan Avocado Production and Export Volumes in 2017. Retrieved May 8, 2020, from http://www.fao.org/faostat/en/#data/

FAOSTAT. (2020). Kenya Avocado Statistics: Production and Acerage. Retrieved April 17, 2020, from http://www.fao.org/faostat/en/#data/QC

Federico G. (2007). Market integration and market efficiency: The case of 19th century Italy. *Explorations in Economic History*, *44*(2), 293–316. https://doi.org/10.1016/j.eeh.2006.02.003

Ghosh M. (2011). Agricultural Policy Reforms and Spatial Integration of Food Grain Markets in India. *Journal of Economic Development*, *36*(2), 15–37.

Gitau R., & Meyer F. (2019). Spatial price transmission under different policy regimes: A case of maize markets in Kenya. *African Journal of Agricultural and Resource Economics, 14*(1), 15–27.

Goletti F., & Raisuddin A. (1995). Structural determinants of market integration: the case of rice markets in bangladesh. *The Developing Economies, 2*(June).

González-Rivera G., & Helfand, S. M. (2001). The extent, pattern, and degree of market integration: A multivariate approach for the Brazilian rice market. *American Journal of Agricultural Economics, 83*(3), 576–592. https://doi.org/10.1111/0002-9092.00179

Goundan A., & Tankari M. (2016). *A Dynamic Spatial Model of Agricultural Price Transmission: Evidence from the Niger Millet Market*. Dakar, Senegal.International Food Policy Research Institute Discusiion paper 01536.

Government of Kenya. (2019). Agricultural Sector Transformation and Growth Strategy, 1–221. Report Retrieved on 15 january 2020, from http://www.kilimo.go.ke/wp-content/uploads/2019/01/ASTGS-Full-Version.pdf

Griesbach J. (2005). *Avocado growing in Kenya*. *World Agroforestry Centre (ICRAF)*. Nairobi. Retrieved on 18 January 2020, from http://outputs.worldagroforestry.org/cgi-bin/koha/opac-detail.pl?biblionumber=35521

Gyau A., Mbugua M., & Oduol J. (2016). Determinants of participation and intensity of participation in collective action: evidence from smallholder avocado farmers in Kenya. *Journal on Chain and Network Science, 16*(2), 147–156. https://doi.org/10.3920/JCNS2015.0011

Haji, B. J. (2014). *Module on Agricultural Marketing and Price Analysis*. School of Agricultural Economics and Agribusiness, Haramaya Univeristy, Ethiopia. Retrieved from https://www.coursehero.com/file/33094767/Module-on-Agricultural-Marketing-and-Price-Analysis-Finalpdf/

Hakizimana C., & May J. (2018). Agriculture and poverty reduction: a critical assessment of the economic impact of the avocado industry on smallholder farmers in Giheta – Burundi. *Forests, Trees and Livelihoods*, *27*(4), 203–216. https://doi.org/10.1080/14728028.2018.1515041

Hatemi-J A. (2018). Hidden panel cointegration. *Journal of King Saud University - Science*, *32*(1), 507–510. https://doi.org/10.1016/j.jksus.2018.07.011

Hillen, J. (2019). AES Prize Essay Market Integration and Market Efficiency under Seasonal Tariff Rate Quotas. *Journal of Agricultural Economics*, *70*(3), 859–873. https://doi.org/10.1111/1477-9552.12355

Horticultural Crop Directorate. (2014). *National Horticulture Market Information System Baseline Survey Report*. National Horticulture Market Information Secretariat. Nairobi, Kenya.

Horticultural Crops Directorate. (2017). *Horticulture Validated Report 2015-2016. Agriculture, Fisheries and Food Authority (AFFA), horticulture performance report*. Nairobi, Kenya. Retrieved from https://www.agricultureauthority.go.ke/wp-content/uploads/2016/09/Horticulture-2015_2016-Validated-Report3.pdf

Horticultural Crops Directorate. (2019). *Horticulture Validated Report 2017-2018. Agriculture, Fisheries and Food Authority (AFFA), horticulture performance report*. Nairobi, Kenya. Retrieved from http://www.agricultureauthority.go.ke/wp-content/uploads/2016/05/Horticulture-Validated-Report-2014-Final-copy.pdf

Jena P. K. (2016). Commodity market integration and price transmission: Empirical evidence from India. *Theoretical and Applied Economics*, *23*(3), 283–306.

Jones C., Gordon M., & Webber M. (2010). Achieving Synergies through Clustering Kenyan Avocados. In *Building Competitiveness in Africa's Agriculture (pp 73-76)*. Webber M. and Labaste P. (Ed.), Washington D.C: The World Bank.

Kabbiri R. (2018). *Dairy Market Integration: Impact of Mobile Phones on Price Transmission and Marketing Channels in Uganda.* Doctoral Thesis. Ghent University.

Kabbiri R., Dora M., Elepu G. & Gellynck X. (2016). A Global Perspective of Food Market Integration: A Review. *Agrekon*, *55*(1–2), 62–80. https://doi.org/10.1080/03031853.2016.1159589

Ke, S. (2015). Domestic market integration and regional economic growth-china's recent experience from 1995-2011. *World Development*, *66*, 588–597. https://doi.org/10.1016/j.worlddev.2014.09.024

Kenya National Bureau of Statistics. (2018). *Basic report on well-being in Kenya : based on the 2015/16 Kenya Integrated Household Budget Survey (KIHBS)*. Retrieved from https://www.knbs.or.ke/download/basic-report-well-kenya-based-201516-kenya-integrated-household-budget-survey-kihbs/#. Nairobi, kenya

Kenya National Bureau of Statistics. (2019). *2019 Kenya Population and Housing Census Volume I : Population By County and Sub-County* (Vol. I). Retrieved from http://www.knbs.or.ke. Nairobi, Kenya

Kenya National Bureau of Statistics. (2020). Consumer Price Index. Retrieved April 3, 2020, from https://www.knbs.or.ke/consumer-price-indice/

Machoka P. & Chege G. (2011). *Ideal Matunda Ltd : Investment Challenges*. *United States International University. Nairobi, Kenya*

MacKinnon, J. (2010). *Critical Values for Cointegration Tests* (Economics Department, Queen's University). *Working Paper* No. 1227.

Minot N. (2011). Transmission of World Food Price Changes to Markets in Sub-Saharan Africa. *International Food Policy Research Institute*, (January). Retrieved from http://www.ifpri.org/publication/transmission-world-food-price-changes-markets-sub-saharan-africa

Moritz S. & Bartz-Beielstein T. (2017). ImputeTS: Time series missing value imputation in R. *R Journal*, *9*(1), 207–218. https://doi.org/10.32614/rj-2017-009

Morlacco M. (2017) Market power in Input Markets: Theory and Evidence form French Manufacturing. Yale University. Retrived from https://www.tsefr.eu/sites/default/files/TSE/documents/sem2018/Jobmarket2018/JMP/morlacco_jmp.pdf

Mose L. (2007). *Who Gains , Who Loses ? The Impact of Market Liberalisation on Rural Households in Northwestern Kenya*. Doctoral Thesis. Wageningen University.

Mtumbuka W. S., Mapemba, L., Maonga B., & Mwabumba M. (2014). Spatial price integration among selected bean markets in Malawi: A threshold autoregressive model approach. *Food Policy*, *35*(4), 16. https://doi.org/10.1016/j.foodpol.2010.03.002

Mwambi M., Oduol J., Mshenga P., & Saidi M. (2016). Does contract farming improve smallholder income? The case of avocado farmers in Kenya. *Journal of Agribusiness in Developing and Emerging Economies, 6*(1), 2–20. https://doi.org/10.1108/JADEE-05-2013-0019

Negassa A., Myers R., & Gabre-Madhin E. Z. (2003). Analyzing Grain Market Efficiency in Developing Countries: Review of Existing Methods and Extensions to the Parity Bounds Model. *MTID Discussion Paper*, *63*.

Neusser K. (2016). *Time series econometrics*. *Introductory Econometrics: A Practical Approach*. Bern: Springer International Publishing AG Swtizerland. https://doi.org/10.1017/cbo9780511814839.012

Oktarina Y. (2015). Rice Market Integration in Southern Sumatra, Indonesia. *International Journal of Agricultural Research, Innovation and Technology*, *5*(2), 1–6. http://www.ijarit.webs.com

Oluoch-kosura W. (2017). *Performance of Kenya's Agriculture : A Macro-Economic Perspective*. University of Nairobi, Kenya. Retrieved from https://ec.europa.eu/jrc/sites/jrcsh/files/3_JRC_KenyaWS_Kosura.pdf

Omiti J., Otieno D., Nyanamba T., & McCullough E. (2009). Factors influencing the intensity of market participation by smallholder farmers: A case study of rural and peri-urban areas of Kenya,. *African Journal of Agricultural and Resource Economics, African Association of Agricultural Economists, 3*(1), 1–26. https://doi/DOI:%2010.22004/ag.econ.56958

Omolo P., Tana P., Okwach E., Onyango H., & Okach K. O. (2011). Analysis of avocado marketing in trans-Nzoia district, Kenya. *Journal of Development and Agricultural Economics*, *3*(7), 312–317.

Ongeri B. (2014). Small Scale Horticultural farming along the Kenyan Highways and Local economic development: Exploring the effect of factor prices. *International Review of Research in Emerging Markets and the Global Economy (IRREM) An Online International Research Journal*, *1*(3), 2311–3200. Retrieved from www.globalbizresearch.org

Pan F. (2019). Evolution of Agricultural Spatial Market Integration : Evidence from the Hog Market in China. *Journal of Agricultural and Applied Economics*, *51*, 349–367. https://doi.org/10.1017/aae.2019.7

Penzhorn N., & Arndt C. (2002). Maize markets in Mozambique: Testing for market integration. *Agrekon*, *41*(2), 146–159. https://doi.org/10.1080/03031853.2002.9523590

Rahman K. & Palash M. (2018). Fed Cattle Market Integration in Mandatory Price Reporting Era. *Open Agriculture, 3*, 348–355. https://doi.org/10.1515/opag-2018-0038

Rapsomanikis G., Hallam D., Conforti P. (2006). Market integration and price transmission in selected food and cash crops of developing countries: Review and Applications. In H. D. Sarris Alexander (Ed.), *Agricultural Commodity Markets and Trade* (pp. 187–217). https://doi.org/10.1016/j.foodpol.2013.11.001

Rashid S. (2007). The "Law" of One Price: Implausible, Yet Consequential. *The Quarterly Journal of Austrian Economics*, *10*(1), 77–88. https://doi.org/10.1007/s12113-007-9001-7

Rashid S. (2004). *Spatial Integration of Maize Markets in Post-liberalized Uganda* (MTID Discussion paper No. 71). International Food Policy Research Institute. Washington D.C.

Rashid S., Minot N., Lemma S. & Behute B., (2010). Are Staple Food Markets in Africa Efficient? Spatial Price Analyses and Beyond. In *A paper presented at the COMESA Policy Seminar "Food Price Variability: Causes, Consequences, and Policy Options" on 25-26 January 2010 in Maputo, Mozambique under the COMESA –MSU-IFPRI African Agricultural Markets Project AAMP*. Retrieved from https://ageconsearch.umn.edu/bitstream/58562/2/AAMP_Maputo_12_spatial_price_variation.pdf

Ravallion Martin. (1986). Testing Market Integration. *American Journal of Agricultural Economics*, *68*(1), 102–109.

Castillo V., Juan S., & García-Cortijo, Ma Carmen. (2015). Price adjustment in world wine markets: A cointegration analysis. *Wine Economics and Policy*, *4*(2), 128–135. https://doi.org/10.1016/j.wep.2015.05.004

Söderbom M., R. B. (2009). *Market Integration and Structural Transformation in a Poor Rural Economy* (No. 4856). *Policy Research Working Paper No 4856.* Poverty Reduction Group, The World Bank

Statista. (2020a). Avocado Global Production. Retrieved April 18, 2020, from https://www.statista.com/statistics/593211/global-avocado-production-by-country/

Statista. (2020b). Export Value of Avocados Worldwide in 2017. Retrieved April 18, 2020, from https://www.statista.com/statistics/938561/major-exporters-avocado-export-value/

Statista. (2020c). Volume of fresh avocado imports into the United States from 2012/13 to 2018/19. Retrieved April 18, 2020, from https://www.statista.com/statistics/592522/volume-of-fresh-avocado-imports-us/

Steglich M., Keskin, E., Hall A., & Dijkman J. (2009). *Are International Market Demands Compatible with Serving Innovation Capacity in Kenya's Horticulture Industry* (Working paper No. 2009-009). United Nations University, Maastricht Netherlands.

Sudha R. (2017). Measures of Skewness and Kurtosis. *Research and Biostatistics for Nurses, 200*, 412–412. https://doi.org/10.5005/jp/books/13016_16

Sunga Chalwe. (2017). *An analysis of spatial market integration: A case of Zambian dry bean markets connected by informal trade to Tanzania and the Democratic Republic of Congo*. MSc. Thesis, University of Pretoria. Retrieved from https://scholar.sun.ac.za

Thomson Reuters Foundation. (2018). Kenya's ground-down coffee farmers switch to avocado amid global boom. Retrieved January 19, 2020, from https://www.reuters.com/article/us-kenya-agriculture-climatechange/kenyas-ground-down-coffee-farmers-switch-to-avocado-amid-global-boom-idUSKBN1JW1AV

Tsiboe F., Dixon B., Wailes. E. (2016). Spatial dynamics and determinants of Liberian rice market integration. *African Journal of Agricultural and Resource Economics*, *11*(3), 183–196.

Varela G., Calì M., Pape U. & Rojas E. (2016). *Market Integration and Poverty: Evidence from South Sudan . Policy Research Working Paper* No. 7564. Trade and Competitiveness Global Practise Group. The World Bank. https://doi.org/doi:10.1596/1813-9450-7564

Vargova L., & Rajcaniova, M. (2018). Spatial Price Transmission of Milk Prices Among the Visegrad Countries. *Visegrad Journal on Bioeconomy and Sustainable Development*, *6*(2), 79–83. https://doi.org/10.1515/vjbsd-2017-0014

Waluube P. (2009). *An Analysis of the Rice Market Integration in Uganda*. MSc. Thesis Makerere Universuty Uganda.

Wasilwa L., Okoko E., & Njuguna J. (2017). *Status of Avocado Production in Kenya*. Kenya Agricultural and Livestock Research Institute. Nairobi.

Wooldridge J. (2012). *Introductory Econometrics: A modern Approach 5th Edition*. *Applied Discrete-Choice Modelling*. Michigan State University.

https://doi.org/10.4324/9781351140768-8

Worako T. K. (2015). Transactions costs and spatial integration of vegetable and fruit market in Ethiopia. *Ethiopian Journal of Economics*, *24*(1), 89–130.

World Bank. (2020). World Development Indicators: Consumer Price Index. Retrieved April 3, 2020, from https://data.worldbank.org/indicator/FP.CPI.TOTL?locations=KE

World Food Programme. (2007). *PDPE Market Analysis Tool : Market Integration*. *World Food Programme*. Rome.

Yoon J. (2017). *Three essays on Spatial, Vertical price transmission and Asymmetric Supply Response in Livestock Sectors*. Doctoral Thesis. *University of Missouri-Columbia*. https://doi.org/10.1016/j.sbspro.2015.04.758

Zakari S., Ying L., Song B. (2014). Market integration and spatial price transmission in Niger grain markets. *African Development Review*, *26*(2), 264–273. https://doi.org/10.1080/00036846.2020.1726862

Zewdie H. (2017). Spatial market integration and price transmission for papaya markets in Ethiopia. *Journal of Development and Agricultural Economics*, *9*(5), 129–136. https://doi.org/10.5897/jdae2016.0777

ANNEXES

Annex 1: Data request letter

Faculty of Bioscience Engineering

Department of Agricultural Economics - Head of Department Prof. dr. ir. Marijke D'Haese

Your reference	Our reference	Date
		13/02/2020

To whom it may concern

This letter confirms that Baroness Chelangat Bor is a bona fide student of the International Master of Science in Rural Development at Ghent University. As a partial requirement for the fulfilment of this degree she is undertaking a research thesis titled "Spatial Market integration in improving Rural Livelihoods: A case of Avocado farming in Kenya".

She would therefore like to have access to following data:

- Time Series monthly avocado prices for the period 2010-2019 in the towns of interest of this study. These towns are Nairobi, Kisii, Eldoret and Kisumu.

Your assistance will be appreciated.

GHENT UNIVERSITY
Faculty of Bio-science Engineering
Department of Agricultural Economics
Coupure Links 653, 9000 Gent, Belgium
TEL. 09 264 59 27 – FAX 09 264 62 46

Yours sincerely

Prof. Dr.ir. Stijn Speelman
Department of Agricultural Economics
Ghent University, Belgium

UNIVERSITEIT GENT

Faculty of Bioscience Engineering - Coupure links 653 - B-9000 Ghent - Belgium
tel. +32 9 264 93 75 - fax +32 9 264 62 46 – Stijn.speelman@UGent.be - http://www.fbw.ugent.be

Annex 2: RStudio Scripts

```
#read the avocado dataset
library(readxl)  #to read the data
newavo <- read_excel("F:/Project/Data Analysis/newavo.xlsx")
View(newavo)
dim(newavo)
summary(newavo)


#read colum 1 as dates
dates<-seq(as.Date("2010-01-01"), length=120, by ="months")  #read column 1 as dates
library(xts)  #to create time series
avo_ts1<-xts(newavo[ ,-1],order.by = dates)  #creating the ts
View(avo_ts1)
class(avo_ts1)
names(avo_ts1) #names of Columns


#subset the data to select relevant columns
#select only the real prices columns
r.prices<-avo_ts1[ ,c(7,8,9,10,11)]
head(r.prices)
View(r.prices)


#cleaning data to estimate missing values
#leave outliers since they could account for seasonality
library(imputeTS)  #to estimate missing values
clean<-na_interpolation(r.prices,option="linear")   #linearimputation selected
View(clean)


#descriptive stats
library(psych)  #to produce neat summary
describe(clean) #show descriprive statistics
```

```
#create a new dataframe with three columns
#grouping the data
#to enable plotting of the series
temp<-data.frame(index(clean),stack(as.data.frame(coredata(clean))))
View(temp)
#change names of the columns
names(temp)[1]<-"Year"
names(temp)[2]<-"Prices"
names(temp)[3]<-"City"
head(temp)
dim(temp)
#check the structure of the ts
str(temp)
class(temp)


#Creating the plots
library(ggplot2)  #to create the graphs
library(scales)   #to scale x and y axis
ggplot(temp,aes(x=Year,y=Prices,color=City))+geom_line()
p1<-ggplot(temp,aes(x=Year,y=Prices,color=City))+geom_line()
p1<-p1+ggtitle("Price Series")+scale_x_date(labels = date_format("%Y"), breaks='2 years')
p1<-p1+scale_y_continuous(breaks = c(500,1000,1500,2000,2500,3000,3500), limits = c(200,3500))
p1
p1<-p1+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
             panel.background = element_blank(), axis.line = element_line(colour = "black"))
p1
#Ploting the boxplots of each city separately
p2<-ggplot(data = temp, mapping = aes(x = City, y = Prices)) +geom_boxplot()
p2<-p2+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
             panel.background = element_blank(), axis.line = element_line(colour = "black"))

p2
#plotting the series side by side
p3<-ggplot(data = temp, mapping = aes(x = Year, y = Prices,color=City)) +
  geom_line() +
  facet_wrap(facets = vars(City))
p3<-p3+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
             panel.background = element_blank(), axis.line = element_line(colour = "black"))
p3
```

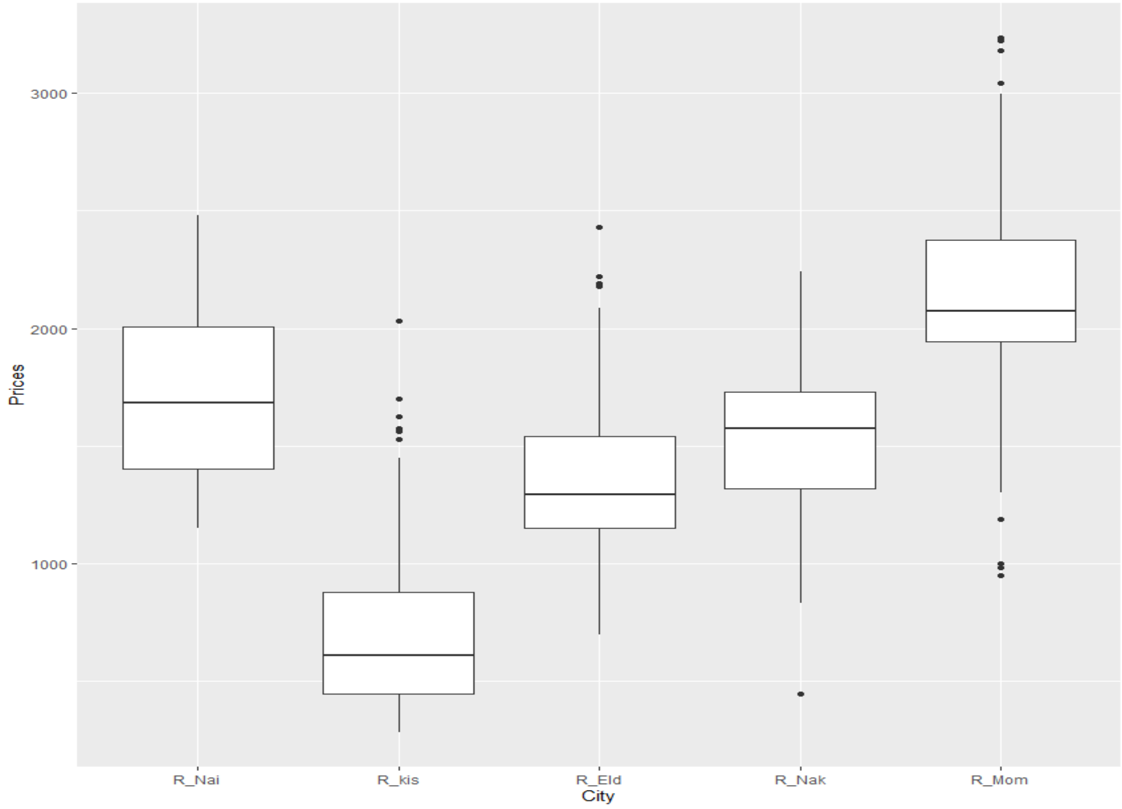
```
#second way to plot
p4<-ggplot(data = temp, mapping = aes(x = Year, y = Prices, color = City)) +
  geom_line() +
  facet_grid(rows = vars(City), cols =  vars(NULL))
p4<-p4+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
             panel.background = element_blank(), axis.line = element_line(colour = "black"))
p4
#conclude p3 is better than p1 and p4


#creating new time series variables for prediction
#avo.nai=prices in Nairobi, avo.kis=prices in Kisii, avo.eld=prices in eldoret
#avo.nak=price in nakuru, avo.mom=prices in mombasa
dim(clean)
head(clean)
avo.nai<-ts(clean[ ,1], start = c(2010,1), frequency=12)
avo.kis<-ts(clean[ ,2], start = c(2010,1), frequency=12)
avo.eld<-ts(clean[ ,3], start = c(2010,1), frequency=12)
avo.nak<-ts(clean[ ,4], start = c(2010,1), frequency=12)
avo.mom<-ts(clean[ ,5], start = c(2010,1), frequency=12)
autoplot(avo.nai)
autoplot(avo.kis)
autoplot(avo.eld)
autoplot(avo.nak)
autoplot(avo.mom)


#transform the variable to make them stationary
#take the first difference of each series to get rid of the trends in series
#difference selected based on ADF results from eviews
dy.nai<-diff(avo.nai,lag=1,differences = 1)
dy.kis<-diff(avo.kis,lag=1,differences = 1)
dy.eld<-diff(avo.eld,lag=1,differences = 1)
dy.nak<-diff(avo.nak,lag=1,differences = 1)
dy.mom<-diff(avo.mom,lag=1,differences = 1)
autoplot(dy.nai)
autoplot(dy.kis)
autoplot(dy.eld)
autoplot(dy.nak)
autoplot(dy.mom)
```

```
#check for seasonality
#plotting combined seasonality plots for each market
library(ggplot2)
library(cowplot) #to combine the plots
nai.s<-ggsubseriesplot(dy.nai)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                                     panel.background = element_blank(), axis.line = element_line(colour = "black"))
Kis.s<-ggsubseriesplot(dy.kis)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                                     panel.background = element_blank(), axis.line = element_line(colour = "black"))
eld.s<-ggsubseriesplot(dy.eld)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                                     panel.background = element_blank(), axis.line = element_line(colour = "black"))
nak.s<-ggsubseriesplot(dy.nak)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                                     panel.background = element_blank(), axis.line = element_line(colour = "black"))
mom.s<-ggsubseriesplot(dy.mom)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                                     panel.background = element_blank(), axis.line = element_line(colour = "black"))
plot_grid(nai.s,Kis.s,eld.s,nak.s,mom.s,
          labels = c('Nairobi','Kisii','Eldoret','Nakuru', 'Mombasa'),
          ncol = 3)
#Nai, eld and mom have seasonality
#kis and nak experiences seasonality only in the months of Jul to Dec
#conclude seasonality exists in the markets but not similar patterns
#variables have both seasonality and trend
```

```r
#creating model for prediction
library(forecast) #to predict
#Arima model
#model for nairobi
arima.nai<-auto.arima(avo.nai,d=1,D=1, stepwise = FALSE,approximation = FALSE, trace=TRUE)
checkresiduals(arima.nai)
print(summary(arima.nai))  #residual sd=145.72
#model for Kisii
arima.kis<-auto.arima(avo.kis,d=1,D=1, stepwise = FALSE,approximation = FALSE, trace=TRUE)
checkresiduals(arima.kis)
print(summary(arima.kis))  #residual sd=220.55
#model for Eldoret
arima.eld<-auto.arima(avo.eld,d=1,D=1, stepwise = FALSE,approximation = FALSE, trace=TRUE)
checkresiduals(arima.eld)
print(summary(arima.eld))  #residual sd=247.45
#model for Nakuru
arima.nak<-auto.arima(avo.nak,d=1,D=1, stepwise = FALSE,approximation = FALSE, trace=TRUE)
checkresiduals(arima.nak)
print(summary(arima.nak))  #residual sd=210.11
#model for mombasa
arima.mom<-auto.arima(avo.mom,d=1,D=1, stepwise = FALSE,approximation = FALSE, trace=TRUE)
checkresiduals(arima.mom)
print(summary(arima.mom))  #residual sd=352.58

#creating the forecast graphs
nai.fcst<-forecast(arima.nai)
f1<-autoplot(nai.fcst)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                         panel.background = element_blank(), axis.line = element_line(colour = "black"))
kis.fcst<-forecast(arima.kis)
f2<-autoplot(kis.fcst)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                          panel.background = element_blank(), axis.line = element_line(colour = "black"))
eld.fcst<-forecast(arima.eld)
f3<-autoplot(eld.fcst)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                        panel.background = element_blank(), axis.line = element_line(colour = "black"))
nak.fcst<-forecast(arima.nak)
f4<-autoplot(nak.fcst)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                        panel.background = element_blank(), axis.line = element_line(colour = "black"))
mom.fcst<-forecast(arima.mom)
f5<-autoplot(mom.fcst)+theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
                        panel.background = element_blank(), axis.line = element_line(colour = "black"))
plot_grid(f1,f2,f3,f4,f5,ncol = 3)
```

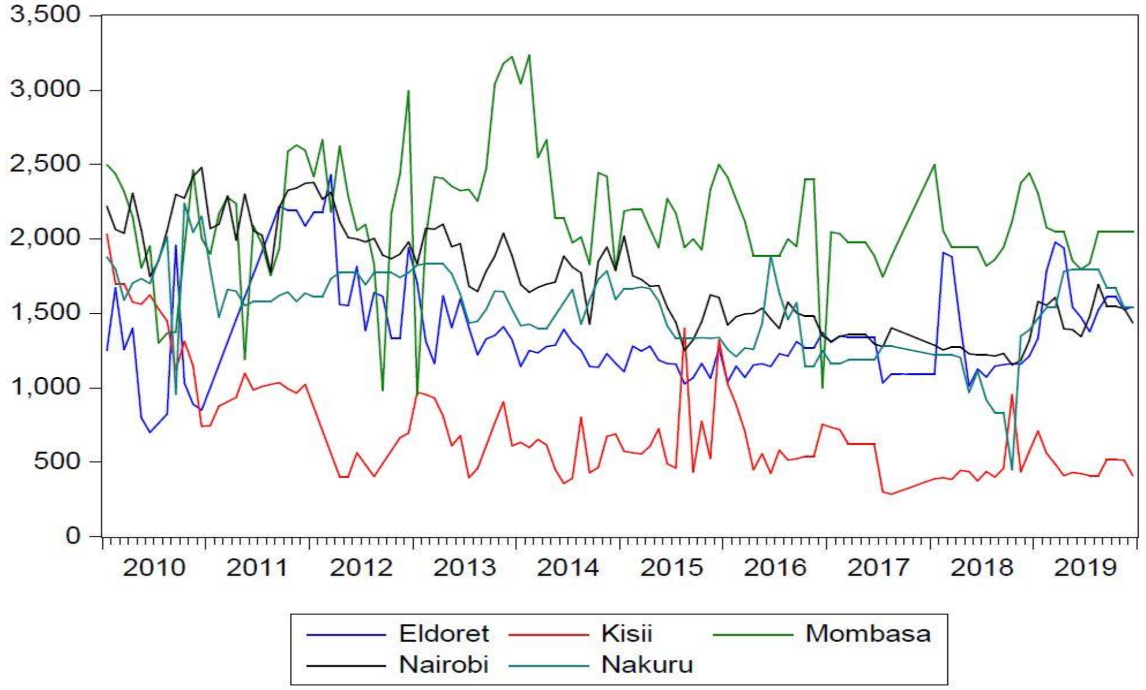
```
#checking for comovement
#work with the dataset clean
library(Hmisc)  #library for correlation
head(clean)
class(clean)
rcorr(clean, type="pearson") # type can be pearson or spearman
rcorr(as.matrix(clean))   #include p values
#plotting correlogram
library(GGally) #package to plot
p5<-ggcorr(clean)
p5<-ggcorr(clean,
           label = TRUE,
           label_alpha = TRUE)
p5



#download the clean dataset to continue analysis with Eviews
library(xlsx)
setwd("F:/Project/Data Analysis")
write_xlsx(x = temp, path = "cleandata.xlsx", col_names = TRUE)
```

Annex 3: Boxplot of the timeseries data showing outliers

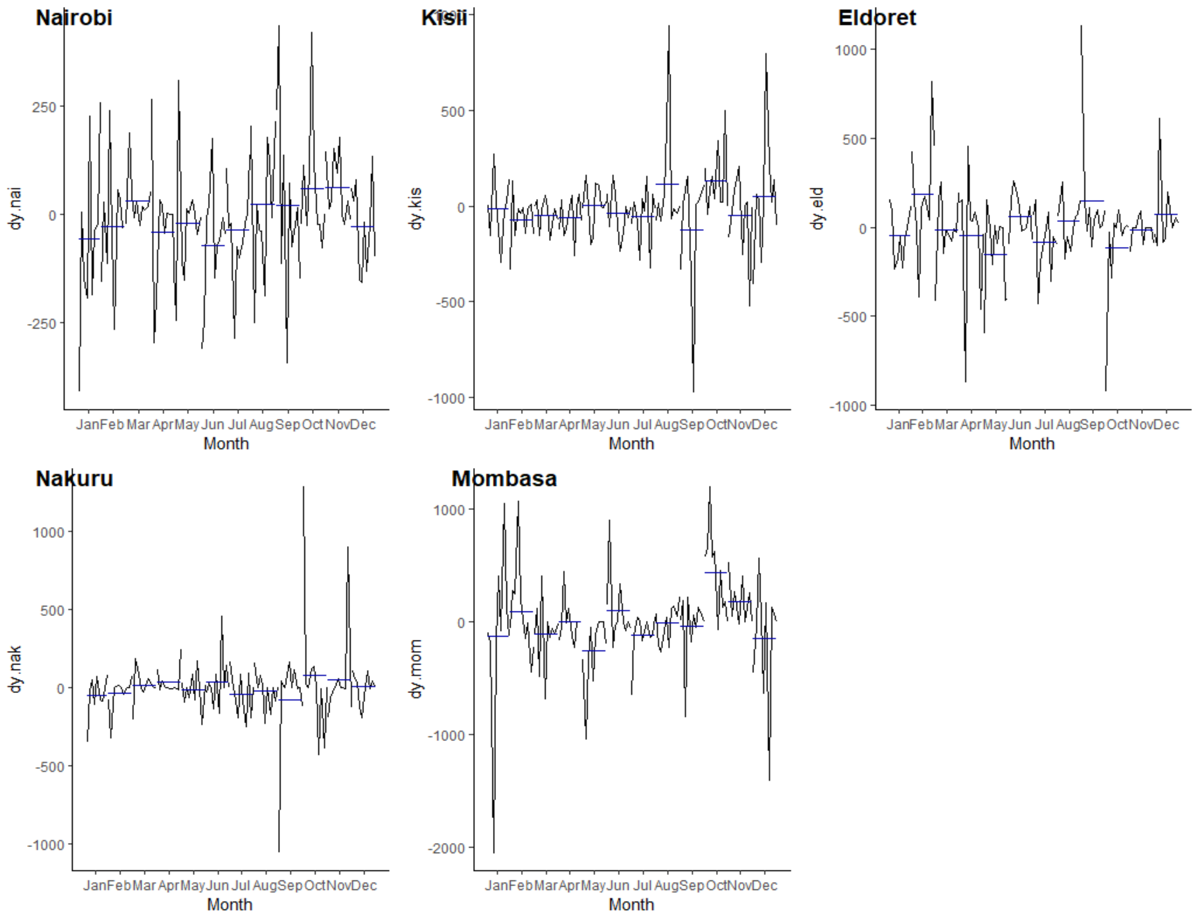


Observation: there are a few outliers in data from Kisii, Eldoret, Nakuru and Mombasa

Annex 4: The local avocado market prices

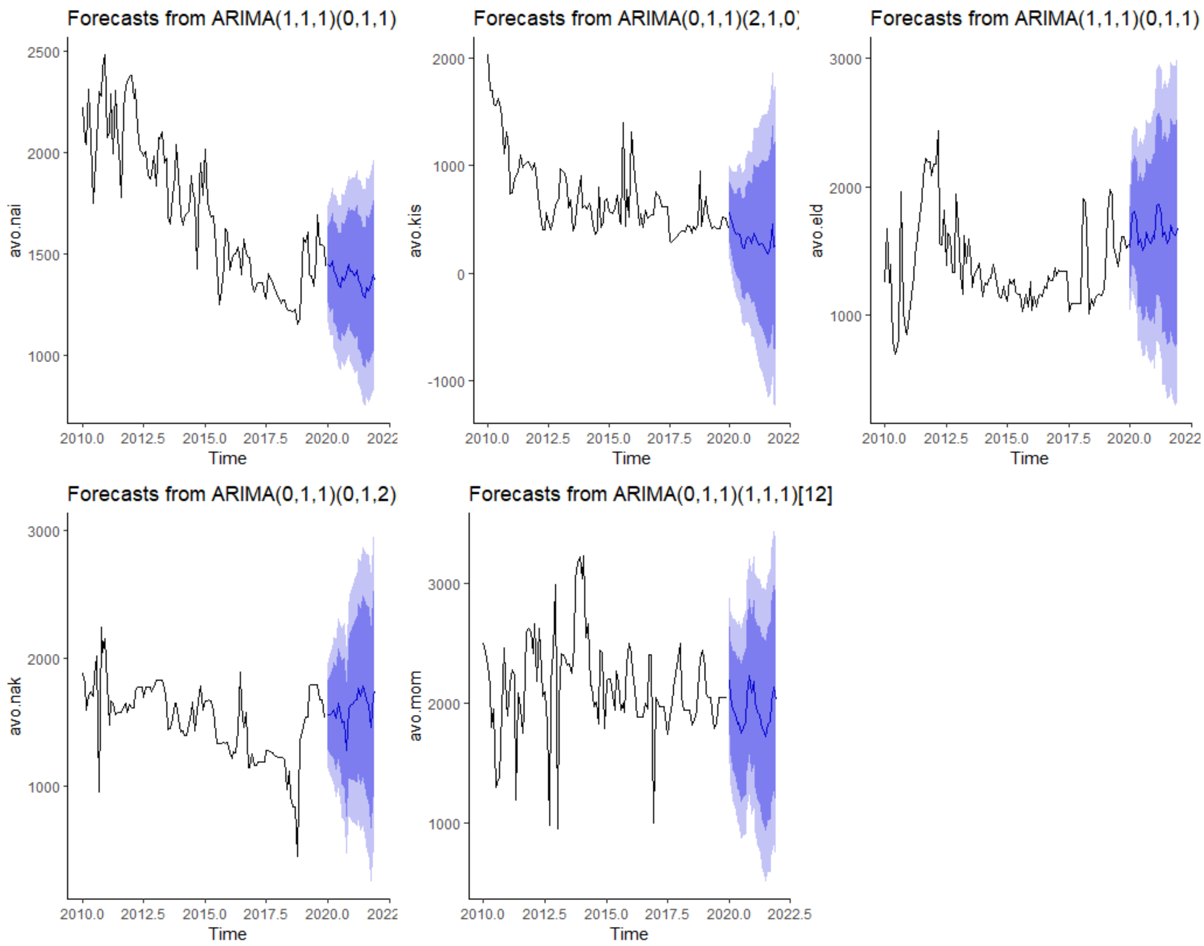


Annex 5: Seasonality graphs for the selected domestic markets



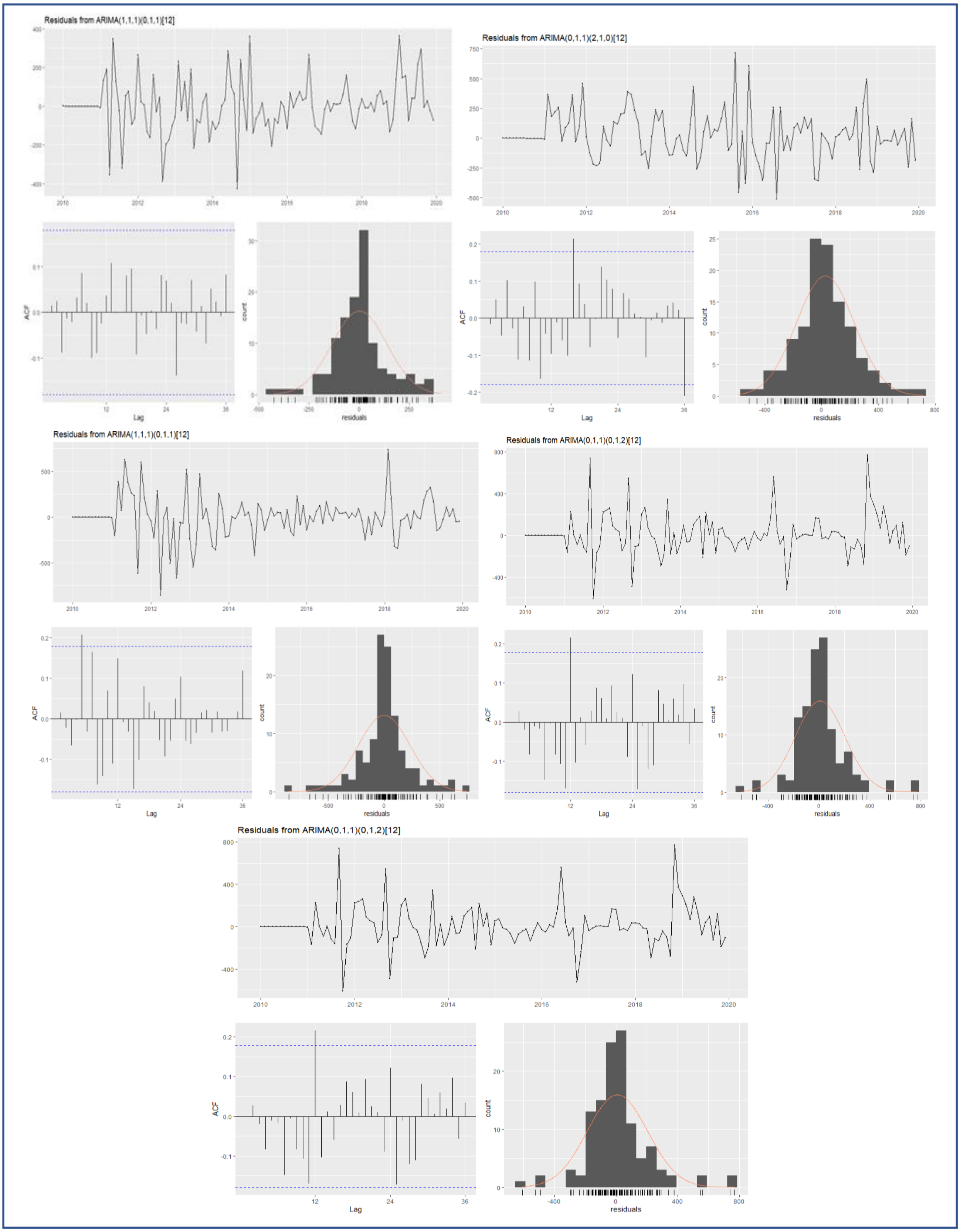
Observation: Seasonal price variations exist in all markets

Annex 6: Price forecast for the selected domestic markets



Observation: Series are likely to exhibit the same downward trend observed in the last ten years

Annex 7: Residuals from the Arima models



Observation: Minimal or no autocorrelation in the residuals; models are good for prediction

Annex 8: ADF at level

	Model	t. Statistic	P. Value
Prices in Nairobi	Intercept	-2,4500	0,1305
	Trend and intercept	-4,6510	0,0014
	None	-0,8598	0,3415
Prices in Kisii	Intercept	-3,6814	0,0056
	Trend and intercept	-4,0782	0,0089
	None	-2,3320	0,0196
Prices in Nakuru	Intercept	-4,4509	0,0004
	Trend and intercept	-5,2014	0,0002
	None	-0,9119	0,3193
Prices in Eldoret	Intercept	-4,2201	0,0009
	Trend and intercept	-4,2394	0,0054
	None	-0,8499	0,3458
Prices in Mombasa	Intercept	-6,4583	0,0000
	Trend and intercept	-6,4541	0,0000
	None	-0,5628	0,4714

Observation: Price series are non-stationary

Annex 9: ADF at 1st Differencing

	Model	t. Statistic	P. Value
Prices in Nairobi	Intercept	-12,8739	0,0000
	Trend and intercept	-12,8175	0,0000
	None	-12,9076	0,0000
Prices in Kisii	Intercept	-16,9029	0,0000
	Trend and intercept	-16,9295	0,0000
	None	-16,8889	0,0000
Prices in Nakuru	Intercept	-16,4751	0,0000
	Trend and intercept	-16,4190	0,0000
	None	-16,5425	0,0000
Prices in Eldoret	Intercept	-13,3610	0,0000
	Trend and intercept	-13,3098	0,0000
	None	-13,4193	0,0000
Prices in Mombasa	Intercept	-9,6385	0,0000
	Trend and intercept	-9,5949	0,0000
	None	-9,6805	0,0000

Observation: All series are stationary at 1st differencing

Annex 10: OLS Regression Model for all markets

Dependent Variable: KIS_DF1
Method: Least Squares
Date: 04/29/20 Time: 16:29
Sample (adjusted): 2010M02 2019M12
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MOM_DF1	-0.049425	0.049767	-0.993136	0.3227
NAI_DF1	0.054266	0.142671	0.380358	0.7044
NAK_DF1	-0.081192	0.098731	-0.822354	0.4126
ELD_DF1	-0.081394	0.087234	-0.933058	0.3528
C	-13.49664	19.50090	-0.692104	0.4903

R-squared	0.016327	Mean dependent var	-13.63650
Adjusted R-squared	-0.018188	S.D. dependent var	210.5566
S.E. of regression	212.4627	Akaike info criterion	13.59652
Sum squared resid	5146005.	Schwarz criterion	13.71329
Log likelihood	-803.9928	Hannan-Quinn criter.	13.64394
F-statistic	0.473049	Durbin-Watson stat	2.770575
Prob(F-statistic)	0.755425		

Annex 11: ADF test on residuals (all markets)

Null Hypothesis: RESIDALL has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.55749	0.0000
Test critical values:	1% level	-2.584707	
	5% level	-1.943563	
	10% level	-1.614927	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESIDALL)
Method: Least Squares
Date: 05/27/20 Time: 21:17
Sample (adjusted): 2010M03 2019M12
Included observations: 118 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDALL(-1)	-1.394413	0.084216	-16.55749	0.0000

R-squared	0.700875	Mean dependent var	1.730185
Adjusted R-squared	0.700875	S.D. dependent var	349.0774
S.E. of regression	190.9186	Akaike info criterion	13.35001
Sum squared resid	4264638.	Schwarz criterion	13.37349
Log likelihood	-786.6506	Hannan-Quinn criter.	13.35954
Durbin-Watson stat	2.070117		

Annex 12: OLS Kisii and Eldoret

Dependent Variable: KIS_DF1
Method: Least Squares
Date: 04/29/20 Time: 15:44
Sample (adjusted): 2010M02 2019M12
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELD_DF1	-0.041739	0.078068	-0.534652	0.5939
C	-13.53308	19.36133	-0.698975	0.4860

R-squared	0.002437	Mean dependent var	-13.63650
Adjusted R-squared	-0.006089	S.D. dependent var	210.5566
S.E. of regression	211.1966	Akaike info criterion	13.56012
Sum squared resid	5218669.	Schwarz criterion	13.60683
Log likelihood	-804.8271	Hannan-Quinn criter.	13.57909
F-statistic	0.285853	Durbin-Watson stat	2.802461
Prob(F-statistic)	0.593905		

Annex 13: ADF on residuals (Kisii and Eldoret)

Null Hypothesis: RESIDKE has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.91154	0.0000
Test critical values:	1% level	-2.584707	
	5% level	-1.943563	
	10% level	-1.614927	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESIDKE)
Method: Least Squares
Date: 05/27/20 Time: 21:26
Sample (adjusted): 2010M03 2019M12
Included observations: 118 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDKE(-1)	-1.411314	0.083453	-16.91154	0.0000

R-squared	0.709670	Mean dependent var	1.810932
Adjusted R-squared	0.709670	S.D. dependent var	353.5502
S.E. of regression	190.5010	Akaike info criterion	13.34563
Sum squared resid	4246002.	Schwarz criterion	13.36911
Log likelihood	-786.3922	Hannan-Quinn criter.	13.35516
Durbin-Watson stat	2.085033		

Annex 14: OLS Kisii and Nairobi

Dependent Variable: KIS_DF1
Method: Least Squares
Date: 04/29/20 Time: 16:05
Sample (adjusted): 2010M02 2019M12
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NAI_DF1	-0.006651	0.134240	-0.049545	0.9606
C	-13.68032	19.40397	-0.705027	0.4822

R-squared	0.000021	Mean dependent var	-13.63650
Adjusted R-squared	-0.008526	S.D. dependent var	210.5566
S.E. of regression	211.4522	Akaike info criterion	13.56254
Sum squared resid	5231310.	Schwarz criterion	13.60925
Log likelihood	-804.9711	Hannan-Quinn criter.	13.58151
F-statistic	0.002455	Durbin-Watson stat	2.804211
Prob(F-statistic)	0.960570		

Annex 15: ADF on residuals (Kisii and Nairobi)

Null Hypothesis: RESIDKN has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.97484	0.0000
Test critical values:	1% level	-2.584707	
	5% level	-1.943563	
	10% level	-1.614927	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESIDKN)
Method: Least Squares
Date: 05/27/20 Time: 21:27
Sample (adjusted): 2010M03 2019M12
Included observations: 118 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDKN(-1)	-1.413314	0.083259	-16.97484	0.0000

R-squared	0.711206	Mean dependent var	1.956147
Adjusted R-squared	0.711206	S.D. dependent var	354.0878
S.E. of regression	190.2853	Akaike info criterion	13.34336
Sum squared resid	4236393.	Schwarz criterion	13.36684
Log likelihood	-786.2585	Hannan-Quinn criter.	13.35290
Durbin-Watson stat	2.093141		

Annex 16: OLS Kisii and Mombasa

Dependent Variable: KIS_DF1
Method: Least Squares
Date: 04/29/20 Time: 16:12
Sample (adjusted): 2010M02 2019M12
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MOM_DF1	-0.041397	0.047901	-0.864211	0.3892
C	-13.79260	19.32327	-0.713782	0.4768

R-squared	0.006343	Mean dependent var	-13.63650
Adjusted R-squared	-0.002150	S.D. dependent var	210.5566
S.E. of regression	210.7828	Akaike info criterion	13.55620
Sum squared resid	5198237.	Schwarz criterion	13.60291
Log likelihood	-804.5937	Hannan-Quinn criter.	13.57516
F-statistic	0.746861	Durbin-Watson stat	2.781601
Prob(F-statistic)	0.389240		

Annex 17: ADF on residuals (Kisii and Mombasa)

Null Hypothesis: RESIDKM has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.74979	0.0000
Test critical values:	1% level	-2.584707	
	5% level	-1.943563	
	10% level	-1.614927	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESIDKM)
Method: Least Squares
Date: 05/27/20 Time: 21:28
Sample (adjusted): 2010M03 2019M12
Included observations: 118 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDKM(-1)	-1.402127	0.083710	-16.74979	0.0000

R-squared	0.705692	Mean dependent var	1.974597
Adjusted R-squared	0.705692	S.D. dependent var	351.5407
S.E. of regression	190.7113	Akaike info criterion	13.34784
Sum squared resid	4255385.	Schwarz criterion	13.37132
Log likelihood	-786.5224	Hannan-Quinn criter.	13.35737
Durbin-Watson stat	2.083481		

Annex 18: OLS Kisii and Nakuru

Dependent Variable: KIS_DF1
Method: Least Squares
Date: 04/29/20 Time: 16:25
Sample (adjusted): 2010M02 2019M12
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NAK_DF1	-0.050549	0.091571	-0.552025	0.5820
C	-13.78158	19.36059	-0.711837	0.4780

R-squared	0.002598	Mean dependent var	-13.63650
Adjusted R-squared	-0.005927	S.D. dependent var	210.5566
S.E. of regression	211.1796	Akaike info criterion	13.55996
Sum squared resid	5217829.	Schwarz criterion	13.60667
Log likelihood	-804.8176	Hannan-Quinn criter.	13.57893
F-statistic	0.304732	Durbin-Watson stat	2.802685
Prob(F-statistic)	0.581984		

Annex 19: ADF on residuals (Kisii and Nakuru)

Null Hypothesis: RESIDKNK has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.96739	0.0000
Test critical values:	1% level	-2.584707	
	5% level	-1.943563	
	10% level	-1.614927	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESIDKNK)
Method: Least Squares
Date: 05/27/20 Time: 21:29
Sample (adjusted): 2010M03 2019M12
Included observations: 118 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDKNK(-1)	-1.412735	0.083262	-16.96739	0.0000

R-squared	0.711025	Mean dependent var	1.987117
Adjusted R-squared	0.711025	S.D. dependent var	353.5348
S.E. of regression	190.0476	Akaike info criterion	13.34086
Sum squared resid	4225816.	Schwarz criterion	13.36434
Log likelihood	-786.1110	Hannan-Quinn criter.	13.35040
Durbin-Watson stat	2.096763		

Annex 20: Granger Causality

Pairwise Granger Causality Tests
Date: 06/04/20 Time: 22:23
Sample: 2010M01 2019M12
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
LAG_KISII does not Granger Cause LAG_ELDORET	106	0.07609	0.9268
LAG_ELDORET does not Granger Cause LAG_KISII		0.34574	0.7085
LAG_MOMBASA does not Granger Cause LAG_ELDORET	106	1.39994	0.2514
LAG_ELDORET does not Granger Cause LAG_MOMBASA		2.70680	0.0716
LAG_NAIROBI does not Granger Cause LAG_ELDORET	106	0.48316	0.6182
LAG_ELDORET does not Granger Cause LAG_NAIROBI		0.32622	0.7224
LAG_NAKURU does not Granger Cause LAG_ELDORET	106	0.80681	0.4491
LAG_ELDORET does not Granger Cause LAG_NAKURU		1.81820	0.1676
LAG_MOMBASA does not Granger Cause LAG_KISII	106	0.21335	0.8082
LAG_KISII does not Granger Cause LAG_MOMBASA		0.60638	0.5473
LAG_NAIROBI does not Granger Cause LAG_KISII	106	0.30070	0.7410
LAG_KISII does not Granger Cause LAG_NAIROBI		2.62341	0.0775
LAG_NAKURU does not Granger Cause LAG_KISII	106	0.11043	0.8956
LAG_KISII does not Granger Cause LAG_NAKURU		1.16913	0.3148
LAG_NAIROBI does not Granger Cause LAG_MOMBASA	106	0.67097	0.5135
LAG_MOMBASA does not Granger Cause LAG_NAIROBI		1.57619	0.2118
LAG_NAKURU does not Granger Cause LAG_MOMBASA	106	0.04601	0.9551
LAG_MOMBASA does not Granger Cause LAG_NAKURU		0.10534	0.9001
LAG_NAKURU does not Granger Cause LAG_NAIROBI	106	1.53213	0.2211
LAG_NAIROBI does not Granger Cause LAG_NAKURU		6.98891	0.0014

Observation: No causality relationships in the markets