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

by

Chalwe Sunga

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Agriculture (Agricultural Economics) in the Faculty of AgriSciences at Stellenbosch University

Supervisor: Dr Cecilia Punt
Co-supervisor: Mrs Lulama Ndibongo Traub

March 2017

Stellenbosch University https://scholar.sun.ac.za

Declaration

By submitting this thesis, I declare that the entirety of the work contained therein is my own original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2017

Copyright © 2017 Stellenbosch University
All rights reserved

Stellenbosch University https://scholar.sun.ac.za

Dedication

I dedicate this work to my father Samuel Sunga, my mother Annety Malwa Sunga, my sisters Mwaba and Ngosa, my brothers Sunga, Chabu, Chiluba, Kaluba and Lubilo and my fiancé Mwenya Kwangu. You are my greatest blessing and the reason for my hard work.

Stellenbosch University https://scholar.sun.ac.za

Acknowledgements

I would like to express my gratitude to a number of people that directly or indirectly helped to make this work a success. Firstly, I am indebted to my supervisor Dr. Cecilia Punt for the many roles she played in ensuring this thesis comes to a completion. I am grateful for her patience, concern, guidance and support rendered throughout my study. But most of all, thank you for the support when times were hard.

I am also grateful to my co-supervisor Mrs. Lulama Ndibongo Traub without whom this study would not have started in the first place. I thank her for introducing me to this interesting field of study, for the time accorded to me and the invaluable guidance rendered throughout my study. Special thanks go to Dr. William Burke of Stanford University for the assistance and guidance with specific portions of my analysis and Dr. Luke Mugode for the valuable comments on the initial drafts of this thesis.

I would also like to sincerely appreciate the government of Netherland through the Niche funders and Mulungushi University for financing my entire study. Special thanks go to Prof. Olusegun Yerokun and Dr. Moses Daura of Mulungushi University who played various roles and ensured finances were available on time.

My work would not have been a success without the many individuals that supplied me with data from various institutions across Zambia, DRC and Tanzania, who I also appreciate. Special thanks go to the entire staff membership of the Department of Agricultural Economics, for the comprehensive training I have received.

The prayers, support and encouragement I received from my family and friends also deserve a special mention. I particularly thank my parents for envisioning my academic life way before I imagined school existed. I thank them, as well as my siblings for the faith they have always had in my academic capabilities and every support I received. My friend Mavis and Pastor Progress, your support is always appreciated. Special thanks go to my dearest friend and fiancé, Mwenya Kwangu who had to endure long years of my absence. I also thank him for being such a pillar in so many aspects. I will forever be grateful.

Lastly but not the least, I offer my deepest gratitude to God for the gift of life, the opportunity to study at Stellenbosch and for bringing to completion what He began in me, all by His grace. Thank you Lord for your grace, the experience and valuable lessons learnt during my study period.

Stellenbosch University https://scholar.sun.ac.za

Abstract

Intra regional trade has the potential to contribute to food supply balance between surplus and deficit countries. However, this critical role can only be accomplished if surplus and deficit zones across countries are integrated. Most previous studies examining integration in food markets in Eastern and Southern Africa (ESA), partly attribute weak inter country market integration to restrictive trade policies and transfer costs. Yet, little evidence has been gathered to examine how international markets free from direct political influence may perform.

This thesis examines spatial integration between ESA dry bean markets where inter-market trade is predominantly conducted through informal channels. By focusing on a pair of markets in Zambia and Tanzania, and a pair of markets in Zambia and the Democratic republic of Congo (DRC), the study employed the Myers and Jayne (2012) extension of the Threshold Autoregressive (TAR) model, which explicitly incorporates transfer costs and allows the long run price equilibrium relationship to vary depending on the magnitude of inter-market bean trade. The analysis also adopted the Gonzalo and Pitarakis (2002) approach in locating the value and number of trade based thresholds. The study combined bean prices, transfer cost and trade volume data covering the period January 2006 to June 2016, for Kitwe, in Zambia and Lubumbashi, in the DRC; and Kasama, in Zambia and Mbeya, in Tanzania.

The empirical results revealed significant variations between the studied market pairs. Firstly, the study found no evidence to support informal trade based threshold effects in either market pairing, suggesting that the functioning of informal markets is independent of exogenous limitation to trade. Secondly, results indicated that there is a long run price equilibrium relationship between Kasama and Mbeya, implying that the two markets are integrated. In the case of Lubumbashi and Kitwe however, results indicated that the two markets are segmented. The latter finding implies that any significant price deviations above transfer cost between Lubumbashi and Kitwe may continue to grow without any tendency to equilibrium. Lastly, the adjustment process to price shocks, as measured by the speed of price transmission, is more rapid between Kasama and Mbeya markets (1.72 months) than Lubumbashi and Kitwe markets (5.3 months) despite both markets being dominantly connected by informal trade.

This study therefore concludes that unless other market operating environment aspects are improved, a policy focus on informal trade and intra-regional trade liberalization in Eastern and Southern Africa may not by itself always guarantee integrated intra-regional food markets.

It is therefore recommended that the food market operating environment be improved beyond simply liberalising regional trade.

Stellenbosch University https://scholar.sun.ac.za

Opsomming

Die potensiaal van intra-streekse handel om by te dra tot streeks voedsel balans tussen surplus en tekort lande, is geïdentifiseer. Hierdie kritieke rol kan egter slegs tot stand gebring word as surplus en tekort sones regoor lande geïntegreer word. Die meeste vorige studies wat integrasie in voedselmarkte in Oos- en Suider-Afrika (ESA) ondersoek, skryf swak tussen-land mark integrasie deels toe aan beperkende handelsbeleid en oordragskoste. Tog is min bewyse versamel om te ondersoek hoe die internasionale markte vry van direkte politieke invloed kan presteer.

Hierdie tesis ondersoek ruimtelike integrasie tussen ESA droëboon markte, waar tussen-mark handel hoofsaaklik gedoen word deur informele kanale. Deur te fokus op die markpaar Zambië en Tanzanië, en die markpaar Zambië en die Demokratiese Republiek van die Kongo (DRK), het hierdie studie die Myers en Jayne (2012) verlenging van die Drempel Outoregressiewe (TAR) model gebruik, wat oordragkoste insluit en die langtermyn ewewigsprys verhouding laat wissel na gelang van die omvang van tussen-mark boonhandel. Die analise het ook die Gonzalo en Pitarakis (2002) benadering gevolg om die waarde en aantal handel gebaseerde drempels te bepaal. Die studie het boonpryse, oordragkoste en handelvolume data gekombineer vir die tydperk Januarie 2006 tot Junie 2016, vir Kitwe, Zambië en Lubumbashi in die DRK; en Kasama, in Zambië en Mbeya, Tanzanië.

Die resultate dui op beduidende verskille tussen die bestudeerde markte. Eerstens, die studie het geen bewyse gevind om informele handel op grond van drumpeleffekte in enige van die markpare te ondersteun nie, wat daarop dui dat die funksionering van die informele markte onafhanklik van eksterne beperkings op handel is. In die tweede plek, dui resultate daarop dat daar 'n langtermyn prysewewig verhouding tussen Kasama en Mbeya is, wat impliseer dat die twee markte geïntegreer is. In die geval van Lubumbashi en Kitwe is egter bevind dat die twee markte gesegmenteer is. Laasgenoemde bevinding impliseer dat enige beduidende prysafwykings bo oordragkoste tussen Lubumbashi en Kitwe mag voortgaan om te groei sonder enige neiging tot ewewig. Ten slotte, die aanpassingsproses na prysskokke, soos gemeet deur die spoed van prysoordrag, is vinniger tussen Kasama en Mbeya markte (1.72 maande) as tusssen Lubumbashi en Kitwe markte (4.7 maande) ondanks die feit dat beide markte oorheersend verbind word deur informele handel.

Die studie se gevolgtrekking is dus dat tensy ander mark bedryfsomgewing aspekte verbeter, 'n beleid gefokus op informele handel en liberalisering van binne-streekhandel nie op sigself altyd geïntegreerde binne-streeks voedsel markte sal waarborg nie.

Dit word dus aanbeveel dat die voedselmark bedryfsomgewing verbeter word met meer as net die liberalisering van streekshandel.

Stellenbosch University https://scholar.sun.ac.za

Table of content

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

List of figures

Stellenbosch University https://scholar.sun.ac.za

List of tables

Stellenbosch University https://scholar.sun.ac.za

List of abbreviations

ADF	Augmented Dickey Fuller
AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
COMESA	Common Market for Eastern and Southern Africa
CSO	Central Statistics Office
DRC	Democratic Republic of Congo
EG	Engle and Granger
ERB	Energy Regulation Board
ESA	Eastern and Southern Africa
EWURA	Energy and Water Utilities Regulation Authority
FAO	Food and Agriculture Organization
FEWSNET	Famine Early Warning Network
FTA	Free Trade Area
GDP	Gross Domestic Product
GP	Gonzalo Pitarakis
HQ	Hannah-Quinone
ICBT	Informal Cross Border Trade
INS	Institut National de la Statistique
ITC	International Trade Centre
kg	Kilogram
km	Kilometers
KPSS	Kwaitkowski Philips Schmidt Shin
LOP	Law of One Price
MSECM	Markov Switching Error Correction Model
MSM	Markov Switching Model
MSVAR	Markov Switching Vector Autoregressive Model
MT	Metric Tonnes
NBS	National Bureau of Statistics

Stellenbosch University https://scholar.sun.ac.za

OLS	Ordinary Least Square
PBM	Parity Bound Model
PP	Phillips Perron
SADC	Southern African Development Community
SEECM	Single Equation Error Correction Model
SSA	Sub-Saharan Africa
TAR	Threshold autoregressive model
USA	United States of America
USD	United States Dollar
VAR	Vector Autoregressive
VECM	Vector Error Correction Model

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 1: INTRODUCTION

1.1 Background

The occurrence of the global food price crisis in 2007/2008 brought food markets to the forefront of world attention. As stakeholders examined this event, one of their key concerns was the transmission of these prices into domestic markets of developing countries, mainly because the poorest people spend most of their income on food (Cranfield, Preckel and Hertel, 2007). For policy makers in Eastern and Southern Africa (ESA), however, the functioning of food markets and food price behaviour has long been a subject of attention. Various initiatives, including market liberalization, have since been pursued in an attempt to eliminate main market distortions and ultimateley stimulate an efficient and integrated agricultural market. Yet, in many of these countries, achieving such a market system still remain a dream.

At the same time, global agricultural trade has been undergoing tremendous directional shifts. The emerging trading patterns seem to indicate that countries are increasingly trading within regional and subregional economic blocs as oposed to trading with countries overseas (Amikuzuno, Setsoafia & Seini, 2015; ITC [International Trade Centre], 2016b). Although official statistics of intra-regional trade are yet to substantially improve in ESA, the existence of various bilateral and multilateral trading agreements is testimony of how the region is a part and parcel of this development.

Amikuzuno *et al*. (2015), explain that the push for intra regional trade liberalization is based on two main reasons, (1) that the well-functioning of markets would ensure smooth exchange of goods between surplus and deficit countries and (2) that the price mechanism in well functioning markets will lead to economic efficiency and optimal allocation of resources. Integration of food markets therefore lies at the heart of modern debates concerning market liberalization (domestic and regional) and price stabilization policies (Golleti & Babu, 1994; Baulch, 1997a). It is also argued that market integration is a precondition for successful economic integration (Artingi-Ego, Opoloti & Drale, 2006).

By definition, spatial market integration refers to the extent to which commodity markets in geographically separated locations share a common long run price equilibrium relationship on a homogenious good. Barret and Li (2002) consider two markets as integrated, if there is tradability and contestability between them. The authors describe tradability as physical flow of a commodity between markets and contestability as when arbitrage between markets is fully exploited, leaving market agents indifferent about trading (Barret & Li, 2002). In Fackler and Goodwin (2001), market integration is simply the extent to which supply and demand shocks arising in one market location is transmitted to other market location(s). Price transmission is therefore at the core of integration analysis (Goodwin & Schroeder, 1991; Goletti *et al.,* 1995; Kabbiri *et al.,* 2016) and hence the two terms are used interchangeably. It occurs when a change in the price of a good in one location, causes a price change in a similar good in another location.

Stellenbosch University https://scholar.sun.ac.za

This concept has become a major issue over the past few decades because of price stabilization and food security concerns (Akhter, 2016). In the absence of market integration, price signals will not be transmited between food deficit and food surplus areas (Baulch, 1994; Baulch, 1997a; Muyatwa, 2000), agricultural producers will fail to specialise according to their comparative advantage (Baulch, 1997a), macro level price stabilization policies will not effectively influence micro level decisions (Moser *et al.*, 2009) and most policy objectives in the agriculture sector will be undermined (Baulch, 1994). In addition, a well integrated market system will ensure regional balance between deficit and surplus zones and between food and non food producing regions (Delgado,1986; Muyatwa, 2000). The importance of integration analysis has also been stressed on the basis that it sheds light on (1) how long a localised scarcity can be expected to last (Ravallion, 1986), (2) the extent to which a country (region) is vulnerable to external market shocks, and (3) spatial market efficiency (an economic equilibrium condition whereby all potential profitable arbitrage opportunities are exploited) (Barret & Li, 2002; Negassa, Myers & Gabre-Madhin, 2003).

In the context of ESA, agriculture plays an important role in the regional economy and a key contributor to GDP and employment (Van Rooyen, 2000). Because prodution is mostly rainfall dependent, ecological conditions create disparities between suplus and deficit regions. Yet the main consumption zones rarely coincide with the main production areas. This is particularly true for the regions food staple, beans[1](*Phaseolus vulgaris*) (Hillocks *et al*., 2006). According to FAO (2016), Rwanda and Burundi have the highest bean per capita consumption in ESA. Malawi, Zimbabwe and Kenya are the main importers (Katungi *et al.*, 2009) while Tanzania, Uganda and Kenya are the major bean producers (Akibode & Maredia, 2011; Siddiq & Uebersax, 2013; FAO, 2016). A well integrated market system is thus crucial in bridging up these supply disparities.

In Zambia, beans are the second most important legume crop after groundnuts. They have since been identified as holding potential for food security and income generation and therefore a national policy target crop for crop diversification (Hamzakaza, 2014). Estimates indicate that for the ten-year period between 2004 and 2013, Zambia produced an average of 59,408 tons of beans annually (Tembo & Sitko, 2013).

Despite its economic importance and the national recognition, Zambia remains a net importer of beans. Local production supplies about 60% of total national demand, while the remaining 40% is supplied by imports mainly from ESA and particularly Tanzania (Hamzakaza *et al.*, 2014; Muimui *et al.*, 2016). By FAO (2016) records, Tanzania is the world's 7th largest producer of beans and is also bean self-sufficient (Bese *et al.*, 2009). Katanga province in the Democratic Republic of Congo (DRC) is Zambia's key export destination. However, official statistics overshadow this fact and understate the importance of intra ESA beans trade because a significant volume is traded through informal channels. For instance, while Zambia's official imports (exports) from the rest of the world was recorded at 410 tons (7,263 tons) in 2015, informal imports (exports) from Tanzania and DRC alone stood at 7,263 tons (3,792 tons) (ITC, 2016b; FEWSNET). These figures alone suggest that informal trade with the selected

[1] For purposes of this study, the term "beans" refer to common edible dry beans.

Stellenbosch University https://scholar.sun.ac.za

partners constituted about 96% and 35% of total (formal plus informal) imports and exports respectively. Yet by ITC (2016b) statistics (formal) the three countries barely trade in beans with each other.

Lesser and Moisé-Leeman (2009) defines informal cross border trade (ICBT) as trade consisting of commodities exchanged across the border that either pass through unofficial routes avoiding customs checks and recording points, or go through the official routes, gets subjected to customs control, but involve illegal practices at the customs office, such as deliberate misclassification of the goods, under invoicing, bribery and misdeclaration of the country of origin. It is also trade in merchandise, that may be legal imports or exports on one side of the border and illegal on the other side or *vice versa*, on account of not having been subjected to statutory border formalities such as customs clearance (Afrika & Ajumbo, 2012). The informality therefore only lies in the fact that the trader directly or indirectly escapes regulatory border procedures set by government (Ogalo, 2010) but the trade itself has strong ties with the formal sector (Little, 2010). Ogalo (2010) and Little (2010) are further of the view that cross border trading through informal channels will continue to thrive due to the several challenges associated with formal trade.

Given the important role of beans in Zambia as elsewhere in ESA, the relevance of informal channels and the importance of integrated markets in bridging inter-country supply-demand gaps, it is imperative that an investigation into the spatial integration of bean markets linked by informal trade is conducted. Such an investigation forms the core objective of this study.

Several authors have studied market linkages within the framework of market integration. Such studies vary across commodities, countries and statistical methodology. Nevertheless, very little research has focussed on pulses and bean markets in particular. To the best of my knowledge, the only studies in ESA (and Africa) exploring dry bean cross border markets from an integration or price transmission perspective are Korir *et al.* (2003) and Mauyo *et al*. (2007), both of which focussed on markets dominated by formal trade between Tanzania and Kenya, and Tanzania and Uganda respectively.

This study aims at extending this literature by analysing intra-regional market integration in ESA, focussing on the Zambian dry bean markets dominantly connected through informal trade to Tanzania and DRC.

1.2 Statement of the problem

Zambia, Tanzania and the DRC are prime examples of informal bean trading partners in ESA. The three countries are all members of the Southern African Development Community (SADC), while Zambia and DRC are also signatory to the Common Market for Eastern and Southern Africa (COMESA). Under the SADC trade protocol of 2008, dry beans trade among member countries attracts a 0% tariff (Bese *et al.*, 2009). Under COMESA's simplified trade regime policy, consignments valued at US$1,000 or less are expected to clear with minimal paper work and little or no border inspection.

Stellenbosch University https://scholar.sun.ac.za

Despite these concerted measures, statistics reveal that cross border bean trade throughout ESA, is predominantly conducted through informal channels. This may be due to the fact that although these agreements have reduced physical tariffs, policies regulating intra-regional formal trade in pulses still involve high bureaucratic government procedures (Korrir *et al.*, 2003; ITC, 2016a), that leaves formal trade less attractive especially for small scale traders. Myers and Jayne (2012) demonstrate that domestic and border regulation policies have a direct effect on the functioning of food markets and spatial market integration in particular. Formal trade procedures further increase the cost of trading and has capacity to limit the extent of spatial arbitrage exploitation.

One can thus argue that bean markets dominated by informal trade may be functioning well given that they are connected by a free trading regime and are not bound by border protocol and general government influence. It is also possible, however, that these markets may be largely unintegrated given that international trade through informal channels is marred with several risk factors, that have capacity to increase market costs and negatively affect market integration. These include risk of good confiscation once caught (Little, 2010), inadequate market infrastructure, inefficient market information (Little, 2010; Afrika & Ajumbo, 2012), border harassment (Little, 2010) and the perpetual civil unrest in the DRC (FEWSNET, 2015). Moreover, empirical literature (Engel & Rogers, 1996; Kouyate & Von Cramon-Taubadel, 2016) consistently reveal that the mere presence of a border is reason enough to impede market integration. The question of whether such markets are integrated or not, therefore, remains an empirical issue, forming the central focus of this study.

Burke and Myers (2014) argue for the importance of such studies in giving insight into how cross border markets confronted by any other market related challenges except government influence would perform. This is particularly critical to SADC member states working towards the formation of a free trade area. Despite this fact however, little research has been conducted to assess and investigate the functioning of dry bean markets in ESA. It is particularly unclear whether cross border bean markets are integrated and in particular those linked by informal trade.

Korir *et al.* (2003) employed static price correlation method to examine integration of bean markets connected by formal trade between Tanzania and Kenya. The study conducted on a mere 24 monthly price observations (2000-2001), found very weak integration between selected markets. The authors attributed this to trade restrictions imposed by the two countries in form of export and import levies as well as the bureaucratic procedures in their bean international trade. Mauyo *et al.* (2007) also investigated market integration in bean markets connecting Kenya and Uganda and dominated by formal trade. Contrary to the findings in Korrir *et al.* (2003), their static price correlation results for a similar number of monthly price observations (24), indicated high levels of integration between the selected markets. The study similarly suggested a focus on eliminating trade obstacles between the two countries to enhance integration. These included road infrastructure development, import and export tariffs and other institutional barriers.

Stellenbosch University https://scholar.sun.ac.za

While these studies provide a baseline for this research, the use of price correlation method, much less based on very few observations make their conclusions unreliable. It has been proven over and above that prices in two markets may correlate even in the absence of market integration and vice versa (Ravallion 1986; Alexander & Wyeth, 1994; Barrett, 1996; Fackler & Goodwin, 2001). Moreover, even amidst their conflicting results, no study has followed up to examine dry beans cross border market integration in ESA.

The current trends relating to regionalisation of markets suggests countries will continue to depend on intra-regional trade to meet their domestic food demand. The fact that production of beans is concentrated in one or a few regions within and across countries of ESA, seasonal and influenced by ecological conditions and that the main consumption regions rarely coincide with the main supply points, suggests there exist a natural tendency to trade beans among ESA countries. If markets in the region are functioning well (integrated), price disparities between surplus and deficit markets will not vary excessively (Rashid & Minot, 2010). This entails that, there is a need to investigate market integration dry bean markets across ESA country borders, so as to ascertain the extent to which supply demand gaps, such as existing in Zambia and DRC, could be balanced out through intra-regional trade.

In addition, some studies have demonstrated that the degree of spatial price transmission may be likely to be sensitive to physical trade flows (e.g. González-Rivera & Helfand, 2001a; Myers & Jayne, 2012) while others have shown that high degrees of market integration can occur in periods of no trade between markets (Stephen *et al.*, 2012). Nevertheless, trade based thresholds have been argued on three grounds, (1) if no trade (zero trade) occurs between markets it may mean transfer costs are higher than price differential, no arbitrage opportunity exists and hence the likelihood of an equilibrium price relationship existing during this period is low (Burke & Myers, 2014), (2) medium trade may imply markets are working well to simply maintain the existing relationship (Myers & Jayne, 2012) and (3) a different trading regime may be observed at high trade volumes as physical flow reaches the capacity limit for transportation (Coleman, 2009; Burke, 2012).

It therefore becomes imperative to not only establish whether the selected Zambian dry bean markets are integrated with their Tanzanian and DRC counterparts but also explore conditions under which price transmission occurs. This study explored inter-market trade volume as a determinant of market integration.

1.3 Study significance

Cross border trade has great potential to contribute to domestic objectives of stable food prices as well as food security. However, this critical role can only be accomplished if surplus and deficit areas across countries are integrated. Given that the demand-supply gap in the Zambian dry bean industry demands an international trade orientation to achieve domestic bean sufficiency, it is crucial to examine the nature of existing market relationships with her neighbours. This study is therefore justified in order to establish the extent to which intra-regional trade and cross border trade, in particular, can be expected to contribute to domestic dry bean supply balance.

Stellenbosch University https://scholar.sun.ac.za

The study makes three contributions to the literature of food market integration literature. The first is that it examines dry bean cross border markets (instead of the commonly studied cereal markets) dominated by informal trade as opposed to formal trade. The second is that it employs a better measure of price transmission, multiple regime TAR model, which explicitly combines trade, price and transfer cost data as opposed to the static price correlation approach used in the only existing evidence on cross border bean market integration. Lastly, it provides evidence on the relative importance of trade volume on market integration.

Findings from this study are relevant for policy makers, especially those addressing regional related food security objectives and wish to influence food market integration between ESA countries. Also, policy makers in ESA have intervened less in domestic bean markets, relative to other markets such as maize. These factors should provide an optimal environment for smooth price transmission (Serra *et al.*, 2006). Findings from this study therefore will provide insight into how nearly liberal markets would operate at regional level. They could also act as baseline information for similar studies in future.

1.4 Study objectives

The general objective of this study was to determine the nature and extent of integration in cross border dry bean markets connected by informal trade between Zambia and Tanzania, and Zambia and DRC.

The study focussed on two specific market pairs, (1) Kasama in Zambia and Mbeya in Tanzania; and (2) Kitwe in Zambia and Lubumbashi in DRC under the following specific objectives:

1. Determine whether a common long run price equilibrium relationship exists in each market pair.
2. Measure the speed of price transimision between markets in each pair.
3. Analyse the effect of variations in inter-market trade flow volume on the degree and speed of price transmision.

1.5 Research questions

These objectives were met by addressing the main question: are dry bean markets dominantly linked by informal cross border trade between Zambia and Tanzania, and Zambia and DRC integrated?

To answer this question, three sub questions examined for the two market pairs, (1) Kasama in Zambia and Mbeya in Tanzania; and (2) Kitwe in Zambia and Lubumbashi in DRC were:

1. Is there a common long run price equilibrium relationship between each market pair?
2. How long does it take for a price change (shock) in one market to be fully transmited to the other market in each pair?
3. Does the degree of price transmission between each market pair vary depending on the amount of beans flowing between markets? If so, how?

Stellenbosch University https://scholar.sun.ac.za

1.6 Study hypotheses

The study explored the following hypotheses:

1. There is a long run price relationship between selected market pair, Kasama in Zambia and Mbeya in Tanzania; and Kitwe in Zambia and Lubumbashi in DRC. Hence markets are well integrated with each other.
2. Price transmission between the selected markets is rapid. This is because exploiting arbitrage opportunities between the selected markets is both time and cost saving given that traders do not follow formal international trade procedures and government control.
3. The level of intermarket informal trade has a significant effect on market integration.

1.7 Research methodology

Necessitated by the objectives and research questions, this study adapted the threshold Single Equation Error Correction Model (SEECM), introduced by Myers and Jayne (2012) as the main framework for measuring market integration. This framework has also been applied by Ndibongo *et al*. (2010), Burke (2012) and Burke and Myers (2014) and allows for price transmission dynamics to vary depending on the level of inter-market trade flow. Following this method, the study employed a combination of secondary time series datasets; retail bean prices, informal trade volumes and diesel price data obtained from several sources for the period January 2006 to June 2016. The analysis was conducted in 4 main steps:

Step 1: Identify the optimal trade based threshold values and select the optimal SEECM. The study used the Gonzalo and Pitarakis (2002) to test whether informal trade volumes has any significant effect in influencing integration. The presence of statistically significant thresholds implied estimating all steps below in multiple regimes otherwise the analysis proceeded on full sample data.

Step 2: Test individual time series for the presence of unit root and determine the level of integration. The study employed Augmented Dickey-Fuller (ADF), Phillips Perron (PP) and KPSS tests. The three tests together determined whether a series was considered stationary (I(0)) or non-stationary (I(1)).

Step 3: If bean and diesel price series in each market pairing were non-stationary (I(1)), the study analysed cointegration in order to establish whether prices shared a common long run relationship. The study used Engle-Granger two step and the Johansen cointegration methods. Information from Steps 2 and 3, then guided the choice between estimating a "stationary", "cointegration" or a "partial cointegration" price transmission model (refer to table 4.1).

Step 4: Estimate the appropriate SEECM model, and measure the speed and extent of dry bean price transmission between selected markets.

Data was analysed in STATA 14, while Microsoft Excel 2016 and Eviews 9 were also employed at various stages. (Refer to chapter 4 for a comprehensive discussion of data analysis).

Stellenbosch University https://scholar.sun.ac.za

1.8 Limitations of the study

This thesis was concerned with investigating spatial market integration within the framework of price transmission. Given that recent evidence of trade statistics between Zambia and her neighbours, as elsewhere in ESA, overwhelmingly support the importance of informal channels in regional bean trade, in comparison to formal trade volume figures, the study limited itself to investigating integration in markets dominated by informal trade. Tanzania and DRC were chosen because they are the major informal bean trading partners with Zambia.

The study strictly incorporated informal trade data as recorded by FEWSNET, even when in some instances some small proportions of formal data was recorded. As argued by several authors studying informal trade routes (e.g. Little, 2010; Ama *et al.*, 2013; Burke & Myers, 2014), it is almost impossible to get an exhaustive aggregate figure of informal transactions. This is because recording data for such trade, often carried out in small quantities per trader (e.g. 50 kg bag), constantly requires monitoring staff at the border. Secondly, traders may choose non monitored routes to cross products between countries or simply cross at night when monitors are off. Lastly, Zambia is linked to DRC through 3 border points; Kasumbalesa, Mokambo and Kipushi. Beans can cross through either of them. However, FEWSNET has monitors only at Kasumbalesa and Kipushi. Considering these issues, it is possible that trade volume figures used in this study may not be certainly exhaustive and may not escape measurement errors with potential to affect the trade based threshold determination. A thorough investigation into the accuracy, quality and reliability of trade data however falls beyond the scope of this study. Officials at FEWSNET however believe trade volume shares captured through their monitoring system is consistent and can thus be relied on as a good proxy for actual informal trade flows (Burke, 2012).

Furthermore, the trade data employed in this study are aggregate quantities of beans reported from the manned borders, 4 between Zambia and Tanzania and 2 between Zambia and DRC. As argued by Burke (2012) and Burke and Myers (2014), the aggregate figures represent a better proxy of the actual trade flowing between the investigated market areas than quantities from one particular border. However, investigations into whether or not the crossed beans are exclusively sold in the studied markets is beyond the mandate of this study.

Finally, the study examined the period January 2006 to June 2016 based on data availability. This period should be enough to get an insight of how bean markets in the region function. However, the commodity of study, beans (*Phaseolus vulgaris)*, was employed as mixed beans without any varietal distinctions. Any potential price differences across varieties are thus not captured. This study is also not concerned with assessing the impact of trade agreements existing between the three countries on bean price transmission (since informal trade defeats the importance of such agreements). Neither does it concern itself with empirically establishing the underlying causes of the findings linked to the long run relationship. The latter would expand the study scope, as this would require primary data from the three countries. Nevertheless, it will explore the relative importance of inter-market trade flow variations as well as the absence of government influence on the performance of cross border market

Stellenbosch University https://scholar.sun.ac.za

integration. Conclusions drawn are in turn of importance in effective formulation of regional and domestic trade policies.

1.9 Thesis outline

To achieve the goal of this research, the thesis is organised into six chapters. Chapter 1 has put the study into context highlighting the problem, objectives, justification and potential study limitations. Chapter 2 provides a detailed review of the relevant literature on the concept of market integration and price transmission in order to understand the underlying economic theories of this study. It begins with an in-depth discussion on the theoretical foundation, working definitions, econometric models applied in integration analysis over time and provides a basis for the model used in this study. The chapter further reviews previous studies in the three countries and across the globe with a special focus on agriculture commodity markets across country borders. Chapter 3 examines the bean industry in detail. The aim is to lay a foundation to understand the bean markets in Zambia, Tanzania and DRC. The chapter begins with a glance at the global industry, before fine tuning it to the ESA context at which level several aspects relating to the industry are discussed, including production, trade, consumption and related trade policies. The chapter further puts informal cross border trade into perspective. Chapter 4 outlines the methodological framework followed in analysing data. It describes the study area and data employed in the analysis before outlining procedures followed in time series property testing. The chapter also discusses the main econometric model for the analysis as well as the procedure followed in locating and selecting the threshold values. The results from these procedures are reported in chapter 5 beginning with a general description of data. Thereafter results covering unit root tests, cointegration and price transmission model for the Lubumbashi and Kitwe pair, are presented first, followed by the results for the market pair Kasama and Mbeya. Finally, chapter 6 concludes the thesis with a summary of the study, policy recommendations and suggestions for future research.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 2: A THEORETICAL REVIEW OF SPATIAL MARKET INTEGRATION

2.1 Introduction

The relationship between markets in space has a long history in economics. Geographical markets are particularly of relevance to agriculture given that the net production regions rarely coincide with the net consumption zones. Despite this importance, the first formal methodology to investigate market integration was only developed in the 1950s and first applied to agriculture in 1967 (Baulch, 1997a). Over the last decades, however, market integration analysis has been central to contemporary debates concerning market reforms, price stabilization policies and domestic market response to regional and global market shocks (Baulch, 1997a).

This chapter presents an in-depth review of relevant literature relating to the concept of spatial market integration in order to put the objectives of this study into their theoretical and analytical context. The chapter begins by giving an economic foundation of the subject of spatial market integration before defining key terminologies that help clarify the context in which they are used in this study. The rest of the chapter is structured as follows; section 2.4 discuses key drivers of market integration. Section 2.5 provides an in-depth discussion on a range of methods used to measure spatial integration. Given that the main modelling framework employed in this study allows for price transmission to vary depending on the level of trade, this discussion extends into section 2.6 where alternative methods used to select and locate thresholds are presented. The aim is to provide a rationale for the choice of the modelling framework and the procedure employed in locating thresholds in this study. Finally, the chapter reviews previous research on spatial market integration in agriculture, particularly focussing on cross border markets.

2.2 Theoretical foundation of spatial market integration

The theory of spatial market integration is founded on the equilibrium condition known as the Law of One Price (LOP) (Fackler & Goodwin, 2001; Amikuzuno, 2009). This law states that any homogeneous good traded easily and freely between geographically separated markets, should sell at the same common currency price, once transfer costs are accounted for. However, two main conditions must hold for the LOP to be valid (Officer, 1986). The first is that markets involved must be perfectly competitive so as to allow perfect spatial arbitrage. The second is that the underlying commodity must be homogeneous (identical). In essence, the logic behind this law is simple. If prices in one location are much lower, it would pay for a profit-seeking trader to buy a commodity in that location and sell it in a market location with higher prices. In the process, arbitrage eliminates the price difference below and above the transfer costs. The law thus predicts that spatial arbitrage restores equilibrium prices to equality across well integrated trading markets (Rapsomanikis *et al*., 2003).

Historically, the earliest conceptualization of this principle in economic theory of commodity markets can be tied to a tradition that stretches back to Augustus Cournot and Alfred Marshal

Stellenbosch University https://scholar.sun.ac.za

(Sexton, Cling & Carman, 1991). Cournot (1838, cited by Marshal, 1890, p.189) laid down a foundation for commodity market integration when he defined a market as:

> "*not any particular market place in which things are bought and sold, but the whole of any region in which buyers and sellers are in such free intercourse with one another that the prices of the same goods tend to equality easily and quickly.*"

Marshal (1890, p.189) reinforced Cournot's definition by contending that:

> *"the more nearly perfect a market is, the stronger is the tendency for the same price to be paid for the same thing at the same time in all parts of the market"*

He however, could not ascertain the extent to which conditions in one market would influence another market, especially for markets he classified as "large markets"[2]. This means price formation of a homogeneous good in one market location is a function of buyers and sellers actions in another market(s) (Harriss, 1979). Cassel (1918) extended this idea in his purchasing power parity theorem when he illustrated that the principle of spatial price parity or the LOP also applied to a bundle of goods.

To illustrate the LOP in international markets, consider the competitive price relationship given as (Ardeni, 1989):

$$P_t^a = P_t^b E \qquad (2.1)$$

where P_a and P_b are prices for a homogeneous good at time t, traded in domestic market, a and foreign market b respectively and E is the exchange rate. Equation 2.1 represents the LOP in its strict (absolute version) form (Ardeni, 1989). Each market has its own demand and supply, and hence autarky prices at each point in time. If no obstacle to free trade exist and a price difference enough to cover transfer costs emerges, an opportunity for making profit then exists by moving the commodity from lower priced market (e.g. market a) to a higher priced market (e.g. market b). Assuming this opportunity is exploited, demand and supply will increase in markets a and b respectively, resulting in a shift in equilibrium prices in each market. Arbitrage is expected to persist until prices differ only by transfer costs ($P_t^b - P_t^a = Tc$). At this point, the spatial arbitrage condition is attained and no incentive to trade exists. There are however several versions of this law (Fackler & Goodwin, 2001).

The fact that it is denoted as a "law" reflects the confidence placed in its ability to hold (Fackler & Goodwin, 2001). However, because of its restrictive assumptions that treats the world as frictionless and undistorted, it rarely holds in real world experiments. Even in highly traded commodities, Rogoff (1996) contends that adherence to the LOP is an exception rather than the rule. Despite this observation, McNew (1996) maintains, it is a necessary condition for market integration to occur and thus still forms the underlying theory of integration of spatial markets.

---

[2] He used an example of the western world as a large market (see page 189)

Stellenbosch University https://scholar.sun.ac.za

2.3 Spatial market integration defined

There is generally no consensus in the literature as regards what spatial market integration is (Fackler & Goodwin, 2001). McNew (1996) observed that market integration is not only less clearly defined, but also that most definitions are based on the analysts' statistical criteria, rather than the economic phenomena. Oftentimes, this concept is also confused with spatial market efficiency (Barrett & Li, 2002). The lack of coherence in the definition and its often alignment to empirical procedures therefore, makes it difficult to compare the growing research output across the globe. Nevertheless, there are definitions that appear to be more widely accepted in contextualizing the concept of spatial market integration.

Before reviewing these definitions, however, it is important to distinguish three types of market integration identified in literature: spatial (which is the focus of this study), vertical and cross commodity. Vertical integration is concerned with the pass through of a commodity price across stages of its marketing chain (producer-wholesale-retail-consumer) or product form. Cross commodity integration relates to integration between two commodities, for example diesel and crude oil. Spatial market integration is concerned with integration of geographically separated markets trading in an identical commodity. For purposes of this study and unless otherwise stated, market integration is used to mean the spatial type of market integration.

2.3.1 Market integration and price transmission

Early literature conceived spatial market integration as the price correlation between trading locations (Amikuzuno, 2009). In this view, spatially separated markets are said to be integrated whenever there is an instantaneous co-movement of prices across them (Goletti *et al.,* 1995; Fackler & Goodwin, 2001). Later on, integration became synonymous with the existence of the LOP and price efficiency (Amikuzuno, 2009). Gonzalez-Rivera and Helfand (2001a) however argue that there must be physical trade between markets in addition to sharing the same long run information (price), for markets to be integrated. The authors thus define integrated spatial markets as a set of geographical "locations that share both the same commodity and the same long run information". This definition underscores the primacy of trade flow as a key mechanism for markets to integrate. In other words, if no physical flow of a commodity between markets exist, integration cannot take place. They further associate larger volumes of inter-market trade to higher degrees of integration.

In contrast, Fackler and Goodwin (2001) recognize the importance of trade flows between integrating markets but argue that it is not a necessary condition for some degrees of market integration. They thus, in line with Negassa *et al.* (2003) define spatial market integration as the extent to which supply and demand shocks arising in one market location is rather transmitted to another market location(s). This understanding is also in line with McNew (1996), who earlier on contended that integration was the only mechanism through which changes in excess demand in one location can be transferred to another location. A point noteworthy in this definition is that market integration is expressed as a degree (completely segmented or perfectly integrated or anything in between), rather than a specific relationship.

Stellenbosch University https://scholar.sun.ac.za

One market can thus be more integrated with another, than is the other with it, depending on the degree of the price transmission ratio between them (Fackler & Goodwin, 2001).

Spatial integration of markets has also been said to occur if, when trade takes place between them, prices in the importing markets equalizes with those of the exporting market plus the transfer costs (Baulch, 1997a). Integration is thus an outcome of inter-market processes. In Barret and Li (2002) however, two markets are integrated if there is tradability and contestability between markets. The authors describe tradability to mean a commodity is traded between markets and argue that while evidence of positive trade flow is sufficient to signal spatial market integration, it does not necessarily imply price equalization. Barret (2001; 2005) is therefore of the view that market integration conceptualized purely on tradability, is consistent with Pareto inefficient distributions. The contestability aspect on the other hand, focuses on full exploitation of arbitrage rents and considers two markets as integrated when market agents face zero marginal returns, leaving them indifferent about trading (Barret & Li, 2002). Put together, Barret and Li's (2002) definition imply market integration refers to the transfer of excess demand or supply from one market to the other, which is evidenced through the transmission of price shocks and trade flow between markets.

In summary, a recurring theme from all the above definitions is that spatial market integration implies a smooth transfer of price information and signals across geographically separated markets (Muyatwa, 2000) and a price shock in one market is felt by another market. This conclusion is consistent with Goletti *et al.* (1995), Meyer (2004), Van Campenhout (2007) and Kabbiri *et al.* (2016), who describe spatial price transmission as the core of market integration. By definition, spatial price transmission occurs when a change in the price of a good in one location, causes a price change in a similar good in another location (Kabbiri *et al.*, 2016). The faster the adjustment takes place, the greater the degree of price transmission.

In this study, Fackler and Goodwin's (2001) definition will be used since the study recognizes that physical trade flow is not a necessary condition for markets to integrate although the level of trade may play a role in the degree of spatial price transmission. Market integration therefore is defined as: the extent to which supply and demand shocks arising in one market is transmitted to another market (s). Price transmission on the other hand is defined as the extent to which a price change in one market causes a price change in another market (Kabbiri *et al.*, 2016). Furthermore, since spatial price transmission forms the core of market integration analysis (Goodwin & Schroeder, 1991; Goletti *et al.*, 1995; Kabbiri *et al.*, 2016), the two concepts are used interchangeably in this study. This is common in many other integration analysis studies (see for example, Ndibongo *et al*. 2010).

2.3.2 Market efficiency

A close concept to market integration is market efficiency. As stated above (section 2.3.1), some studies have used these two concepts interchangeably and to mean the same thing (Barrett & Li, 2002; Negassa *et al.*, 2003). While they may be related, integration and efficiency are two distinct concepts that need to be treated as such (McNew & Fackler, 1997; Barrett, 2001; Barrett & Li, 2002; Negassa *et al.*, 2003). Understanding the distinction between these two

Stellenbosch University https://scholar.sun.ac.za

terms therefore becomes pertinent when one endeavours to interpret results from the analysis of either. As stated by Negassa *et al.* (2003), neither one of them is necessary nor sufficient for the existence of the other.

Spatial market efficiency is an economic equilibrium condition and a statement about welfare, whereby all potential profitable arbitrage opportunities are exploited (Barrett, 2001; Negassa *et al.,* 2003). A market is efficient if its price can fully and correctly reflect all relevant information about supply, demand and transfer costs (Fackler & Goodwin, 2001). In the case of spatially distinct markets, market efficiency requires that the long run competitive equilibrium is attained and that inter-market transfer costs are minimized (Barret, 2001). Although spatial markets can attain both integration and efficiency with or without physical trade flow, conditions under which they do so differ. In the absence of trade, market efficiency occurs when the spatial price differential is less than transfer costs (Negassa *et al.,* 2003). However, if the price differential is greater than transfer costs, the underlying markets are said to be inefficient with or without trade (Negassa *et al.,* 2003). On the other hand, integration in the absence of trade occurs if the underlying markets belong to the same trading network or in a case where a state institution fixes prices adjusted to all market location (Cirera & Arndt, 2008).

In summary, efficiency in spatial markets imply all arbitrage opportunities have been exploited a condition that need not be fulfilled for markets to be considered integrated.

2.3.3 Competitive market equilibrium

Other than efficiency, price transmission and integration itself, the study of spatial market integration makes reference to the economic concept of competitive market equilibrium and therefore needs to be contextualized in this study. Competitive market equilibrium relates to a condition when all extraordinary profits are exhausted through competitive pressures regardless of whether it results in inter-market trade flow or not (Barret & Li, 2002).

2.4 Factors affecting market integration

In practice, spatial market integration is discerned when two or more markets in different locations share a common long run price equilibrium relationship (Fackler & Goodwin, 2001; Rapsomanikis *et al.,* 2003). Approaches to analyzing such market linkages have since helped to tell which markets, for what commodities and what periods are integrated and which ones are not. A key question faced by analyst therefore regards why or why not a particular set of markets for a given commodity would be integrated, or why some markets within the marketing network may be more integrated than others for the same commodity. Goletti *et al.* (1995) contends that the answer to these questions lie in fundamental market conditions that influence the pass through of price shocks. Figure 2.1 presents a broad classification of these factors, followed by a brief discussion of each factor below. It is important to note, however, that while each of these factors has the potential to individually influence market integration, the influence is greater when a combination of them is at play.

Stellenbosch University https://scholar.sun.ac.za

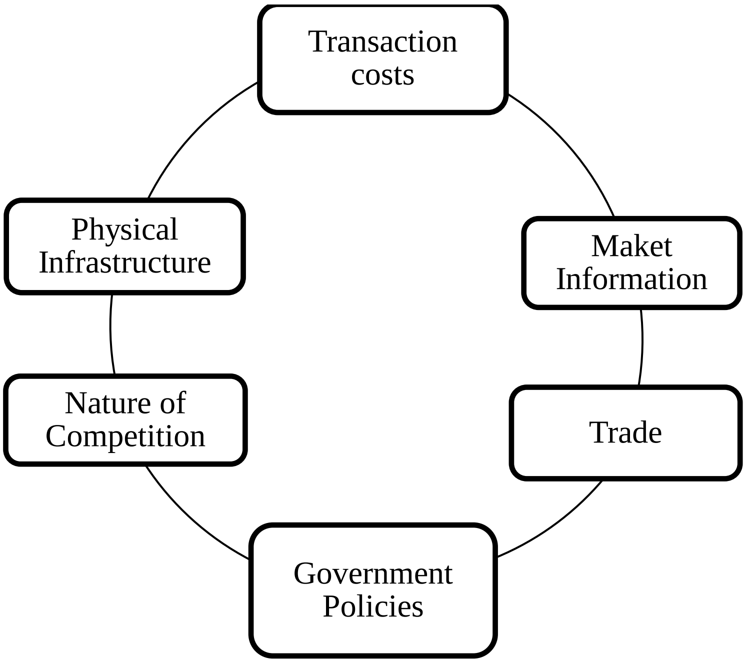


Figure 2.1: Factors influencing spatial market integration
Source: Goletti *et al.*1995; Rapsomanikis *et al*. (2003) and Kabbiri *et al.* (2016)

Transaction costs: the cost of transacting a commodity between markets play a crucial, if not the most important, role in determining the extent to which markets in space integrate. Sufficiently high transfer costs, which may arise due to limitations in transport capacity (such as poor roads) (Minten & Kyle 1999), inter-market distance, the time required to ship goods as well as legal requirements, impede efficient arbitrage that in turn limit the transmission of price signals (Sexton *et al.*, 1991; González-Rivera & Helfand, 2001b). Distance in particular has been found to contribute largely to trade costs, with every increase reducing the likelihood of markets to integrate (Kouyate & von Cramon-Taubadel, 2015).

The problem of transaction cost is further heightened by the presence of an international border between trading markets. Trading across borders comes with additional direct and indirect costs which increase trade costs and reduce chances of cross country markets integrating (Kouyate & Von Cramon-Taubadel (2015). Recent evidence (Engle & Rogers, 1996; Versailles, 2012; Kouyate & von Cramon-Taubadel, 2015) has since shown that there is a higher probability for two markets located in the same country to be integrated, than two markets situated in different countries even when distance between them is the same. In summary, transfer costs have an inverse effect on market integration, the lower the costs therefore the higher the chances of integration.

Physical infrastructure: closely related to transaction costs, is physical infrastructure. Kabbiri *et al.* (2016) notes that a market is a complex institution whose performance depends on several factors including storage, credit, communication and transportation facilities. Especially in cross country trade, poor marketing facilities increase transaction costs and consequently limit arbitrage as well as the flow of information between markets (Sexton *et al.*, 1991). As a result, changes in international markets may not fully transmit to domestic markets. Some studies

Stellenbosch University https://scholar.sun.ac.za

(Loveridge, 1991) have shown, for instance, that the quality of the road connecting markets has influence on spatial market relationships while others have (Minten & Kyle, 1999) found that transportation cost on poor roads can be twice as much as on simple gravel roads.

Nature of competition (market power): Although empirical evidence regarding the role of competition and market power in spatial food market integration is limited, some studies, such as that of Ndibongo *et al.* (2010), have attributed poor (lack of) market integration on market power. The law of one price also directly hinges on this factor, in that it is only valid if competition between markets is perfect (Officer, 1986). Imperfect competition in one or more markets or collusive behaviour among traders hinder market integration as traders (regions) with monopoly power may return price differentials higher prices than would be naturally determined by transfer costs (Rapsomanikis *et al*., 2003; Keats *et al.*, 2010).

Market information is another factor with a direct effect on spatial integration. The nature of information flow across markets (domestic or international) determine the extent to which traders exploit arbitrage opportunities. If traders possess perfect information regarding the market condition, they can project price changes and effectively exploit any arising arbitrage opportunities. Availability of market information is thus, the most basic factor determining effective trader participation in arbitrage.

Trade: Although spatial integration can occur in the absence of physical trade flow (Fackler & Goodwin, 2001; Stephens *et al*., 2012), trade volumes may play a role in the degree of spatial price transmission (Myers & Jayne, 2012). González-Rivera and Helfand (2001a) associates larger volumes of trade with high degrees of integration on the basis that they contribute to reducing transfer costs. Some empirical (Stephens *et al*., 2012; Myers & Jayne, 2012) evidence have however revealed the opposite.

Government policies: government border and domestic policies play an important role in the functioning of markets. Their actions, mainly through price stabilization policies, trade restrictions and regulations on credit, transport and exchange rates, directly influence market functioning and in turn affect integration of spatial markets positively or negatively. While distortions introduced in the form of domestic price support mechanisms, may weaken the link between domestic and international markets, policy instruments such as prohibitive import tariffs, export bans and other international trade intervention mechanisms may isolate domestic markets and prevent full transmission of price signals to and from international markets. Listorti (2009) argues that even the very existence of specific trading agreements that results in different trading blocs with varying degrees of market integration, can hinder cross country price transmission. Earlier on, article XXIV of the General Agreement on Tariffs and Trade (GATT) recognized this sentiment, identifying institutional support as a source of violations of the LOP Miljkovic (1999, cited by Listorti, 2009*).*

On the other hand, it is also possible that government intervention, especially in domestic markets, can force similar price changes in spatial markets, e.g. through direct price dictation (floor price), in which case a rather "non-natural" integration may occur. Such markets may have minimal influence from changes in international prices below and above the floor price

Stellenbosch University https://scholar.sun.ac.za

and hence limit their integration with international markets (Rapsomanikis *et al*., 2003). In general volatility in government intervention may work for the good or bad of markets (Goletti *et al.*, 1995).

In practice, however, conventional techniques for measuring market integration do not endogenously incorporate these determinants so as to disentangle each factors contribution to the findings. Conclusions drawn are simply based on whether or not markets share a common long run relation and to what extent. While this study will follow a similar path, it will however explore the relative importance of physical trade volume and the absence of direct government influence on integration of cross border markets.

2.5 Methods and techniques used to analyse spatial market integration

The issue of how to measure integration in spatially separated markets occupy a voluminous literature in economics and agricultural economics in particular. Variations in the definition of the concept and advances in time series econometrics have led to the development of a broad range of empirical procedures since the pioneering work of Lele (1967) and Jones (1968) (Baulch, 1997b). Nevertheless, Amikuzuno (2009) argues that none of the methods to date is so flawless that it can be regarded as appropriate in all cases and contexts. Researchers, therefore, choose a measurement approach based on the underlying research question but most importantly, data availability. In particular, Barrett (1996) and Van Campenhout (2007) note that the problem of data availability has in many instances reduced integration analysis to "price only" based techniques. Approaches that combine price, trade flow and transaction cost data, however make better inference on market integration (Baulch, 1997b; Myers & Jayne, 2012).

There are generally three main classes of methods used to analyse market integration; the static price correlation approach, dynamic methods and regime switching methods. Static price correlation approach is purely based on price data and measures integration by estimating the extent to which prices of an identical commodity in two markets correlate. Similarly, dynamic models rely on price data alone, but are not limited to a pair-wise test and distinguish between short and long term integration. These methods include the Granger causality (Granger, 1969), the Delgado variance decomposition (Delgado, 1986), the Ravallion model (Ravallion, 1986) and the standard cointegration methods (Engle & Granger, 1987). On the other hand, regime switching models combine price and transaction cost data at their minimum and analyse integration in a non-linear approach. Methods in this category include the Parity Bound Model (PBM) (Baulch, 1997a), the Markov Switching Model (MSM) (Hamilton, 2001) and Threshold Autoregressive (TAR) model (Tong, 1978; Balky & Fomby, 1997).

2.5.1 Static price correlation and bivariate methods

The earliest empirical conceptualization of spatial market interaction is associated with the point location model, first discussed by Enke (1951) and Samuelson (1952), and extensively developed by Takayama and Judge (1964). However, modern market integration analysis is

Stellenbosch University https://scholar.sun.ac.za

originally associated with the pioneering work of Lele (1967) and Jones (1968), who employed static price correlation method to measure integration in agricultural markets (Baulch, 1997b; Fackler & Goodwin, 2001).

This method tests integration by estimating the correlation of prices for an identical commodity in two markets (Hossain & Verbeke, 2010). The idea is that the extent of price formation in one market is related to the process of price formation in another market and thus can be indicated by price correlation coefficients (Harriss, 1979).

The basic structure as applied in early studies, takes the form of the LOP given as:

$$P_t^a = \alpha + \beta P_t^b + \varepsilon_t \qquad (2.2)$$

Where P_t^a and P_t^b are prices in spatially separated markets a and b at time t, ε_t is the error term and α and β are parameters to be investigated. Since no lags are included, the underlying assumption of this method is that arbitrage condition holds instantaneously. Markets a and b are then discerned as integrated if $\alpha = 0$ and $\beta = 1$. The value of the coefficients range between zero (absence of a relationship) and one (perfect correlation). The higher the coefficient therefore the higher the integration.

The advantage with this method is that it is simple to conceptualise, calculate and interpret results. Nevertheless, there are many limitations to the price correlation approach as a measure of market integration. The principal weakness is that prices in two markets may correlate even in the absence of market integration and *vice versa* (Harriss, 1979; Ravallion 1986; Alexander & Wyeth, 1994; Barrett, 1996). Fackler and Goodwin (2001) argue that a simultaneous effect of such common factors as inflation, climate change and production seasonality could result in price correlation even among markets that are not connected in any way. This approach therefore, has potential for spurious results and may not reflect the implication of spatial integration.

Other weaknesses of spatial price correlation are that, (1) it does not imply causality and is thus problematic when applied on markets with bidirectional trade flow (Timmer, 1974; Cirera & Arndt, 2008), (2) it is also limited to pair wise market analysis and cannot be applied on a system of markets (Delgado, 1986), (3) the assumption of instantaneous price adjustment and hence omission of lags, leads to measurements that completely omit dynamic price adjustment processes (Ravallion, 1986), and finally, (4) it may overstate the absence of market integration if a delay in the flow of market information causes a lag in the response to price changes (Barrett, 1996; Negassa *et al.*, 2003). In Barrett's (1996) view, price correlation method is simply weak and cannot be relied on for market integration analysis.

2.5.2 Dynamic methods

Recognizing the economic importance of market integration analysis and the multiple problems associated with bivariate correlation approach, scholars and researchers began making attempts to address these shortcomings. The mid 1980s saw some significant results of these efforts

Stellenbosch University https://scholar.sun.ac.za

mainly from the works of Delgado (1986) and Ravallion (1986). Earlier the causality test had been developed by Granger (1969). These authors pioneered what is known today as dynamic regression models of market integration which recognize lagged relationships and therefore account for dynamic nature of price relationships (Fackler & Goodwin, 2001). The underlying theoretical motivation is that markets may not instantaneously adjust to regional shocks in most cases, for instance because of the bulk nature of interregional trade and a delay in arbitrage activities caused by costly transportation, delivery lags and traders lack of perfect foresight of market conditions. They are thus considered a better measure of market integration despite being primarily price based tests. Some of these tests are discussed below.

2.5.2.1 Granger Causality test

According to Goletti and Babu (1994), if two markets exhibit a linear constant relationship in the long run (integrated), then there must be some causal influence running between them. Granger (1969) formalized this notion of causality in terms of lead and lag relationships. He defined causality between two series A_t and B_t as follows, A_t causes B_t if A_t poses information not available in B_t but helps forecast B_t (Fackler & Goodwin, 2001). The Granger causality approach thus measures the extent to which current and past price changes in one market location explain price changes in another market, as well as the direction of market influence (Baulch, 1997b). The cause-effect can be unidirectional, bidirectional or no causality at all.

According to Fackler and Goodwin (2001) this test is typically carried out within the Vector Autoregressive (VAR) framework, by regressing one market prices on the lagged values of the other. Baulch (1997b) however presents it in an error correction format estimated as:

$$\Delta P_t^A = \alpha_0 + \sum_{k=1}^{n} \alpha_k \Delta P_{t-k}^A + \beta P_{t-n-1}^A + \sum_{l=1}^{m} \theta_l \Delta P_{t-1}^B + \delta P_{t-m-1}^B + \varepsilon_t \qquad (2.3)$$

Where α_0 is the constant, *n* and *m* are the number of lags and P_t^A and P_t^B are prices in markets A and B at time *t*. Prices in Market B Granger-causes prices in market A if, $\delta = 0$ and $\theta_1 = 0$ for all *l*=1...m. If prices in A and B Granger-cause each other, it implies prices are simultaneously determined (Gupta & Mueller, 1982; Goletti & Babu, 1994; Baulch, 1997b). If on the other hand, one market causes prices in more than one market, without its prices being granger-caused by them, then such a market is referred to as a central market (Goletti & Babu, 1994).

While this approach may provide some inference regarding the existence of statistically significant lead/lag relationships among variables as well as allow for simultaneous price determination, it does not say anything about the nature of existing price relationships between markets (Fackler & Goodwin, 2001). In other words, accepting the causality hypothesis does not imply a cause-effect relationship between variables, but only means that price in one market can be used to explain price behaviour in another (Mutambatsere *et al.*, 2006). At all times, therefore, Granger causality test requires a supplementary test that can make inference on the actual nature of price relationship before drawing any valid conclusions regarding market

Stellenbosch University https://scholar.sun.ac.za

integration. Granger, (1969) also notes that since this method does not take into account seasonal patterns and non-stationarity in price series, it has potential for spurious results. Results from this test must thus be interpreted with care (Fackler & Goodwin, 2001) as they are neither necessary nor sufficient for the existence of market integration (Barrett, 1996).

2.5.2.2 Delgado Variance Component approach

This method is a direct attempt to address some of the weaknesses of static bivariate methods rather than an attempt to improve Granger causality test. Instead of pair wise testing, as in static bivariate models, Delgado (1986) pioneered a variance component method to measuring market integration for a whole marketing system, which in Negassa *et al.*, (2003) view is rather a variance decomposition approach. This test implicitly assumes constant transport and transaction costs and removes any existing common trends and seasonality in price series prior to testing for integration (Negassa *et al.*, 2003). Pairs of markets are then discerned as integrated when spatial price spreads between them equal the transfer costs for a given season.

The principle weakness of this method however is that it is based on a test of contemporaneous price relationships and never allows for dynamic relationships among prices (Negassa *et al.*, 2003).

2.5.2.3 The Ravallion model

Ravallion (1986) proposed an extension of the static bivariate method into a dynamic model of spatial price differences, that distinguishes between short and long run relations among price variables. This model permits Individual price series to have their own dynamic (autoregressive) structures while being interlinked with other market prices and hence, overcomes the inferential dangers of simple bivariate methods arising from correlation and seasonality (Mutambatsere *et al.*, 2006). The theoretical motivation for Ravallion's proposal, lie in the recognition that markets may take time to respond to shocks from other markets. Ravallion pictured a radial spatial market framework in which a group of subsidiary (regional) markets are linked to some central (single) reference market. He then argued that price formation in each regional market is a function of trade with the central market (Negassa *et al.*, 2003). Mathematically, the model is based on the regression of the general form (Ravallion, 1986):

$$P_{1t} = \sum_{j=1}^{n} \alpha_{1j} P_{1t-j} + \sum_{k=2}^{N} \sum_{j=0}^{n} b_{1j}^{k} P_{kt-1} + X_{1t} C_i + \varepsilon_{1t} \qquad (2.4)$$

Where P_{1t} and P_{1t} are prices in the regional market i and central market 1 at time *t* respectively, j is the lag length and X_{1t} are factors influencing regional markets such as seasonality and policy. Market segmentation is then discerned when $b_{1j} = 0$, implying that the central market does not influence price formation in the regional market. If on the other hand, price changes in the central market are instantaneously reflected in regional market, $b_{10} = 0$ and $\alpha_{1j} = b_{1j} = 0$, it denotes short run integration. Finally, long run integration, which imply sluggish response

Stellenbosch University https://scholar.sun.ac.za

in regional markets to changes in central market prices, occur when the restriction $\sum_{j=1}^{n} \alpha_{1j}$ $+\sum_{j=1}^{n} b_{1j}$ =0.

The Ravallion method has been a subject of modifications in recent years. Timmer (1987) revisited this approach by applying different assumptions regarding the central and the regional market prices. Faminow and Benson (1990) introduced base point pricing in the Ravallion model. Alexander and Wyeth (1994) applied econometric and cointegration techniques to the Ravallion model.

Despite its influence, the Ravallion model is not without limitation. Mutambatsere *et al.* (2006) argues that the principle weakness of the model lies in its own assumptions. The Ravallion model assumes that transfer costs will not significantly influence spatial market integration. Any violations to this assumption however, may significantly bias the results. Secondly, the central market hypothesis implies that the model can only capture integration between the central and regional markets with a direct linkage. A regional market may however, be indirectly linked to the central market, through subsidiary markets and still be integrated with it. Moreover, the central market is arbitrarily assumed rather than discovered from the data (Silvapulle & Jayasuriya, 1994). Barret (1996) also contends that the assumption of direct link and the fact that inter-seasonal trade reversals are ruled out, violates the radial market assumption, which forms the very basis of the Ravallion approach.

2.5.2.4 Cointegration analysis

Alternatively, integration under dynamic approaches is examined using cointegration analysis. While the above methods utilize econometric modelling techniques to test integration, they draw conclusion on an implicit assumption that price series are stationary. In reality, however, economic time series exhibit non-stationary behaviour more often than not (Engle & Granger, 1987). This is especially true for nominal price series, the basic dataset for integration analysis (Fackler & Goodwin, 2001).

A stationary time series is one in which the mean and variance do not vary systematically over time (time independent). In contrast, stochastic properties (mean and variance) of non-stationary data fluctuates substantially over time (time dependent). Granger and Newbold (1974) then argued that since integrated series often wonder (non-stationary), a simple regression of one on the other may wrongfully accept the hypothesis of integration, even when the series are independent. Fackler and Goodwin (2001) concurs with this sentiment by adding that, the presence of non-stationarity in price series invalidates inference of conventional approaches on market integration. The cointegration approach therefore not only offers an alternative measure of market integration, but also takes these issues into consideration and tests market integration by examining whether prices overtime, wonder within a fixed band (Baulch, 1997b).

This approach is originally associated with the work of Engle and Granger (1987). Even though cointegration was first formally introduced by Granger (1981, 1983), it was only studied further by Engle and Granger (1987), and Engle and Yoo (1987). Engle and Granger (1987) argued

Stellenbosch University https://scholar.sun.ac.za

intuitively that, while individual time series variables can be non-stationary (I (1)) in levels, there might be a combination of them that could in fact be stationary (I (0)). Any two or more variables exhibiting such patterns are thus said to be cointegrated, *i.e*. they may drift apart in the short run but still move closely together in the long run. In the context of market integration, markets are thus integrated if they exhibit co-integration, otherwise they are segmented.

Cointegration analysis as originally proposed by Engle and Granger (1987) is a stepwise procedure. In the first step, individual series are tested to determine the order of integration[3] using appropriate unit root tests. If two series are integrated of the same order (I(1)), the second step estimates the residuals (u_t), from a co-integrating regression generally given as (Brooks, 2008):

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t \qquad (3.5)$$

where y_t is the dependent variable at time t, β_1 is the constant, x_{2t} and x_{3t} are the independent variable at time t, and $u_t$is the error term, which is later tested for unit root, using appropriate unit root tests from step 1. Rejecting the null of unit root in the residuals means series are cointegrated. Additionally, the short run relationships can be determined by constructing an error correction model (Engle & Granger, 1987). In case series are cointegrated, the error correction model can also be used to detect which of the series move to restore equilibrium (Mann, 2012).

Despite its conceptual and computational simplicity, the outcome of this test might be affected by the choice of the dependent variable in the model (Goodwin & Schroeder, 1990; Konya, 2004). Moreover, the inability of the cointegration approach to explore all cointegration possibilities in a multivariate setting (Myers, 1994; Negassa *et al.*, 2003) means it can only handle pair wise tests.

Johansen (1988) revisited cointegration approach by employing maximum likelihood method to test for cointegrating relationships among several series. Unlike the Engle-Granger approach which relies on residuals, the Johansen (1988) approach relies on the relationship between the rank of the matrix and its characteristic roots (eigen values). The procedure is conducted on a general system presentment as (Muyatwa, 2000):

$$\Delta x_t = \Gamma_0 + \sum_{j=1}^{p-1} \Gamma_i \Delta x_{t-j} + \pi x_{t-p} + v_t \qquad (3.6)$$

where Δx_t is the vector of the first difference of endogenous variables (x_t), Γ_0 is the constant, π is the rank of the VECM matrix, Δx_{t-p} are vectors of the lagged values of the first of difference values and V_t is the normally distributed error term, and are (*n* x *n)* coefficient matrices of unknown parameters. In a multivariate case the number of cointegrating relations is used as an indicator for the strength of market integration (McNew & Fackler, 1997). For a bivariate case, a rank (π) of zero implies segmentation or no cointegration between variables

[3] Order of integration refers to the number of times a time series variable needs to be differenced in order to become stationary. See section 4.4.1 for a detailed discussion of unit root tests

Stellenbosch University https://scholar.sun.ac.za

while a rank of 2 means the variables were originally stationary (Muyatwa, 2000). A rank of one on the other hand implies cointegration.

Other modifications to the Engle-Granger test came with Phillips and Ouliaris (1988), who introduced two residual-based tests to the model; the variance ratio test and the multivariate trace statistic. Cointegration approach has since been widely adopted as a test for market integration beginning with Ardeni's (1989) application to internationally traded commodity markets (Fackler & Goodwin, 2001).

The principal advantage of this technique is that it is able to overcome problems associated with non-stationarity in time series and hence allow for a consistent inference on market integration. However, drawing conclusions merely on this approach has been criticized for not being sufficient to infer market integration (Baffes, 1991; Barrett, 1996; McNew & Fackler, 1997). The method disregards transfer costs and assumes they are constant through time. If, in reality, transfer costs are non-stationary during the period under review, cointegration analysis may wrongfully reject the hypothesis of integration (Barrett, 1996; McNew & Fackler 1997). Fackler and Goodwin (2001) supports this view by noting that price series in two markets may continue to drift apart with no sign of long run convergence simply due to non-stationarity (volatility) in transfer costs rather than unexploited arbitrage opportunities. Critiques of cointegration have also been made from the necessity and sufficiency perspective (Barrett, 1996; McNew & Fackler, 1997; Negassa *et al.*, 2003). Collectively the authors note that integrated markets need not be cointegrated. Finally, spatial markets may be segmented either because the inter-market margin is greater than transfer costs but for some reason such as information, the arbitrage opportunity is not exploited, or margins are lower than the transfer costs, in which case there is no incentive for arbitrage. Barret (1996) argues that both cases imply absence of efficient arbitrage and yet cointegration only captures the former.

Despite these critiques, cointegration analysis in its traditional form above, is arguably the most applied approach to measure integration. Modern market integration analysis, however, apply cointegration tests as only a pre-test (see for instance, Ndibongo *et al.*, 2010; Myers & Jayne, 2012; Burke & Myers, 2014). This is in line with Alexander and Wyeth (1994), and Fackler and Goodwin (2001) who recommend for complementary models to accompany cointegration results. It is for this reason that this study employed cointegration analysis only as a pre-test to market integration.

2.5.3 Regime switching models

The third class of models used to test integration is regime switching approach. Although dynamic techniques described in section 2.5.2 reflect considerable improvements to the static price testing approach, namely, dynamic adjustment structure, multiple market considerations, direction of influence, time series property considerations and the distinction between short and long run adjustments, they take a narrow view of market integration and rely solely on price data, ignoring the pivotal role of transfer costs and strictly classifying markets as either integrated or not. Baulch (1997b) argues that none of them in fact challenges the basis of price correlation coefficient, in which, markets are discerned to be integrated purely on the basis of

Stellenbosch University https://scholar.sun.ac.za

co-movement of prices, without reference to transfer costs. In Barrett's (1996) view, incorporating trade flows and transaction costs with price data could eliminate most of the earlier weaknesses.

Recent developments in the spatial dynamic literature are thus an attempt to explicitly combine these datasets. A more recent development, however, regards the realization that the speed of markets adjustment to external market(s) shocks, may vary depending on whether the size of the shock is below or above a threshold representing intermarket transfer costs (Myers & Jayne, 2012). Models that consider non-linear specifications may thus be more realistic representations of price transmission dynamics, forming the theoretical motivation for regime switching techniques. The class of models relax the assumption of constant speed to long run equilibrium relationship and allow for regime dependent, price adjustment speeds. In the next section, three of these models, the Parity Bound Model (PBM), Markov-Switching Models (MSM and Threshold Autoregressive (TAR) widely applied in market integration analysis are discussed.

2.5.3.1 The Parity Bound Model (PBM)

The PBM in its recent application is associated with the work of Baulch (1997a, 1997b). Even though Spiller and Haung (1986) and Spiller and Wood (1988) were first to develop the PBM, which was later modified and applied by Sexton *et al.* (1991), the most frequently applied version is Baulch's (1997b), after his innovative modifications and popularizations of the model. Zhao *et al.* (2012) remarks that the desire to improve the reliability of market integration tests led Baulch (1997) to propose the parity bound approach, which explicitly make use of all available market data (prices, trade flow and transaction costs) to examine intermarket relationships.

The PBM assumes that transfer costs determine price efficient bands, known as parity boundswithin which spatial market prices can vary independently (Barrett & li, 2002; Mutambatsere *et al.*, 2006). The central argument is that spatial market integration cannot be measured on price data alone, as such tests fail to account for non-linearity implied by transfer costs, as well as the discontinuity in trade and trade reversals.

Assuming P_t^a and P_t^b are prices of a homogeneous good in geographically separated markets a and b and that trade flows from a to b, PBM tests for integration in markets a and b by distinguishing among three possible trade regimes determined by transfer costs (T_t^{ab}) (Baulch, 1997a). The first regime is found *at the parity bound* where inter-market price differential equal transfer costs ($T_t^{ab} = P_t^b - P_t^a$). As long as there are no impediments to trade, spatial arbitrage will drive prices in the two markets to move together and the spatial arbitrage condition will eventually be fulfilled. Regime 2 occurs *inside the parity bound* where transfer costs are higher than inter-market price differences ($T_t^{ab} > P_t^b - P_t^a$). Trade will not occur since there is no opportunity for rational arbitrage, spatial arbitrage conditions are also not fulfilled. Finally, Regime 3 is located *outside the parity bound,* where transfer costs are lower than the inter-market price differential ($T_t^{ab} < P_t^b - P_t^a$). In this regime, spatial arbitrage conditions are violated whether trade occurs or not, a possible sign of trade impediments (Baulch, 1997a).

Stellenbosch University https://scholar.sun.ac.za

Based on these specifications, the model examines the probability that an observation will fall into either one of the three categories, using a maximum likelihood estimator specified as (Baulch, 1997a):

$$L = \prod_{t-1}^{T}[\lambda_1 f_t^1 + \lambda_2 f_t^2 + (1 - \lambda_1 - \lambda_2) f_t^3] \qquad (3.7)$$

Where $f_t^1,\ f_t^2,\ f_t^3$ takes functional forms for regimes 1, 2 and 3 respectively, and λ_1 and λ_2 are the probabilities for regimes 1 and 2 respectively (Baulch, 1997a). The higher the probability of being in regime 1 the greater the degree of market integration. Barret (1996) supports this procedure stating that the use of likelihood estimator effectively handles time varying transfer costs as well as periods of discontinuous trade.

The PBM, like any other influential model, has been a subject of modifications. Barret and Li (2002) extended PBM to accommodate both market integration and competitive market equilibrium analysis. Negassa and Myers (2007), and Zant, (2012) modified the model by relaxing the assumption of constant regime probabilities overtime.

The principle advantage of the PBM is that it measures integration by incorporating data on transaction costs. However, there are a number of limitations to the PBM approach. From a practical point of view, transaction cost data is rarely available and difficult to measure. Barret (1996) also argues that beyond transport costs there are several unobservable costs associated with doing business. Underestimated transaction costs may thus bias the PBM results to find integration. Secondly, PBM is inherently static, implying that the model tells us the probability of being in either of the regime but says nothing about the persistence of these deviations from equilibrium (Van Campenhout, 2007). Negassa and Myers (2007) further notes that the standard PBM assumes that regime probabilities are constant over time, basically implying that the degree of spatial efficiency between markets does not change overtime. But this may be problematic in practice if for some reason (such as policy change), the probabilities of falling within a trade regimes changes. Other limitations are that (1) it is a trade flow direction specific analysis: in the case of bi-directional trade, therefore, one estimation for each market direction is required, (2) it fails to account for lagged price adjustments, since it only employs contemporaneous price differences and (3) the procedure treats short run deviations from equilibrium as inefficiencies (Negassa *et al.*, 2003)

2.5.3.2 Markov Switching Model (MSM)

The use of MSM models in spatial market analysis is based on the assumption that the model recognizes regime generating processes of time series variables, which are governed by a Markov chain. This class of models is thus most useful in cases where the regime triggering (threshold) variable is nearly impossible to identify (Van Dijk *et al.*, 2002) or cannot be adequately described by use of threshold models (Psaradakis *et al.*, 2004). Hamilton (1995) cites abrupt changes in government policy (for example, the introduction or removal of a legal regulation) as potential sources of such a regime switch. The intuition behind MSM is that time series are subject to recurring and alternating regimes, that switch in a stochastic manner, from one to another with certain probabilities (Ihle, 2009). Proponents of this model thus believe

Stellenbosch University https://scholar.sun.ac.za

regime switching in spatial price behaviour, may be more explained by exogenous factors such as the prevailing state of the trading process or other random immeasurable shocks to the economy, which may influence the behaviour of economic agents, rather than the magnitude of disequilibrium (Ihle, 2009). This idea is central to the model application in market integration analysis.

The principal advantage of this class of models is that, unlike the PBM models, the factors (variables) inducing the regime switch do not have to be specified. This implies that regime switching determinants may not need to be measurable or observable and may not even be known *a priori* by the analyst. The model is thus attractive, given the challenges associated with time series data availability (Hamilton, 1994). Nevertheless, analysts still need to identify the plausible determinant, such as driving forces of trade and prices or other economic variables, that triggered the recurrent switches *a posteriori* (Ihle, 2009).

The origin of Markov models for switching regression is tied to the work of Goldfeld and Quant (1973), who considered parameter estimation in a switching regime pattern (Ihle, 2009). Lindgreen (1978) proposed a comprehensive approach to the analysis Markov models, following Baum's *et al.* (1970) idea (Ihle, 2009). Most recent contributions are, however, associated with the work of Hamilton (1988, 1989), after his extension of the model to time series analysis. A detailed treatment of the model was then proposed by Krolzig (1997) while Pelagatti (2005, 2008, cited in Ihle, 2009) proposed a Markov Switching Vector Autoregressive Model (MSVAM) with time dependent probability transitions. Variants of this model have since been developed and applied in empirical studies with a recent contribution from Zhao *et al*. (2012), who proposed a Markov Switching Error Correction Model (MSECM) with time varying regime transition probabilities.

The standard MSM is flexible[4] and is suitable for both stationary series (when applied as autoregressive models) and non-stationary series (when applied as Vector Error Correction Models (VECM)). In a VECM format, commonly applied in price dynamic analysis, the model is generally formulated as (Ihle, 2009):

$$P_t = v^{s(t)} + \alpha^{(s_t)}\beta^1 P_{t-1} + \sum_{i=1}^{k-1} \Gamma_i^{s(t)} P_{t-i} + \varepsilon_t \qquad (3.8)$$

Where *s(t)* is a regime variable, with a value between 1 and m regimes and is allowed to be unobservable. Parameters v (constant), α (cointegrating vector) and Γ (short run price adjustment) are thus regime dependent.

Although this approach is an improvement to spatial market integration analysis, especially because it can incorporate threshold variables that are completely impossible to measure, it is more commonly applied in financial analysis and business cycle and only rarely applied in price dynamic analysis (Ihle, 2009). The principle limitation of this method is that regime selection requires a thorough and critical analysis. This is because such regimes are exclusively selected from the data sample. It is up to the analyst to make sense of the regime, since there is no immediate economic theory upon which to base the decision (Ihle, 2009). Moreover, the

[4] Refer to Krolzig (1997) and Ihle (2009) for a detailed mathematical formulation

Stellenbosch University https://scholar.sun.ac.za

fact that regimes are identified based on careful scrutiny of economic and other factors prevailing during the data period, leaves the analysis to a far more confirmatory rather than exploratory procedure.

2.5.3.3 Threshold Autoregressive Models (TAR)

Threshold Autoregressive model (TAR) is another technique used to test spatial market dynamics within the framework of regime switching. The use of this model is based on the theoretical idea that time series might exhibit measurable differences in their dynamic behaviour, depending on the size of a particular variable, referred to as a *threshold*. A key modelling characteristic, therefore, is to identify, a priori, the threshold variable (variable that triggers the regime switch). Whenever this variable crosses a critical value, which is known as a *threshold value*, a regime switch occurs. Model parameters are therefore allowed to vary across regimes but must remain constant within each regime. This approach differs from the MSM in that the threshold variable must be quantifiable and known *a priori*.

Price transmission literature within the framework of traditional TAR models uses a fixed but often unobserved transaction cost as a threshold variable (Fackler & Goodwin, 2001). Intermarket price deviations must thus exceed these threshold bands before provoking equilibrating price adjustments that in turn results into market integration (Goodwin & Piggot, 2001). Goodwin and Piggot (2001) portray threshold effects to occur when markets exhibit variations in the way they respond to shocks below and those above a critical threshold.

The basic premise behind TAR models is attributed to Howell Tong the “forefather of threshold ideas” (Goodwin & Piggot, 1999; Mann, 2012). From his series of publications on nonlinear time series models, Tsay (1989) developed a technique for testing threshold effects and modelling TAR processes. Balke and Fomby (1997) extended the threshold model to cointegration framework while Goodwin and Hot (1999) proposed a Threshold Vector Autoregressive (TVARM) model.

Variants of TAR models have since been developed and applied in several empirical studies including, Enders and Siklos (2001), Hansen (2002), Serra *et al.* (2005) and Balcombe *et al.* (2007). The Enders and Siklos (2001) approach in particular is a single threshold model estimate procedure applied by adding a Heaviside indicator function (I_t) directly into the Engle Granger (1987) residual regression equation estimated as (Mann, 2012):

$$\Delta\varepsilon_t = \rho_1 I_t \check{\varepsilon}_{t-1} + \rho_2(1 - I_t)\varepsilon_{t-1} + \mu_t \qquad (3.19)$$

where I_t is the Heaviside indicator function given as:

$$I_t = \begin{cases} 1 & if|q_{t-1}| > \tau \\ 0 & if|q_{t-1}| \leq \tau \end{cases}$$

and ε_t is the mean zero residuals from the cointegrating equation, μ_t is the constant. Nevertheless, other studies have employed a multiple threshold, multiple regime modelling approach.

The principal advantage of TAR models is that they account for non-linearity in price adjustment while testing the speed with which markets restore arbitrage conditions once violated (Abdulai, 2007; Fackler & Goodwin,201). Van Campenhout (2007) compares the TAR and PBM models and argues convincingly for the superiority of TAR models as a measure of market integration.

Nevertheless, TAR models in their traditional setting have been criticized on many grounds. Myers and Jayne (2012), and Burke and Myers (2014) provide four of the major criticisms. Firstly, nearly all applications (see for instance Abdulai, 2000; Goodwin & Piggot, 2001) impose an implicit assumption that transaction costs remain constant overtime. Since transfer cost data is rarely available, analysts believe it is better to assume constant transfer costs than to ignore them altogether. However, easily observable time varying factors driving transfer costs (such as fuel prices) may provide a better proxy of transfer costs. Secondly, traditional TAR models are measured with price data alone. If trade flow data is available and explicitly included in the model, however, they may provide a better insight into the spatial market relationship. This is also in line with González-Rivera and Helfand (2001a) who argue that trade levels influence price transmission. Thirdly, traditional TAR models assume that the equilibrium relationship between prices remain the same with and without trade. It is possible however, that the relationship may vary with variations in trade levels. Finally, when transfer costs are above intermarket price differences, theory suggests a lack of arbitrage opportunity. In contrast, traditional TAR models allow for a long run equilibrium during this period, which could result in considerable bias in parameter estimation.

Stephens *et al.* (2012) attempted to address some of these issues by modifying the TAR model in a way that transmission mechanisms could vary between periods of no trade and periods with trade, but their model can only handle two regimes (one threshold). Myers and Jayne (2012) however address many of these challenges by modifying TAR models to allow for multiple trade regimes (multiple thresholds). The authors consider a spatial market equilibrium relationship given as (Burke and Myers, 2014):

$$p_t^A = \beta_{0i} + \beta_{1i} p_t^B + \beta_{2i} k_t + u_{it} \qquad (3.10)$$

In this equation, p_t^A is the commodity price in the destination market (A), p_t^B is the commodity price for the exporting market (B), β_{0i} is the constant, β_{1i} is the long run equilibrium relationship between bean prices in the two markets A and B, k_t is the transfer cost (diesel price), β_{2i} is the long run relationship between bean prices in the importing market (A) and the transfer cost and u_{it} is the error term. Markets A and B will attain perfect spatial arbitrage condition when, (1) β_1 equate to one (β_1=1), (2) β_2 equate to one (β_2=1) and (3) the constant (β_0 =0) (Burke, 2012). Myers and Jayne (2012) then extend the equation above into a Single Equation Error Correction Model (SEECM) framework which Burke (2012) presents as:

$$\Delta p_t^A = \mu_i + \beta_{1i}\Delta p_t^B + \beta_{2i}\Delta k_t + \lambda_i(p_{t-1}^A - \beta_{1i}p_{t-1}^B - \beta_{2i}k_{t-1}) + \theta_{1i}p_{t-1}^B + \theta_{2i}k_{t-1} + \sum_{j=1}^{n} b_{ji}\left(\Delta p_{t-j}^A - \beta_{1i}\Delta p_{t-j}^B - \beta_{2i}\Delta k_{t-j}\right) + \rho_{1i}\,\Delta p_t^B + \sum_{j=1}^{n} c_{ji}\left(\Delta p_{t-j}^B\right) + \rho_{2i}\Delta k_t + \sum_{j=1}^{n} d_{ji}\,\Delta k_{t-j} + u_{it} \qquad (3.11)$$

Stellenbosch University https://scholar.sun.ac.za

in order to allow for the estimation of the speed of price transmission (λ). All other variables remain as defined in equation 3.10, but θ_1, θ_2, b_i, d_i, ρ_1, and ρ_2 are parameters to be estimated. The model is flexible and can take various forms, depending on the stochastic properties of the underlying data (refer to section 4.5.1 for details on the modelling assumption).

Myers and Jayne (2012) tested this framework on maize markets between Zambia and South Africa. Other applications are found in Ndibongo *et al*. (2010), Burke (2012) and Burke and Myers (2014). Given its several strengths, this study therefore adopted the Myers and Jayne (2012) variant of the TAR model to measure market integration between Zambia and DRC, and Zambia and Tanzania. It is important to however note that this approach is both data (requires price, trade volume and transfer cost data) and computationally burdening, since it requires a separate procedure to identify multiple threshold values. Moreover, as in PBM, transfer cost data is rarely available. The explicit use of price, trade volume and transfer cost data is thus expected to provide a conclusion from this study that is nearer to the truth as regards cross border bean market integration between Zambia and her selected neighbours.

2.6 Methods used to select the number and location of thresholds

Overall, regime switching techniques and TAR models, in particular, improve the measurement of spatial market integration by recognizing the dynamic responses arising from threshold effects and they capture non-linearity. However, economic theory in most cases is not specific about the full structure of these models. In TAR models, particularly, the theory rarely offers an intuitive rationale for *a priori* imposition of the number and location of thresholds (Gonzalo & Pitarakis, 2002; Chen *et al.*, 2012).

The main question analysts then ask when performing econometric analysis involving TAR models is whether the non-linear specification is superior to a linear model (Hansen, 1999). The answer lies in whether the hypothesis of a significant threshold could be rejected or not. TAR modelling thus begins by identifying the threshold variable (q) (variable that triggers the regime switch), locating the threshold values (μ) (critical value at which model parameters are allowed to change) and determine the optimal number of regimes (n). The threshold variable (q) is selected based on economic theory while the threshold values (μ) and the optimal number of regimes (n) require empirical procedures. Figure 2.2 presents a hierarchy of methods frequently employed for this task. Each of these methods is discussed below:

Stellenbosch University https://scholar.sun.ac.za

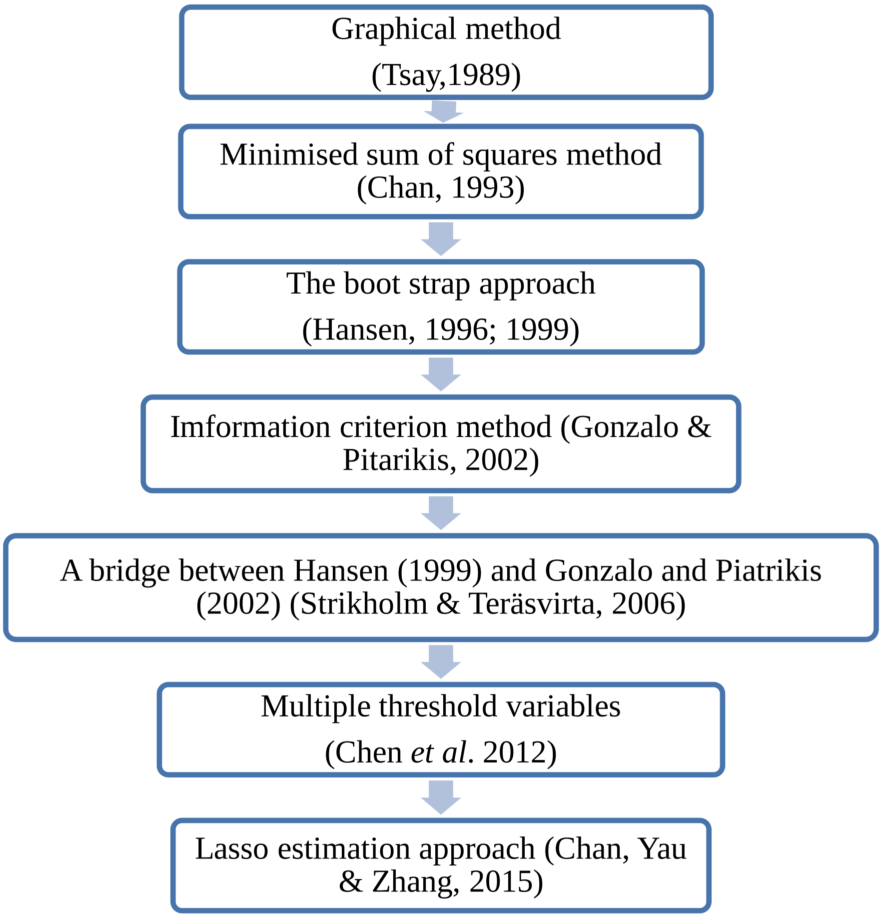


Figure 2.2: Common methods used to locate and select the number of thresholds
Source: Summarised from Mann (2012)

***Graphical method**:* Tsay (1989) proposed a graphical approach to locating the value (v) and the number of thresholds (n). His method is built on Tong and Lim (1980) who pre-specified a set of finite sample profiles and used percentiles to estimate threshold values (μ). However, the method by Tsay (1989) is more direct. The procedure combines linear regression analysis and a scatter plot of several statistics versus the threshold variable (q) to estimate the location (μ) and number (n) of thresholds. To do this, recursive least square estimates are firstly computed based on arranged autoregression size of the threshold variable (q) (Mann, 2012). This is followed by a careful examination of three scatter plots; (1) the standardized predictive residuals versus the threshold variable (q), (2) the t statistics, obtained from the recursive estimates of an arranged autoregressive coefficient, versus the threshold variable(q) and (3) the ordinary predictive residuals versus the threshold variable (q) (Mann, 2012). The break points on the plots represents a possible threshold location (μ) at which point TAR model parameters change

Tsay (1989) argues that these plots are sufficient to directly reveal the location of the threshold value (μ). The scatter plots can also detect the absence of a threshold ($\mu = 0$) in the TAR model. For instance, in a TAR model without thresholds ($\mu = 0$), graph 2 is expected to converge gradually and smoothly (Mann, 2012). If on the other hand at least one significant threshold ($\mu = 1$), graph 2 will gradually and smoothly converge but deviates from

Stellenbosch University https://scholar.sun.ac.za

convergence as soon as (and every time) a threshold (μ) is detected. If there is more than one threshold, it is expected that graph 2 will converge and deviate every time a threshold (μ) is detected.

The primary advantage of this method is that it can detect more than one threshold (n) and their locations (μ). However, this procedure fails to detect a threshold located at extreme percentiles (e.g. 0^{th} or 100^{th} for the dataset involving 0 to 100 values) (Tsay, 1989). Moreover, the precise location of threshold values is subject to researchers' discretion.

Minimised sum of squares method: Chan (1993) proposed an alternative to Tsay's (1989) method. Instead of graphs, this method uses the minimized sum of squared residuals to locate the threshold value (μ). It proceeds by discarding the smallest and largest 10%-15% of the possible threshold values respectively and then searches for a threshold within the middle 80% of the possible threshold values (Mann, 2012). The principle advantage of this approach is that it is more precise and less subjected to threshold selection error. Nevertheless, it is only useful when considering a single threshold (μ) TAR model, since the approach cannot detect multiple (more than one) threshold values (μ).

***The Boot strap approach*:** The method by Hansen (1999) uses a bootstrap procedure to test for thresholds. Unlike Tsay (1989) and Chan (1993), this method aims at identifying the number of thresholds (n) *i.e.* regimes, as opposed to locating the actual values (μ) (Mann, 2012). Given a sample of x observations, Hansen's (1999) test uses linear regression in a sequential threshold estimation procedure, to select the number of regimes (n). Step 1 tests the null hypothesis of a linear model (n=0) against the alternative hypothesis of two regime model (n=1). If the null hypothesis is rejected, the procedure is repeated to test for 3 regimes (n=2) model, which equals 2 thresholds (μ). The procedure continues with an addition of a potential threshold (n) in every subsequent test until the first rejection of the null hypothesis of ($n+1$) regimes. The decision of whether a threshold is significant or not is based on an F-statistic. However, since the distribution of the F-statistic is non-standard due to problems associated with nuisance parameters (Mann, 2012), Hansen (1999) employs the Hansen (1996) bootstrap procedure to determine the significance of the F test. The final decision therefore is based on the p value of the F-statistic.

The principal advantage of this approach is that it allows for the estimation of asymptotically valid p-values for the test of single versus two regime model (Gonzalo & Pitarakis, 2002). It can also handle general TAR models of any order although its validity is primarily established on a single threshold (two regime) model. Nevertheless, the ability of Hansen's (1999) test to detect multiple regimes (n>1) is often brought into question since there is no distribution theory to support the bootstrap-based procedure (Gonzalo & Pitarakis, 2002). Additionally, after comparing this method with the method by Gonzalo and Pitarakis' (2002), Strikholm and Teräsvirta (2005) concludes that the method by Hansen (1999) may not work properly when more than 2 regimes (n>1) (multiple regime) are considered.

Information criterion approach: To solve some of these limitations, Gonzalo and Pitarakis (2002), here after GP, proposed a model selection approach that utilizes information criterion

Stellenbosch University https://scholar.sun.ac.za

functions to select threshold values (μ). This procedure can be implemented either in a sequential or non-sequential approach. Basically, the procedure begins by deciding between no threshold (n=0) (null hypothesis) and one threshold (n=1) by optimizing the objective function given as (Burke, 2012):

$$Q_T(n) = \max_{\tau} \ln \left[\frac{SS_T}{SS_T(\tau)}\right] - \frac{g(T)}{T} Kn \quad (3.15)$$

where $Q_T(n)$is the model with n thresholds, K is the number of parameters for estimation in a given single regime model, SS_T and $SS_T(\tau)$ are the residual sum of squares for a model without thresholds (single regime) and a model with n thresholds respectively. If $Q_T(n=0) > Q_T(n=1)$, the null of no threshold is not rejected, so that the procedure stops and one concludes that a linear model is optimal (Gonzalo &Pitarakis, 2002). However, if $Q_T(n=1) > Q_T(n=0)$, the null hypothesis is rejected, one obtains the parameter estimates and proceeds to step 2. In step 2 the procedure decides between no threshold (n=0) (null hypothesis) and one threshold (n=1) in the subsamples (observations above and below the threshold value (μ). The procedure continues until the grid search in subsequent subsamples cannot detect any threshold. The GP procedure, as in Chan (1993), imposes a lower bound of about 10% to 15% of the minimum observations in each regime in order to allow sufficient model parameter identification.

A unique feature of the GP approach is that the decision between n=0 and n=1 thresholds is based on the penalty term instead of test statistics**.** The penalty term is independent of the threshold location (μ), and prevents model over-parameterization as the number of thresholds (n) increase (Mann, 2012). Hypothesis testing is performed using a combination of Akaike information criteria (AIC) ($\lambda_T = 2$), Bayesian information criteria (BIC) ($\lambda_T = \log T$), (BIC2) ($\lambda_T = 2logT$) and (BIC3) ($\lambda_T = 3logT$) and Hannah-Quinone criteria (HQ) ($\lambda_T = 2logT$). Nevertheless, Gonzalo and Pitarakis (2002) demonstrated the superiority of BIC in finite samples.

The main advantage of this method and therefore the rationale for its choice in locating trade based thresholds in this study is that it is able to jointly detect multiple thresholds (n) and their specific values (μ). Moreover, the fact that it depends on a penalty term instead of test statistics helps avoid complications resulting from nuisance parameters (Mann, 2012). Also, in simulation tests presented in Strikholm and Teräsvirta (2006), the GP approach compared to Hansen (1999), have the best performance by far when more than 2 thresholds are considered. On a negative note, however, the implied significance level may vary considerably between higher (n+1 thresholds) and lower models (n), since GP cannot adequately control asymptotic significance levels (Strikholm & Teräsvirta, 2005)

A bridge between GP procedure and Hansen (1999)*:* Strikholm and Teräsvirta (2006) then developed a method that bridges between Hansen (1999) and the GP approach following a careful simulation study of the two methods. Their method is similar to Hansen (1999) in that it is based on test statistics and focuses primarily on determining the number of thresholds (n). At the same time, the procedure follows a sequential testing framework proposed by GP. Strikholm and Teräsvirta (2006) suggest that their method can be seen as an alternative to

Stellenbosch University https://scholar.sun.ac.za

Hansen's (1999). The first step in this approach tests the null of *n*=0 against the alternative of *n*=1 thresholds using F statistics. The number of thresholds (*n*) is determined by repeating this procedure for subsequent thresholds (μ) until one fails to reject the null hypothesis.

Strikholm and Teräsvirta (2005) conducted simulations to compare their method with the GP and Hansen (1999) methods. The authors concluded that the principle advantage of their procedure over Hansen (1999) is that it is computationally simpler and avoids bootstrap based distribution problems that occurs when the null hypothesis model is incorrectly specified. With respect to the GP approach, the authors showed that the advantage of their procedure is that it has better size properties and overcomes challenges associated with optimal model selection criteria. Despite these strengths, the correct decision frequency in Strikholm and Teräsvirta (2005) is, in most cases, lower than the GP approach (Mann, 2012).

Multiple threshold variable: The method by Chen *et al.* (2012) is quite distinct from all of the above. Instead of one threshold variable (q), the procedure uses the likelihood ratio tests to determine the number of regimes (n) and the threshold values (μ_1 and μ_1) in a TAR model with *two* threshold variables ($q_1$and q_2). A case of two threshold variables is justified when two variables simultaneously determine the dynamic behavior of time series data. To determine the number of regimes, Chen *et al.*, (2012) follow a sequential procedure. The first step considers the null of only one regime (no thresholds) versus the alternative of more than one regime. The null hypothesis is rejected if the bootstrapped p-values, following Hansen (1999), are significantly small. In the second step, Chen *et al.* (2012) tests the hypothesis of a three versus four regime model. The six hypothesis tests are examined in sets of; $q_1 \leq \mu_1$ and $q_2 \leq \mu_1$; $q_1 \leq \mu_1$ and $q_1 > \mu_2$; $q_1 > \mu_1$ and $q_2 \leq \mu_2$; $q_1 > \mu_1$ and $q_2 > \mu_2$. This is the highest number of regimes allowed. However, the procedure can also test an alternative hypothesis of three versus a two regime model.

The primary advantage of this method is that it can handle cases of more than one threshold variable ($q_1$and q_2) in a multiple threshold framework. Nonetheless, the method is computationally burdening and does not allow cases were the number of threshold are unknown.

LASSO estimation approach: Recently, Chan *et al.* (2015) proposed a two-step procedure to estimate the number and location of threshold values (μ). The first step obtains a set of potential thresholds by a least absolute shrinkage and selection operator (LASSO) estimation procedure from a specified maximum number of thresholds *(n)*. "The second step employs a model selection procedure that selects consistency estimates of the thresholds from the set of potential thresholds" Chan *et al.* (2015).

The advantage with this method is that it can detect multiple threshold values (μ). In comparison to the GP approach, Chan *et al.* (2015) conclude that their method gives similar results. However, they note that the two-step procedure performs very well in large samples (above 600), while in small samples the bias and standard deviation of model estimates is high.

Stellenbosch University https://scholar.sun.ac.za

2.7 Empirical evidence of cross border market integration in agricultural commodities

As the number of empirical methods have proliferated, so have the studies examining spatial food market behaviour. These studies occupy a voluminous literature in agricultural economics. Van Campenhout, (2007) attributes this to the growing recognition of the importance of market integration analysis as well as price data availability. Most of all, these studies have been central to contemporary debates concerning domestic market liberalization, price stabilization policies (Baulch, 1994) as well as the domestic impact of global food price crisis of 2007/2008. In Africa, empirical evidence is vast. However, little analysis has been done on cross border pulse market integration, as most research examines cereal markets and at intra country level (see for example, Goletti & Babu, 1994; Muyatwa, 2000; Cirera & Arndt, 2008; Van Campenhout, 2007; Myers, 2013).

To the best of my knowledge, only Korir *et al.* (2003) and Mauyo *et al.* (2007) have attempted to study bean market integration at inter-country level in Africa. Korir *et al.* (2003) employed the static price correlation method to examine markets, connected by formal trade between Tanzania and Kenya. Conducting the study on a mare 24 monthly price observations (2000-2001), the authors found very weak integration between selected markets. Korir *et al.* (2003) attributed this to trade restrictions imposed by the two countries in the form of export and import levies as well as the bureaucratic procedures in their bean international trade. The study therefore recommended free exports and importation of beans across the two countries to improve cross country market integration. Mauyo *et al.* (2007), on the contrary, found high levels of integration between selected wholesale bean markets connecting Kenya and Uganda, using a similar number of price observations (2001-2002) and analytical methods as in Korrir *et al.* (2003). The study similarly suggested a focus on eliminating trade obstacles between the two countries. These included road infrastructure development, tariff and other institutional barriers.

Bean markets have also been a subject of investigation at intra-country level. Nevertheless, evidence is also mixed. Loveridge (1991) investigated the impact of road infrastructure development on bean market integration in Rwanda by examining price differentials between rural and urban markets. The results showed an increased strength of market linkages in the post road paving period. However, it was revealed that the cost of moving beans between rural and urban markets was still high. The study thus recommended investment in the transport sector to reduce these costs. Odiambo *et al.* (2006) employed co-integration and impulse response methods to test for bean market integration in Nairobi, Kenya. Their analysis established that the Nairobi bean markets were not fully integrated, attributing it to inefficient and poor access to market information. The authors thus proposed government intervention in information provision and infrastructural development to improve market linkages. Mayaka (2013) expanded this study out of Nairobi by investigating rural-urban bean market integration across Nairobi, Nakuru, Eldoret and Kitale markets for the period 1994-2011. His analysis using co-integration, contradicted Odiambo's *et al.* (2006) findings and instead found that all markets under study were co-integrated. The traditional TAR model results further established

Stellenbosch University https://scholar.sun.ac.za

that 3 weeks were required for prices in Nairobi to be transmitted to Kitale. Meanwhile, Mtumbuka *et al.* (2014) in Malawi concluded that the domestic bean markets were not fully integrated with each other, when they employed a TAR model to analyse regional market integration. The study also showed that distance and road quality connecting markets had a significant influence on observed results.

Zooming out of the studies on beans, but sticking with the southern and eastern African corridors, Teka and Azeze (2002) analysed integration in their study on livestock cross border trade between Ethiopia and her neighbours Djibouti and Somali. The authors employed correlation coefficients and the Ravallion model. They found that markets in the eastern borderland (Jijiga area) are weakly integrated with their Somaliland counterparts. Earlier on, Teka *et al.* (1999) had found no spatial integration between livestock markets linking Ethiopia and Kenya when they employed correlation coefficients on a three-year data period (1997-1999).

Mutambatsere *et al.* (2006) also investigated regional spatial integration and efficiency among five central maize markets in southern Africa. Their study covered Gaborone in Botswana, Gauteng in South Africa, Blantyre in Malawi, and Maputo and Mocuba in Mozambique. They employed a combination of correlation coefficient, Granger causality, cointegration, PBM and Barret and Li's (2001) extension of the PBM over a data period of 1994-2002. Results even between similar market pairs varied notably across these methods and the study did not state the preferred method in case of such contradictions, creating some confusion in the overall conclusion. The authors, however, concluded that overall the markets included exhibited significant levels of integration. In particular, higher levels of integration were observed between Botswana and South Africa, Malawi and Mozambique and lower degrees between Mozambique and South Africa. Botswana and Mozambique were completely segmented (unintegrated) but highly efficient.

Another study on maize markets was conducted by Ihle *et al.* (2010). The aim was to assess the magnitude and the effect of the border on maize price transmission across Uganda, Kenya and Tanzanian. Although these countries are all members of East African Community (EAC), they pursue varying agricultural trade policies. The results, similar to Mutambatsere *et al.* (2006), showed variations in the degree of integration across country pairs. Overall, however, they revealed that it mattered which border was crossed. Kenya and Uganda were more integrated with each other while Tanzania appeared to be rather isolated from outside price changes. The study further revealed a strongly negative effect on market integration from the Kenya-Tanzanian border. The study however suggested further research to explain such findings.

Meanwhile, Ndibongo *et al.* (2010) employed a multiple trade based threshold TAR model to assess market integration in grain maize markets linking Mozambique (Maputo) and South Africa (SAFEX). The method allowed for the speed of adjustment and price transmission to vary across trade levels as plausibly as possible. The authors combined monthly prices, diesel cost and formal trade flow data. They found that regardless of the model assumption made, there was no statistically significant long run relationship between Maputo maize grain prices

Stellenbosch University https://scholar.sun.ac.za

and the South African SAFEX prices. The outcome of the study thus meant that price deviations between the two countries would be expected to grow, with no tendency towards equilibrium. The authors cited a lack of grain market demand, prohibitive tariffs, existence of market power, poor inventory management and asymmetric information as potential reasons for their observation. Nevertheless, some evidence of a long run relationship was observed between South African grain maize market and Mozambique retail maize meal markets at high imports only. It took Mozambique 10.3 months to fully adjust to grain maize prices in South Africa, clearly indicating very weak integration.

Acosta (2012) revisited the Ndibongo *et al.* (2010) study by employing an asymmetric error correction model on white grain maize for the period 2000 to 2011. This study however revealed that although no short run relationship existed, Mozambique and South Africa were weakly integrated in the long run. The authors found that it took between 7 months (response to price increases) and 11 months (response to price decreases) for Mozambique to fully adjust to price changes in South Africa. The study justified the weak integration on the prohibitive tariffs imposed on maize trade between the two countries.

Amikuzuno and Donko (2012) analysed the border effects on tomato markets between Ghana (deficit market) and Burkina Faso (surplus market). The regime switching VECM, on semi-weekly data found that, prices largely co-move although producer markets exhibited stronger and considerable levels of response to spatial price variations than did the consumer markets. The study demonstrated that the existence of the border weakened transmission of prices in markets connected by informal trade. The authors could not however unravel the extent to which other factors such as imperfect market information and transaction costs could explain the observed results.

Myers and Jayne (2012) investigated maize markets between Zambia and South Africa in their quest to test their modified TAR model to allow for multiple trade based thresholds. The approach combined maize prices, formal trade flow data and transaction costs and assessed the effect of government influence on international price transmission. Their results showed that the two markets were integrated. However, it took Zambia 1.2 months (strong integration) during the low import regime and 7.8 months (weak integration) during the high import regime, to fully respond to maize price changes in South Africa. These two periods coincided with periods when government did not participate in importing maize and when government participated. The study therefore underscored the negative influence of government intervention on price transmission and the fact that transmission may be likely to be sensitive to trade volume.

Burke and Myers (2014) reapplied the Myers and Jayne (2012) approach on cross border maize grain markets dominantly connected by informal trade between Zambia and DRC, and Malawi and Mozambique. After exhaustively searching the data for trade based thresholds, the authors concluded that there was no evidence of trade based threshold effects. Despite this, however, the authors found evidence of long run price equilibrium relationships in both market pairs with rapid price transmission. Estimates of the speed of price transmission indicated that it took 2.5 months for Liwonde to fully respond to Cuamba prices and 2.7 months for Kasumbalesa

Stellenbosch University https://scholar.sun.ac.za

(DRC) to respond to Kitwe (Zambia) price changes. Informal trade constraints between markets had very little disruptive effect on long run price relationships. Nevertheless, Burke and Myers (2014) found occasionally higher levels of transfer costs sufficient to impede trade flows. The authors conclude that a policy focus on informal trade and reduction in their trade costs would result in improved market performance. This conclusion is in line with Korri *et al.* (2003), Mauyo *et al.* (2007), Myers and Jayne (2012) and Acosta (2012).

Cross border market integration has also been studied beyond ESA. Serra *et al.* (2006) found that cross country price transmission was dependent on the distance between markets. The authors assessed leading pork producers and traders (Germany, Spain, France and Denmark) in the European Union using both parametric tests (local linear regression estimator) and the non-parametric TAR model. The results indicated evidence of market interrelationships. However, Spain and Germany, with a relatively longer distance between, them showed symmetric price transmission. Recently, Akhter (2016) investigated rice market integration between the surplus market in India and deficit markets in Bangladesh and Nepal, to determine the impact of government intervention following the food price crisis of 2007/2008. Co-integration results showed that both market pairs (India and Bangladesh, and India and Nepal) were integrated. The authors attributed this surprising finding to widespread informal cross border trade between these markets.

In summary, empirical evidence on cross country market integration is mixed and biased towards maize. Nearly all studies examining markets dominantly connected by formal trade and weakly integrated at least attributed their results to policy intervention as a possible determinant. However, the limited empirical evidence on informal trading markets is also mixed, suggesting that the presence of the border can still weaken price transmission between free trading countries. The contributions of this study to this literature are threefold. The first is that it examines dry beans cross border markets dominated by informal trade as opposed to formal trade. The second is that it employs a better measure of price transmission, multiple regime TAR model, which combines trade, price and transfer cost data as opposed to the static price correlation approach used in the only evidence on cross border beans studies. The third is that it provides evidence on the relative importance of trade volume on market integration.

2.8 Conclusion

This chapter focused on providing an in-depth overview of the theoretical and empirical literature on the subject of spatial market integration. However, there is no general agreement in the literature as to what market integration is, although consensus exists that it is founded on the law of one price. It is also apparent that market integration is influenced by a host of factors including transaction costs, government policies, market information and market infrastructure.

A review of methods used to measure market integration showed that integration analysis has evolved over time from simple static models (that use price data only and classify markets as strictly integrated or segmented over the entire study period) to advanced models (that can combine price, trade volume and transfer costs data and recognize that markets could switch between higher and lower degrees of integration or even segmentation within a given period, a

Stellenbosch University https://scholar.sun.ac.za

more realistic case). Nevertheless, data availability still limits the use of advanced models. The section however argues that when available, advanced models provide better inference on market integration. The study therefore adopted the threshold autoregressive model in data analysis. The next section provides a background to the bean industry in eastern and southern Africa as a focus of this study.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 3: OVERVIEW OF THE EASTERN AND SOUTHERN AFRICAN DRY BEAN INDUSTRY

3.1 Introduction

The common beans (*Phaseolus vulgaris L.*) represents one of Eastern and Southern Africa's (ESA) staple agricultural crops and a popular crop throughout the world. Its cultivation has a long history. However, Africa only learnt of beans around the 16th Century. Since then, the crop has been grown widely across the continent and beyond (Graham & Ranalli, 1997; Gepts, 1998) existing as the world's most important pulse crop, in both production and consumption terms (Broughton *et al*., 2003). They are particularly important from a nutrition, economic and food security perspectives.

The purpose of this chapter is to provide an in depth overview of the common bean industry in order to lay a background to the assessment of its market functioning across Zambia, Tanzania and DRC. Before this is done, however, the chapter takes a look at the global industry in order to situate the local dry bean industries. The rest of the chapter is organised as follows, section 3.3 describes various aspects of the industry at ESA level with a major focus on regional trade. Sections 3.4, 3.5 and 3.6 discuss the bean markets in Zambia, Tanzania and DRC respectively. Trade between Zambia, DRC and Tanzania is discussed in section 3.7. Finally, section 3.8 focuses on various aspect relating to Informal Cross Border Trade (ICBT) as it reflects the single most important form of bean trade across borders of ESA countries. This section particularly highlights such issues as definitions, drivers and constraints.

3.2 Global overview

3.2.1 History and origin

Common beans are one of the oldest crops known to man. Although its history is a subject of debate, some literature records bean cultivation in the Nile valley as early as 2000 BC (Siddiq & Uebersax, 2013). Beans played a role in ancient people's politics and superstitions, with Greek magistrates elected by casting its seeds into the helmet (Siddiq & Uebersax, 2013). It was later discovered as a fine staple food when it provided a subsistent diet to early European settlers, first in America then generations later in the Great Lakes region, who grew it for their consumption and sold excess to their neighbours. Since then, the crop has been grown in nearly every part of the world, except Antarctica (Graham & Ranalli, 1997; Gepts, 1998), existing as the most important food legume crop globally (Beebe *et al.*, 2000; Broughton *et al*., 2003). It provides an important source of dietary nutrients and food especially among the global poor.

The ancestral origin of beans lies in the Andes, a region between northern Peru and Ecuador (Kami *et al.*, 1995). This empirical discovery is, however, widely disputed in recent literature, that now places beans' origin in the Americas central Mexico (Gepts, 1998; Beebe *et al.*, 2000; Bitocchi *et al.*, 2012). The *phaseolus vulgaris L* species then moved up north and south into what is known today as Argentina and northern Mexico (Gepts, 1998). It further spread widely

Stellenbosch University https://scholar.sun.ac.za

and became established as a food crop in Africa long before the colonial era (Wortmann *et al*., 1998; Beebe *et al.*, 2000). Historians particularly believe beans were initiated into Africa by Portuguese explorers in around 1500 and was first cultivated in Tanzania, Mozambique and Kenya of today (Wortmann *et al*., 1998). To date, these countries remain Africa's key sources of beans. Beans are a short season crop, maturing within 3 to 4 months. They are extremely diverse in terms of environmental adaptability, cultivation methods, growth characteristics and utilization.

3.2.2 Global production of dry beans

Worldwide, dry bean[5] production has increased over the last two decades. It also dominates the pulse industry at global level (Figure 3.1). Figure 3.2 illustrates how the world output of dry beans has increased gradually over the last 21 years, from a total of 17.4 million tons in 1994 to 25.1 million tons in 2014. This represents a 44% increase in production. Output growth increased by 4.5% from 1995 to 2004 but picked significantly (32% increase) from 2005-2014 (FAO, 2016). The upward trend in production is, to a large extent, driven by increases in area planted to beans, which increased by 1.4 million hectares and 3.3 million hectares over the two periods respectively (Figure 3.2). However, a yearly trend (Figure 3.2) seems to suggest improved yields could have also been a factor as noted by Akibode and Maredia (2011).

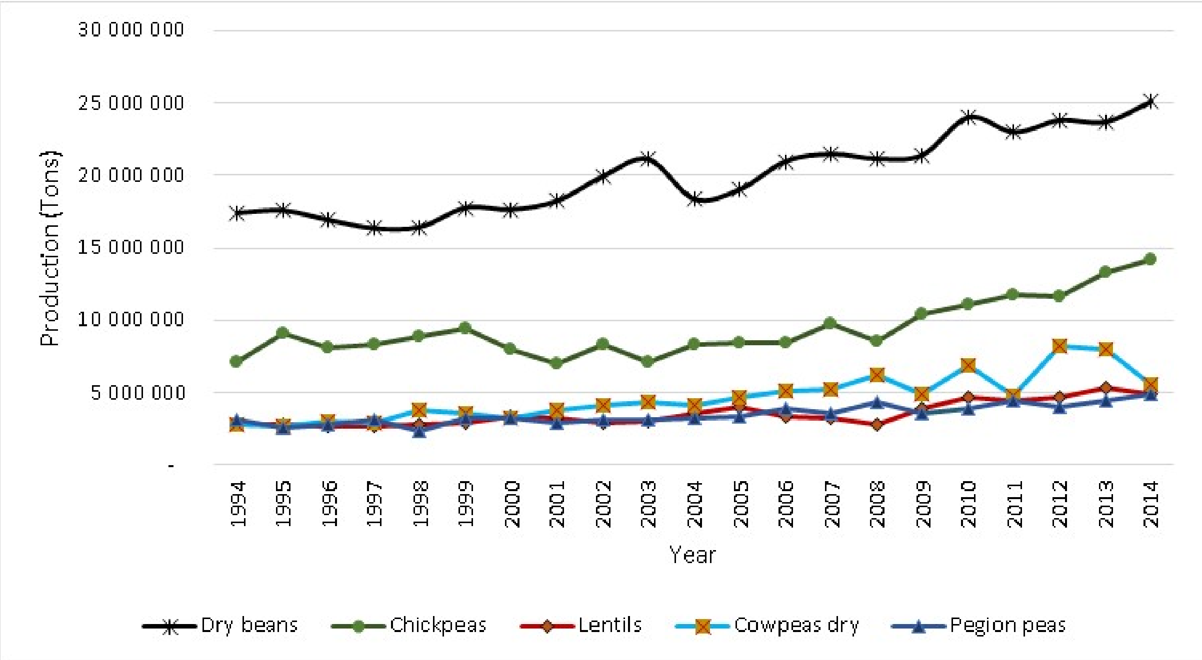


Figure 3.1: Global pulse production by pulse type (1994-2014)
Source: FAO (2016)

[5] FAO data has a number of weaknesses including its general classification of dry beans, country classification into regions, missing data amongst others. A detailed discussion on this is provided in Akibode & Maredia (2011).

Stellenbosch University https://scholar.sun.ac.za

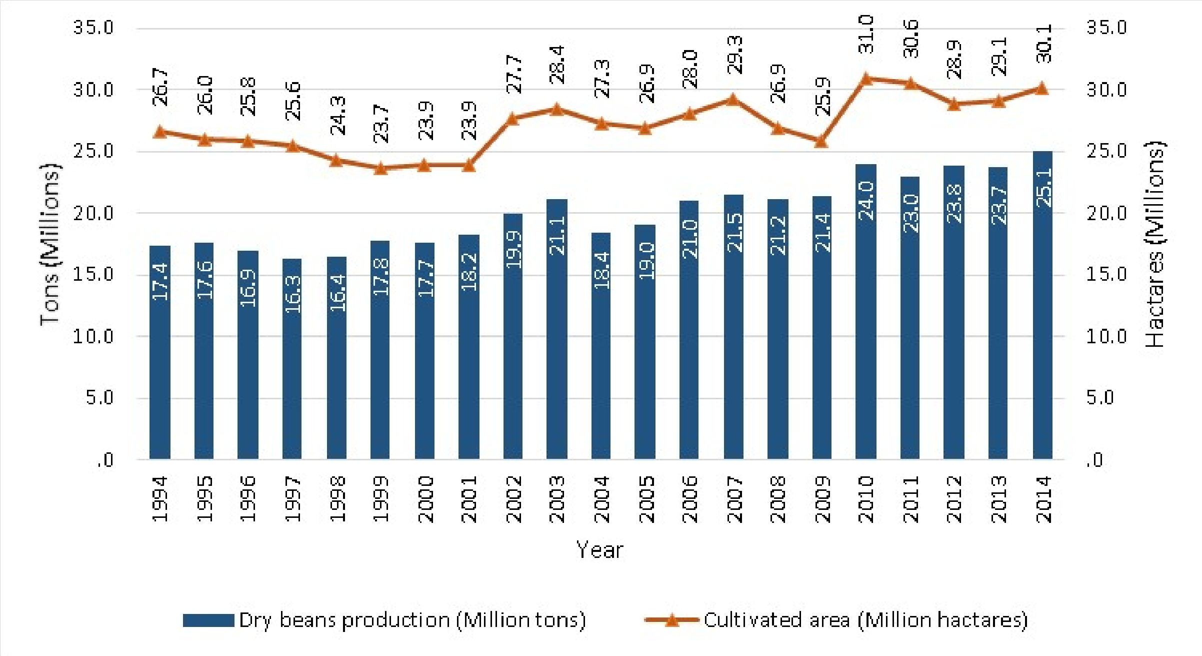


Figure 3.2: Global bean production and area cultivated (1991-2014)
Source: FAO (2016)

Production distribution is heavily concentrated in the Americas and Asian regions, which together are responsible for over three quarters of total world output (Appendix A1). Nevertheless, when compared to other regions and the world as a whole, Africa's annual production grew by a significantly higher rate between 1994-2014 (4.95%), representing nearly twice the world annual growth rate (2.7%) (FAO, 2016).

Figure 3.3 shows that the major producers of beans at country level are India (16.7%), Brazil (15%), Myanmar (11%), China mainland (8%) and United States of America (USA) (6%). These countries together were responsible for over half (56%) of the total world dry bean output in the last two decades.

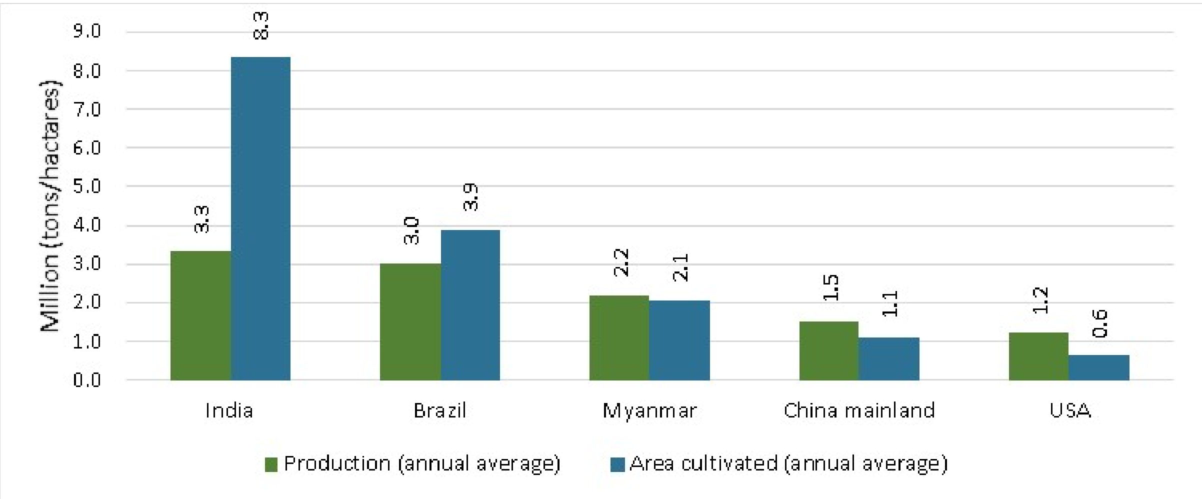


Figure 3.3: Top 5 global producing countries ranked by production quantity (1994-2014)
Source: FAO (2016)

Stellenbosch University https://scholar.sun.ac.za

3.2.3 Global trade of dry beans

In terms of trade, the global market for dry beans is evolving and indicates an upward trend. In a twenty-year period between 1994 and 2013, the world export quantity of dry beans increased by 60% while exports by value increased by 222% (FAO, 2016). Additionally, export quantity grew at a rate of 3% over the last two decades, 1994-2003 and 2004 -2013 compared to the 9% growth rate in value (FAO, 2016). This means that the price of dry beans in the global market is increasing at a higher rate than the quantity traded.

However, international trade of dry beans by FAO (2016) data, represented only 18% of total global production between 2004 and 2014. As shown in Figure 3.4, all regions in comparison to their production, participate minimally in international dry bean trade. In observing the geographical patterns of trade, data reveals that a bulk of dried common beans is traded between countries located within the same region than with countries outside the region (ITC [International Trade Centre], 2016b). This trend is consistent with the emerging intra-regional trading regime being fuelled by regional integration partnerships (Amikuzuno & Donkoh, 2012).

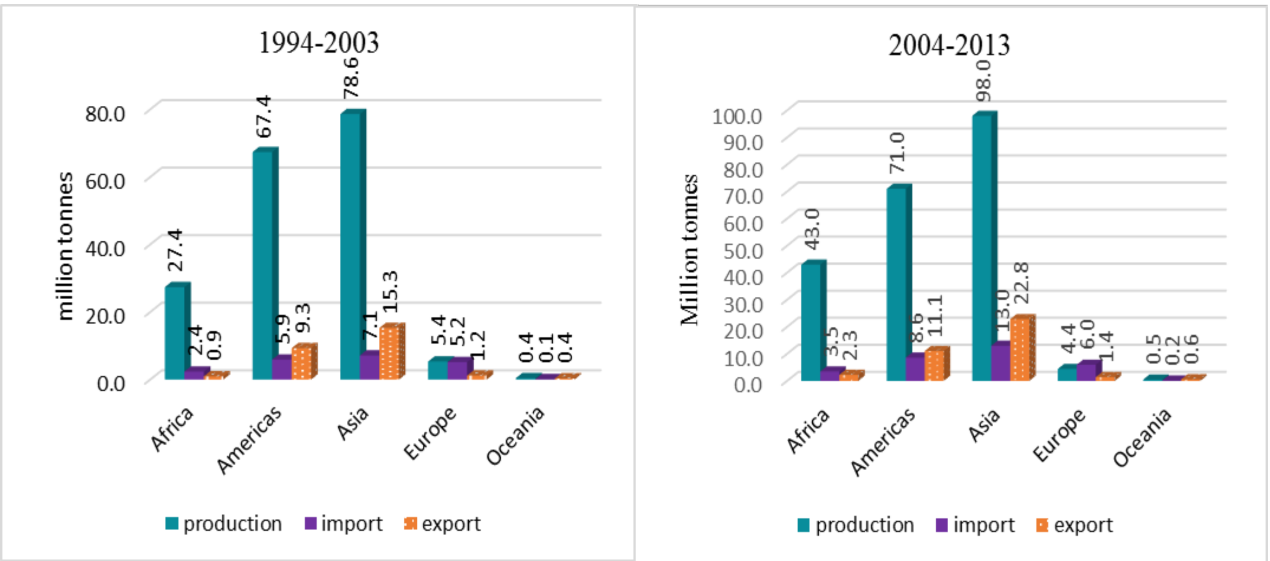


Figure 3.4: Regional share in global bean production, exports and imports (1994-2003) and (2004-2013)

Source: FAO (2016)

The principal global suppliers of beans are Myanmar (31%), China (23%), China mainland (23%), USA (11%) and Argentina (8%) who collectively handle 88% of world traded (exported) volumes (Figure 3.5). This clearly indicates that the world exports of dry beans is highly concentrated in few countries, particularly in Asian and American regions.

Stellenbosch University https://scholar.sun.ac.za

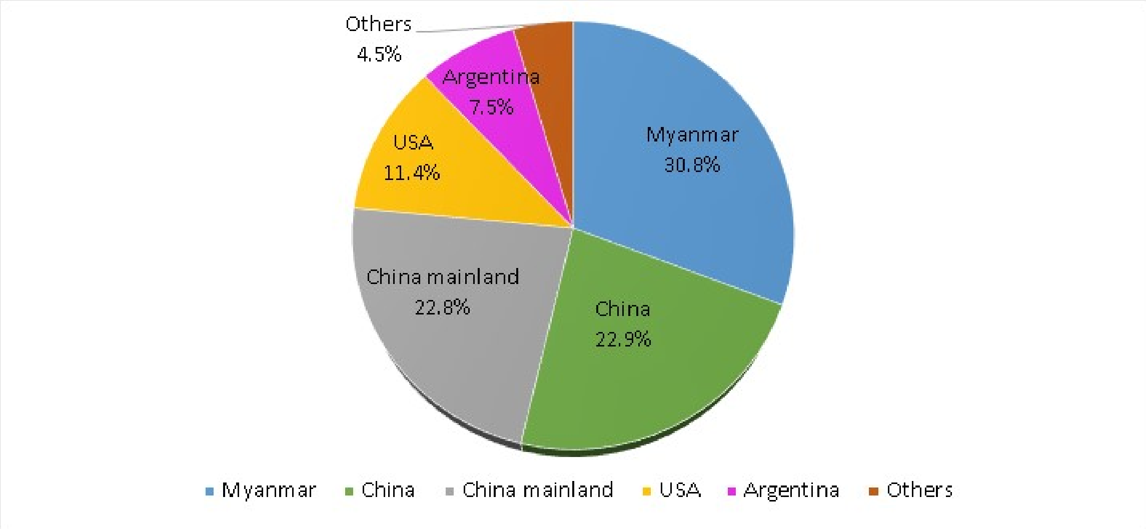


Figure 3.5: Top 5 global exporting countries ranked by export quantity (1994-2013)
Source: FAO (2016)

3.2.3.1 *Global Imports*

When imports are brought into perspective, statistics show that dry bean import volumes grew by 52% between the two decade period 1994-2003 and 2004-2014 (FAO, 2016). Although Africa's share in global import volume remained at 11% in both periods, her trade deficit grew smaller, from 1.5 million tons to 1.2 million tons, while that of Europe grew wider (4 million tons to 4.6 million tons) (FAO, 2016). These statistics suggest Africa could be slowly growing into a net exporter of dry beans.

In essence, the geographical spread of imports is more widely distributed as opposed to its exports (Figure 3.6). Over the twenty one year period from 1994 to 2014, India ranked the world's top importer of beans followed by Brazil (6%), Japan (5.1%), United Kingdom (4.7%) and USA (4.7%). The five countries together were responsible for about a third of world import volume (FAO, 2016). ITC (2016a) attributes the fast-growing pulse demand from Asian countries such as India, to their increasing population, steady economic growth and urbanization.

Stellenbosch University https://scholar.sun.ac.za

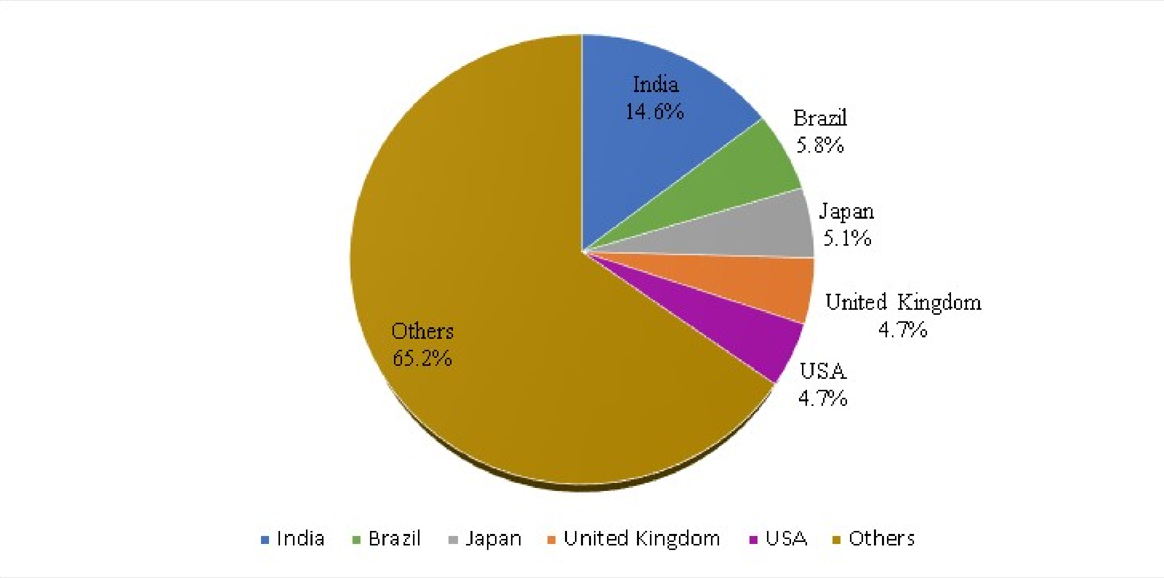


Figure 3.6: Top 5 global importing countries ranked by import quantity (1994-2013)
Source: FAO (2016)

3.3 Beans in Eastern and Southern Africa

The global context of the bean industry and its trade patterns indicate a growing industry, but more also the importance of regional trade. In this section, a perspective of the industry in its wider context at Eastern and Southern Africa (ESA) level is brought into light as well as the trading policy environment.

ESA is home to over half a billion (518.7 million) people (World Bank, 2016). The majority of this population, however, resides in rural areas where agriculture is their mainstay and poverty dominates (Van Rooyen, 2000). Agriculture, further, remains a backbone for many of the region's economies, contributing from as low as 3% in South Africa to as high as 42% in Ethiopian GDP (World Bank, 2016). On the other hand, the prevalence of malnutrition, food inadequacy and under malnourishment continues to be among issues of regional and national concern despite the general progress in most countries economic growth.

The enormous potential of pulses[6] to address many of these issues has been well documented. These include, but not limited to, their nutritional profile (a rich, yet inexpensive source of protein), health benefits, ready demand, disease and drought resistant, minimal production input requirement and a soil fertility booster (Van Heerden & Schönfeldt, 2004; De Luque & Creamer, 2014). Beans in particular is a staple food in ESA (Wortmann *et al.*, 1998; Katungi *et al.*, 2009).

[6]FAO defines pulses as a subgroup of legumes, belonging to the *leguminosae* family and produce edible seeds. This category only includes legumes harvested for dry grains including beans, cowpeas, chickpeas, pigeon peas, lentils, vetches, Lupins, Bambara beans and broad beans. Legumes such as soybeans and groundnuts are not considered as pulses.

Stellenbosch University https://scholar.sun.ac.za

By FAO (2016) estimates, total pulse production in ESA stood at 8.7 million tons in 2014, harvested from a total of 9.5 million hectares. Beans accounted for half of this production (4.3 million tons over 5.1 million hectares). Production is typically a small-scale farmers' activity, characterised by concentration in a few countries (Table 3.2) and specific regions within countries, seasonal and highly influenced by ecological conditions (Ackello-Ogutu & Echssah, 1998; Katungi *et al.*, 2009; FAO, 2016).

Figure 3.7 illustrates that a bulk of regional land allocated to beans is in the Great Lakes region including Uganda, Rwanda, Kenya and Tanzania. Other notable countries with a relatively large area under bean cultivation are Ethiopia and parts of Malawi and Mozambique. Wortmann *et al.* (1998) attributes this distribution to the favourable mean temperature during growing seasons and the density of the rural population. However, the fact that rain fed beans could be grown twice in a year in nearly all East African countries, such as Tanzania (Katungi *et al.*, 2009), could also partially explain this pattern.

Stellenbosch University https://scholar.sun.ac.za

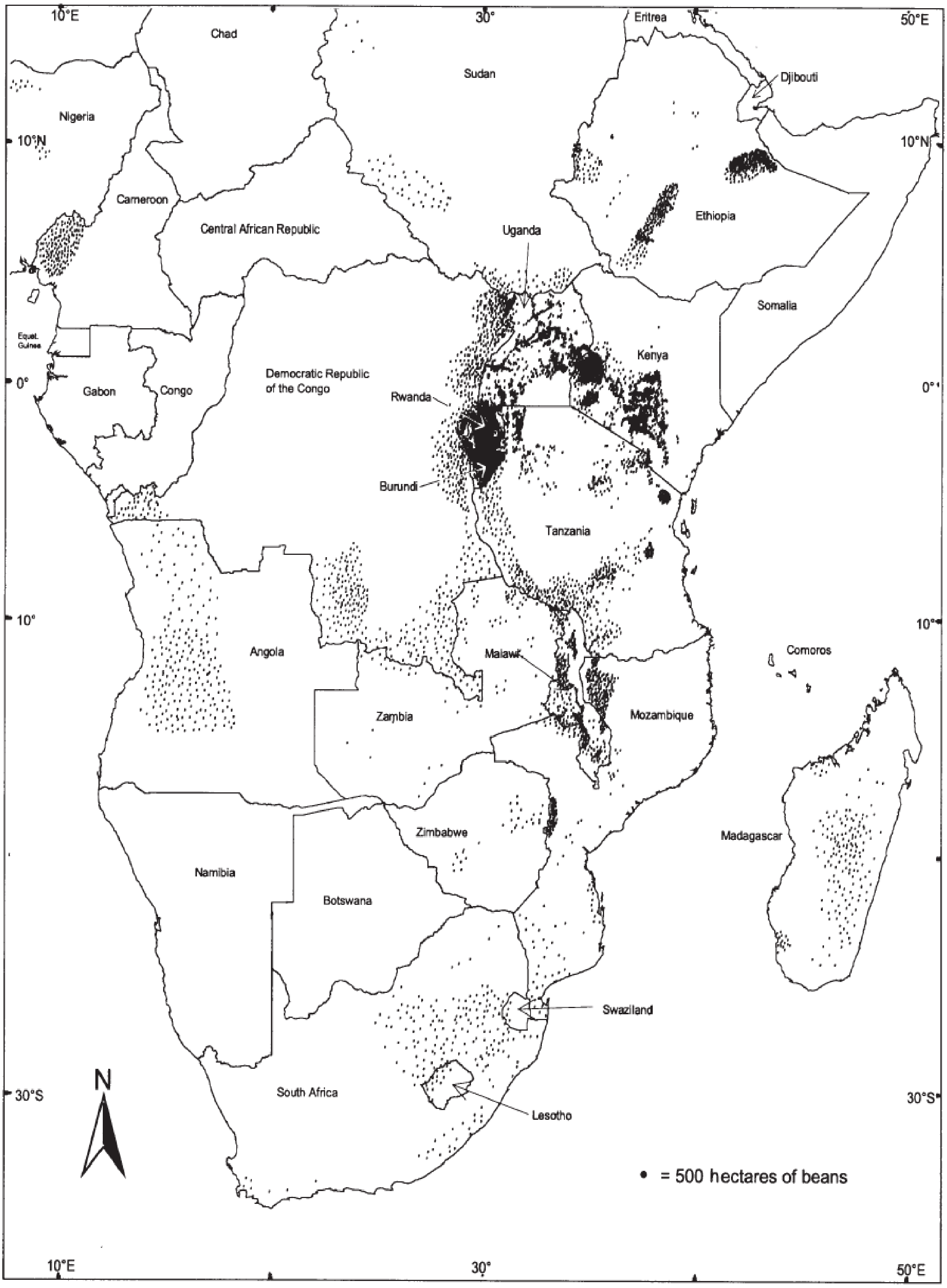


Figure 3.7: Dry bean production distribution in Sub-Saharan Africa
Source: Wortmann *et al.* (1998)

Stellenbosch University https://scholar.sun.ac.za

Table 3.1: Total dry bean production, trade and consumption in ESA (1994-2013)

Country	Production	Export	Import	Total Consumption	Per capita consumption
Tanzania	12,676,379	234,906	56,419	12,497,892	16.3
Uganda	9,407,218	361,565	95,705	9,141,358	16.8
Kenya	8,009,970	117,877	488,725	8,380,818	12.2
Rwanda	4,881,556	27,879	156,717	5,010,394	29.2
Ethiopia	4,050,261	1,125,367	43,091	2,967,985	2.0
South africa	1,283,973	109,480	1,267,208	2,441,701	2.6
Malawi	2,270,550	60,828	25,842	2,235,564	9.0
Madagascar	1,622,932	115,878	14,909	1,521,963	4.3
Zimbabwe	780,492	22,519	169,820	927,793	3.6
Zambia	819,072	8,113	56,745	867,704	3.7
Somalia	356,994	1,059	73,566	429,501	2.6
swaziland	36,181	7,932	37,449	65,698	3.0
Namibia	1,098	977	4,000	4,121	0.1

Source: FAO (2016)

Although beans are traditionally grown for subsistence, it has been steadily growing in importance as a market oriented crop. Wortmann *et al.*, (1998) estimates that approximately 40% of Africa's total production valued at US$452 million, is marketed annually. However, like any other food crop, the current domestic bean marketing is rooted in a historical context.

3.3.1 Historical evolution of agricultural commodity marketing

Prior to 1980s, agricultural marketing in ESA was characterised by heavy state control. Agricultural markets are important avenues for food price stability, income generation and food security. Consequently, governments in the region, as elsewhere in Africa, found it necessary to directly intervene in the operation of these markets. The nature and extent of state intervention varied from country to country depending on the patterns and nature of intervention of the preceding colonial government (Kherallah *et al*., 2002). Nevertheless, interventions in nearly all countries involved official price controls and legal monopolies in produce marketing (Barrett, 1997; Muyatwa, 2000; Kherallah *et al*., 2002). In the case of Zambia, market control was enshrined in the Agricultural Marketing Acts of 1969 and 1989 under which the state enterprise (National Agricultural Marketing Board) set pan-territorial and pan-seasonal[7] prices (Muyatwa, 2000).

However, pressure to liberalise markets came from various sources. Firstly, market control involved use of state subsidies, with significant negative effects on fiscal budgets. Secondly, inflation, stagnant economic growth, increasing dependence on foreign capital resources and shortages of consumer goods engulfed most countries (Kherallah *et al*., 2002). Coupled with donor pressure for structural adjustments, governments began taking steps to liberalizing agricultural markets. The initial attempts at liberalizations were made in the 1980s. The main objectives of these reform policies were to enhance economic efficiency, including eliminating state control of markets, devaluating currencies, liberalizing imports, promoting private trader

[7] Pan-territorial pricing refers to charging the same price throughout the country regardless of differences in marketing factors while Pan-seasonal pricing involves charging the same price throughout the year.

Stellenbosch University https://scholar.sun.ac.za

participation as well as the removal of input and output subsidies (Muyatwa, 2000; Kherallah *et al*., 2002). The state was to seize uniform pricing policies and allow market forces to determine producer and consumer prices.

The pace and extent of reforms however varied across countries and the crop commodity. In some instances, reforms have only been partially implemented, partly due to weak government commitment (Kherallah *et al*., 2002). In other cases, complete policy reversal has occurred driven by external shocks or changing economic conditions. For instance, government intervention is still pronounced in the Zambian maize sector. Kambewa (1997) also documents that the Agricultural Development and Marketing Corporation, a state marketing board still plays a dominant role in bean marketing in Malawi. The board buys beans from farmers at a set price and sells to retailers and institutions throughout the year. African Centre for Biodiversity (2016) adds that beans are included in Malawi's farm input subsidy programmes (FISPs).

Notwithstanding the Malawian case, and without ascertaining the effectiveness of these policies (a subject of many debates), it is true, however, that the dry bean sector is one of the few that can be argued to be fully liberalized in the region. Countries therefore have more or less similar domestic bean marketing policies, with spatial arbitrage a task for the private sector. Farmers and private traders set domestic prices based on competition, without influence from the state. However, traders have an upper hand in price determination, as they are fewer with better market information and market intelligence (Katungi *et al.*, 2009).

3.3.2 Regional trading arrangements

At regional level, however, trade is guided by policies directly or indirectly influenced by regional co-operation. Economic integration has a long history in ESA dating back to 1910, when the Southern African Customs Union (SACU) was established. Over time, several blocs have emerged with different objectives and often overlapping membership (Figure 3.8). However, SADC is the only bloc to which Zambia, DRC and Tanzania are all members (Figure 3.8). The three countries were all signatory to the Common Market for Eastern and Southern Africa (COMESA) until Tanzania pulled out in 2000.

COMESA's aim, since inception, has been to strengthen regional integration through promoting cross border trade and investment (Mishili, 2009). Zambia and DRC are among 16 of the 21 member countries currently participating in the COMESA Free Trade Area (FTA). Recently, the trading bloc launched a Simplified Trading Regime (STR) aimed at motivating small scale regional traders to use formal trading channels. Consignments worth US$1, 000 and less, under this regimes, can clear with minimal paperwork and without inspection by clearing agents.

Similarly, SADC member countries signed a trade and economic development protocol in 1996, leading to the launch of a FTA in 2000. Through this protocol, the bloc set out to liberalize 85% of all regional goods traded by 2008, harmonize customs rules and procedures and eliminate non-tariff barriers. More recently (in 2016), the East African Community (EAC),

Stellenbosch University https://scholar.sun.ac.za

to which Tanzania is a member, SADC and COMESA came together and launched a Tripartite FTA. Zambia, DRC and Tanzania are all signatory to this pact. The tripartite FTA is based on three pillars covering industrial development, market integration and infrastructural development. Through this pact, the three blocs aim to formulate a single customs territory, liberalize intra-regional trade and facilitate wider regional investment. In so doing, it seeks to harmonize the often-conflicting obligations and custom rules across countries belonging to the different blocs.

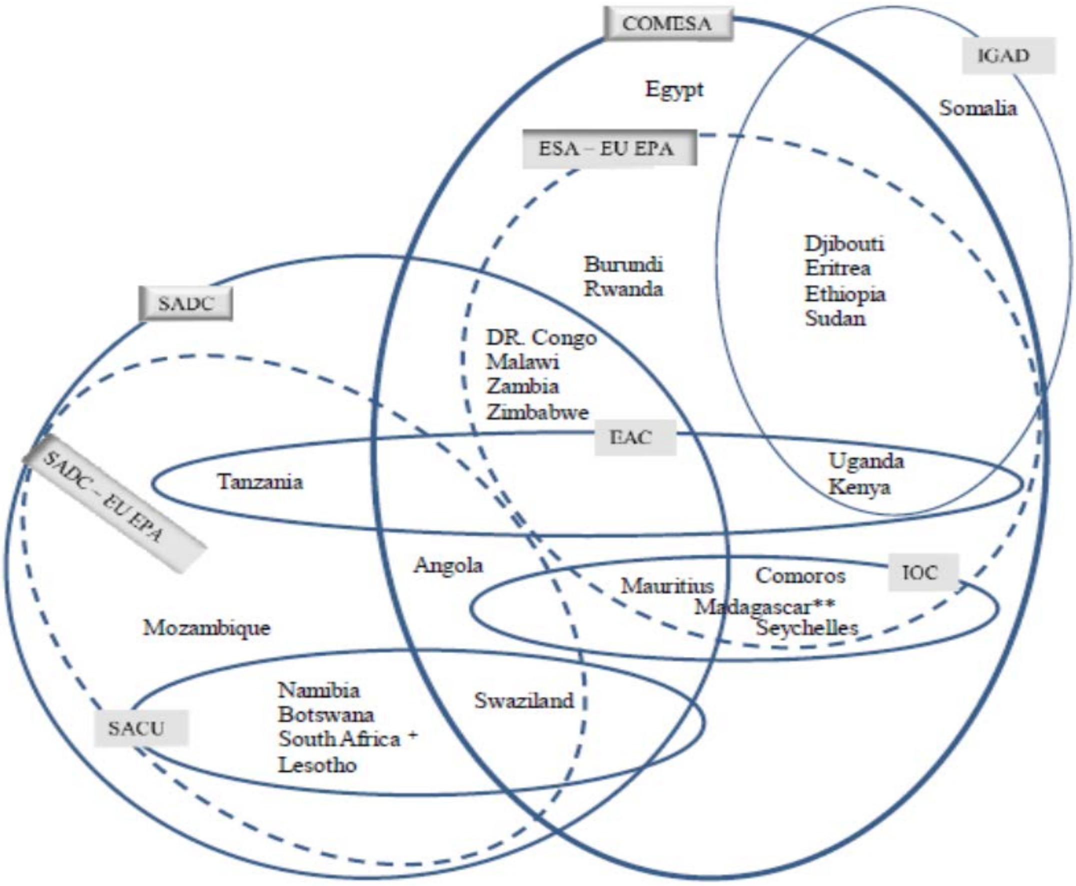


Figure 3.8: Regional trading agreements in ESA
Source: Bilal & Szepesi (2005, cited in Mishili, 2009).

3.3.3 Regional trade policies

However, while regional agreements have succeeded in reducing customs tariffs for pulses[8], formal trade has continued to encounter numerous obstacles that leaves it less attractive. Small scale traders are hindered even more by these obstacles. With the reduction of tariffs, the importance and prevalence of non-tariff barriers have increased. According to Korir *et al.* (2003) for any cross border bean trade, the Tanzanian tropical pesticide research institute requires an import or export permit, a phytosanitary certificate costing US$15 per consignment in addition to a border tax of US$2. Moreover, the Tanzanian government, through the country’s export control act of 2008, often restricts bean exports on food security grounds (Korir *et al.*, 2003). ITC (2016a) concurs with this, describing the administrative procedures as

[8] Bean tariffs are set for 0% in the SADC agreement (Bese *et al.*, 2009).

Stellenbosch University https://scholar.sun.ac.za

often burdensome and costly. They further add that import and export of pulses must always be accompanied by a radiation certificate, whether the importing country requires one or not. This is based on the country's Atomic Energy Act of 2002.

In Zambia, agricultural trade is primarily guided by the commercial, trade and industry policy of 2009, which is currently under review (Bese *et al.*, 2009). In line with the revised sixth national development plan (2013-2016), the trade policy aims at enhancing the competitiveness of local products in regional and international markets. However, Zambia pursues a food self-sufficiency policy. Trade procedures therefore include standard and quality requirements as well as permits to control trade flow while ensuring domestic food security. Bean traders require permits and must also obtain phytosanitary documents. Moreover, a certificate of origin is required for all intra-SADC (COMESA) traders in order to benefit from SADC (COMESA) privileges.

3.4 Bean market in Zambia

Zambia is a landlocked country located in Southern Africa. It shares a common border with eight other countries, Angola and Namibia to the west, Botswana, Mozambique and Zimbabwe to the south, Malawi to the east, Tanzania to the north-east and DRC to the North. Consequently, her major routes to the seaports are through one or more of these countries. The two dominant routes are the Dar es Salaam seaport via Tanzania and the South African sea port via Botswana and Zimbabwe (Muyatwa, 2000). The country's strategic location and proximity to these countries further enhances its regional trade potential.

According to the UNDP (2015), almost 15 million people populated Zambia in 2014, of which 40% were located in the urban areas and 45.8 percent of children classified as malnourished. The country's GDP per capita stood at US$3 800 in 2013 (UNDP, 2015). While copper mining is the major driver of Zambia's economy, agriculture remains a major contributor to the economy at slightly less than 20% of GDP and providing over 70% of employment. Zambia's one million farmers can be divided into small scale, medium and large scale farmers. As of 2003, the small-scale category made up 94% of farmers with only 0.5% classified as large scale (Siegel & Alwang, 2005). In addition, the country has a variety of climatic conditions and soil types favourable for production of a variety of crops.

Until the beginning of 1990's, Zambia's agricultural policies were biased towards maize, with little emphasis on food legume crops such as beans (Hamzakaza *et al.*, 2014). After liberalization, however, the national policy has been adjusted and refocused to lay emphasis on crop diversification identifying beans as a priority food crop (Hamzakaza *et al.*, 2014).

3.4.1 Market Structure

Beans are the second most important legume crop produced in Zambia after groundnuts. They are also among food staples heavily traded across the country border informally, with regular exports to DRC and imports from Tanzania, providing income to producing households. Although produced throughout the country, the majority of the crop commodity, around 60%, is produced in only two of Zambia's ten provinces, Northern and Muchinga (Sichilima *et al.*,

Stellenbosch University https://scholar.sun.ac.za

2016) while the most urbanised provinces, Lusaka and the Copper Belt contain the majority of the net bean purchasers. In essence, the net producing provinces are basically an extension of Tanzania's southern highland, except that they are demarcated by a border.

Transport is an important component in the flow of beans from surplus to net consumption zones. Being landlocked, the country is almost entirely dependent on the road network, as rail options are extremely limited. However, like many other countries in ESA, Zambia suffers from a poorly developed market infrastructure system (Muyatwa, 2000). Poor road networks and infrastructure development has a negative implication on agent participation in output markets both within and across countries and hence increase chances of spatial market imperfection (Loveridge, 1991; Muyatwa, 2000; Kabbiri *et al.*, 2016). In a nut shell, the existence of good road networks connecting major towns may not be a problem, the challenge however is with the state of roads (gravel) in rural areas. Small scale farmers, who dominate the beans industry, are usually found in areas where an all-weather road may not be assured (Ackello-Ogutu & Echessah, 1998; Muyatwa, 2000). Additionally, the gravel road connecting Zambia to Lubumbashi (DRC) is in a bad condition and worsens during rainy season. Rail options are non-existent. Attempts have been underway to upgrade the DRC-Zambia road but until then, movement of beans to DRC will be subject to high transfer costs caused not only by distance between surplus and deficit zones and fuel costs, but also the infrastructure-related challenges.

3.4.2 Production

As noted earlier, beans in Zambia are grown throughout the country, although regional production concentration is evident. Total production has been increasing over the last two decades (Figure 3.9). Data suggests that the increase could more likely be attributed to the increasing land allocated to bean production and less to the increase in productivity. Figure 3.9 shows an increasing trend in both production and area harvested, although local demand for beans still outweighs local production. By Muimui's *et al*. (2016) estimates, Zambia annually produces an average of 50,000 to 60,000 metric tonnes of beans against an estimated national demand of 100,000 metric tonnes. The remaining amount is thus sourced from imports.

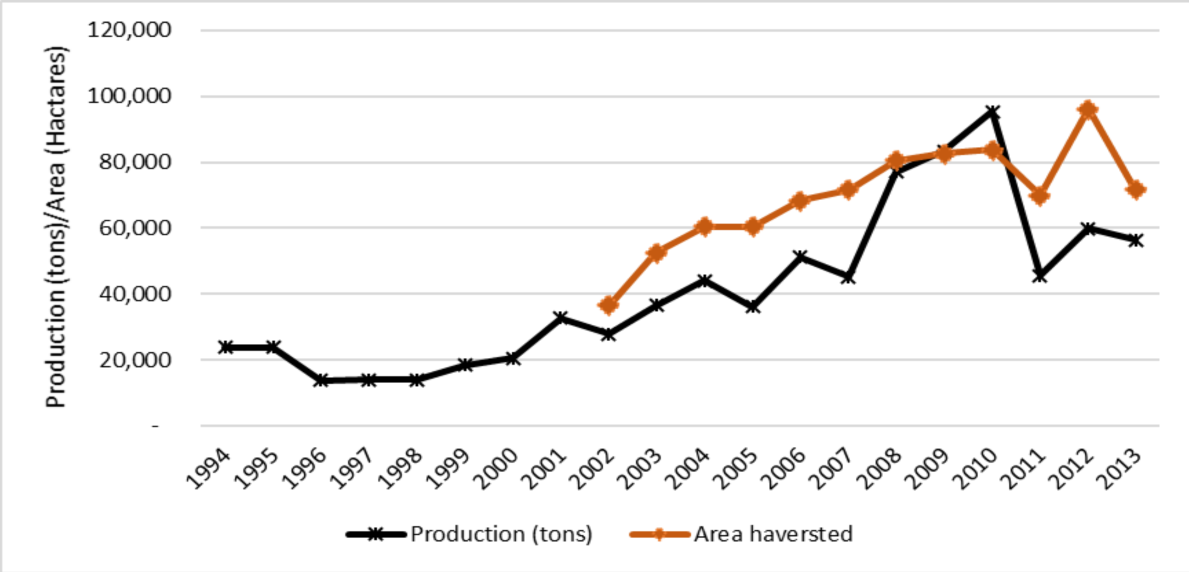


Figure 3.9: Total bean production and area harvested in Zambia (1994-2013)
Source: Authors plot based on data in Zulu *et al.*, (2000), Tembo & Sitko, (2013) and FAO (2016)

Most beans are produced for subsistence using minimal inputs. A typical farmer grows a mixture of varieties, although farmer and market preference determines the varieties grown. Bean production is almost exclusively rain-fed and therefore seasonal and influenced by ecological conditions.

Rainfall varies across the country, dividing Zambia into three distinct agro-ecological zones. Zone I located in the northern part of the country, accounting for 46% of Zambia's land area (Muyatwa, 2000), has the highest annual average rainfall (more than 1,200mm annually), although this has led to extreme soil acidity due to leaching. This Zone is also the bordering region between Zambia and DRC and Tanzania. Zone II (which covers the central and eastern parts of Zambia) receives medium rainfall (800mm annually) that is more evenly distributed throughout the growing season and has fertile soils. Finally, Zone 3 in the south has the most unpredictable rainfall patterns (an annual average of less than 800mm), poorly distributed with higher episodes of drought. The unreliable rainfall patterns in this Zone make crop production risky. Nevertheless, the unimodal rainfall pattern throughout the country, can only allow for one possible harvest (between November and April) per year. The regional differences in climatic conditions thus create natural difference in production potential.

3.4.3 Consumption

Notwithstanding the production differences, bean consumption is by a wide array of people in both rural and urban areas. Since consumption data is virtually unavailable in Zambia, an attempt was made to estimate consumption patterns over the last two decades following Akibode and Maredia' s (2011) approach. Consumed quantities in Figure 3.10, refer to a sum total of local production and imports less exports, while per capita consumption denotes consumption per total population. It should be noted that the analysis employed trade data

Stellenbosch University https://scholar.sun.ac.za

recorded by FAO. Given that such data excludes informal imports and exports, it is possible that the estimated figures here may understate or overstate the actual per capita consumption.

Notwithstanding the data quality, bean consumption shows an upward trend in total quantity and per capita terms (Figure, 3.10). Per capita consumption, particularly, grew at a rate of 2% with an average of 3.7 kg/capita/year between 1994 and 2003 (FAO, 2016). These findings are slightly lower, but proportionately consistent with the 10kg/capita/year reported in Sitko *et al.* (2011) for the 2009/2010 season. They both indicate low levels of consumption.

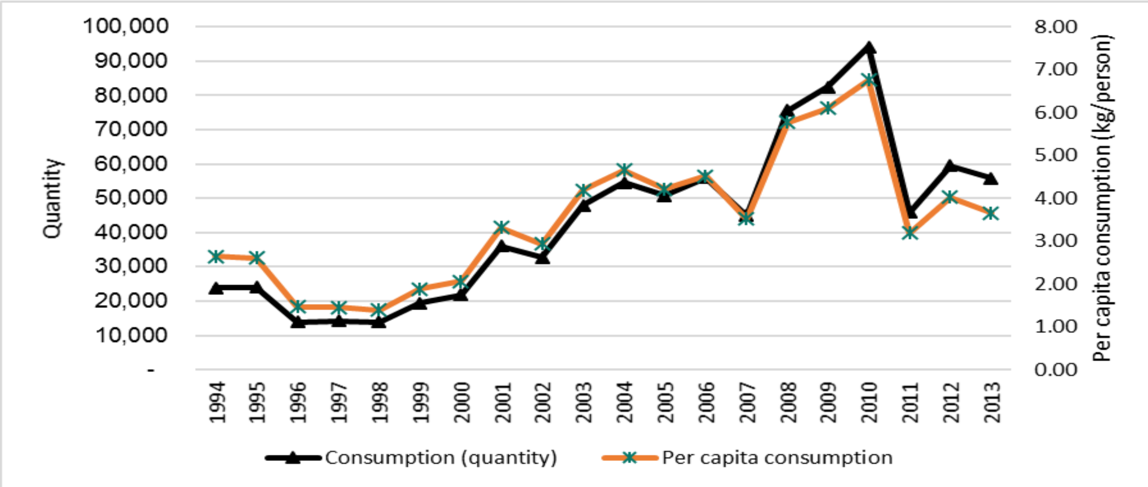


Figure 3.10: Bean consumption in Zambia (1994-2013)
Source: Authors plot based on data in Zulu *et al.* (2000), Tembo & Sitko (2013) and FAO (2016)

The majority of consumers depends on markets for their bean consumption. Domestic demand consists of local household purchases, institutional demand (prisons, schools, hotel/restaurants and hospitals) and urban demand. Bean marketing is entirely a private sector activity, conducted in both formal and informal markets. The marketing chain is basic, flowing from bean sources (farmers or imports), either directly to local consumers, middlemen who come to the farm gate or to intermediaries including retailers, institutions or wholesalers.

The major in-country flow is from northern provinces to the Copperbelt and Lusaka. The liberalized domestic market system provides an opportunity for effective spatial arbitrage. However, factors such as seasonal production, poor/inadequate storage facilities, poor marketing infrastructure, high transportation costs and poor access to market information constrain the extent of arbitrage.

3.5 Bean market in Tanzania

In comparison to Zambia, the bean industry in Tanzania receives more attention from a policy perspective. ITC (2016a) reports several policies directly or indirectly guiding the Tanzanian pulse industry. While the National Agricultural Policy serves as the major long term plan guiding the sector, the national road map for pulses 2016-2020, which is among the most recent

Stellenbosch University https://scholar.sun.ac.za

commitments to the industry, is centred around enhancing the industry's value chain, in a bid to position Tanzania as a key player in the global pulse market.

3.5.1 Production and consumption

Tanzania is the leading producer of beans in ESA and the entire Africa (Table 3.2). The country ranks 7th in world production (Ronner & Giller, 2013; FAO, 2016). In a typical year, Tanzania produces more beans than Namibia, Zambia, DRC, Swaziland, South Africa, Zimbabwe, Malawi, Lesotho, Madagascar, Burundi and Sudan combined (FAO, 2016). Figure 3.11 illustrates that the main producing areas are the mid to high altitude areas, covering Arusha region in the north, the Great Lakes region in the west and Tanzania's neighbouring region with Zambia, the Southern highlands (around Mbeya and Rukwa regions), where temperatures are cooler during growing season and rainfall is more reliable (Hillocks *et al.*, 2006). Mbeya and Rukwa regions are particularly characterized with rainfall ranging between 650 mm to as high as 2600 mm annually, with a warm and arid climate (Bese *et al*., 2009). While beans rank the third largest cultivated crop after maize and cassava in Tanzania, they are second only to maize in the southern highlands (Bese *et al*., 2009).

Tanzania planted an average of 806,636 hectares of beans per year between 1994 and 2004 (FAO, 2016). Unlike Zambia, bean production in most parts of Tanzania occurs during two main cycles. The first runs from October to January while the second runs from March to June. Production in unimodal areas occurs between November and May. Although there are a few commercial farms cultivating beans, production is typically in the hands of small scale farmers who usually intercrop beans with maize (Katungi *et al.*, 2009; Mishili, 2009; Bese *et al.*, 2009). Figure 3.16 shows that total bean production has increased in the last twenty years.

Beans is also considered an important part of the diets of the Tanzanian population. On average Tanzania consumed around 624,895 tons of dry beans each year between 1994 and 2013, that was supplied mostly from domestic production and supplemented by imports, mainly from Kenya (FAO, 2016; ITC, 2016b). Additionally, per capita consumption averaged 16 kg/person/year over the same period which is four times higher than Zambia (3.7 kg/person/year). The major bean deficit areas in the country, shown in brown on Figure 3.11, include Dar es Salaam, Simiyu, Lindi and most central regions. Existing market demand for beans is high, however, trade is constrained by market infrastructure challenges, limited access to market information as well as limited bean inventory management (Ackello-Ogutu & Echessah, 1998; Bese *et al.*, 2009)

Stellenbosch University https://scholar.sun.ac.za

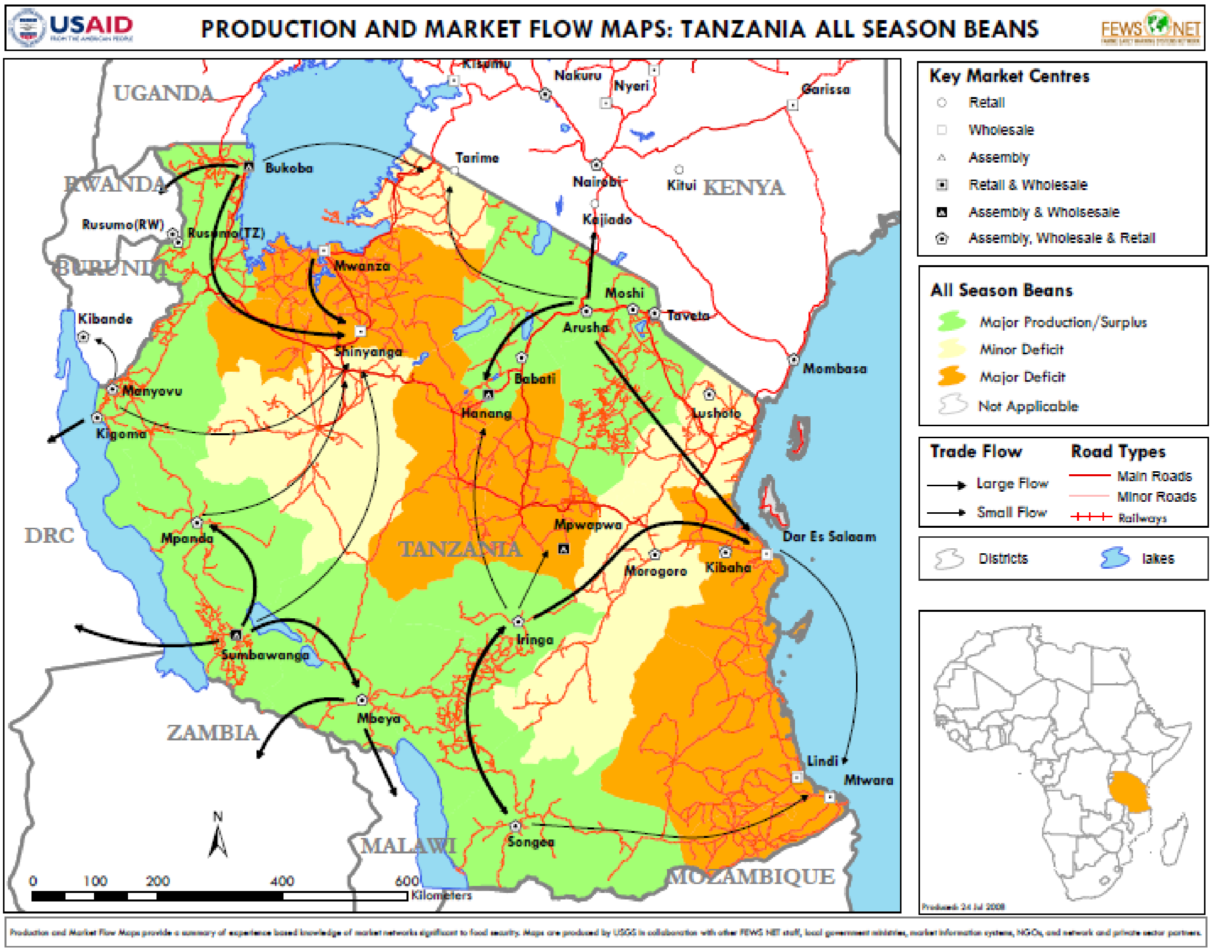


Figure 3.11: Tanzania bean production and trade flow map
Source: fewsnet.net

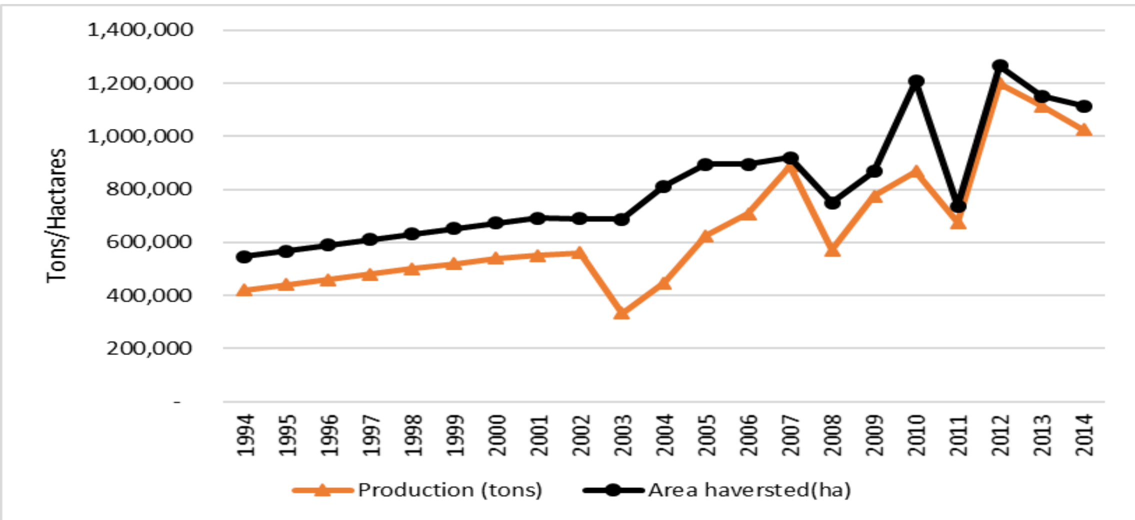


Figure 3.12: Bean production and area cultivated in Tanzania (1994-2013)
Source: FAO (2016)

Stellenbosch University https://scholar.sun.ac.za

3.6 Bean market in the DRC

Located in the north of Southern Africa, DRC is the second largest country in Africa, and is endowed with enormous potential for crop production, including beans. Unlike Zambia and Tanzania, some regions in DRC (equatorial region) have capacity to produce crops throughout the year. Additionally, the main production cycle north of the equator runs from April to November, basically alternating with the region south of the equator, where the cycle runs from October to May. Nevertheless, DRC remains one of Africa's food deficit areas. FEWSNET (2015) attributes this to the non-favourable macro-economic environment which is biased towards the mining sector. More than this, however, widespread internal displacement caused by decades long political conflict, instability and natural disasters render long term investment in the sector one of the risk adventures (FEWSNET, 2015).

Additionally, the country's basic market infrastructure is poor with only about 2% (2, 700km out of 153, 497 km) of all roadways paved (FAO, 2012). Minten and Kyle (1999) concurs with this sentiment adding that the bad road conditions curtail access to some interior areas.

3.6.1 Production and consumption

Annual bean production averaged 145, 654 tons between 1994 and 2013, with a growth rate of 1.7%, clearly outpaced by population growth (2.7%) (Institut National de la Statistique [INS], 2015). Katanga, the bordering province with Zambia which is also home to Lubumbashi, is among DRCs bean deficit areas, producing on average 5,000 tonnes annually (FEWSNET, 2015). The majority of national production (over half) comes from North Kivu and South Kivu provinces, with an annual production 5 to 10 times higher than that of Katanga (FEWSNET, 2015; INS, 2015). However recent statistics indicate significant production potential and growth from Oriental province (INS, 2015). Non-Governmental Organizations (NGO) offer support to the sector by proving technical assistance and access to improved inputs (FEWSNET, 2015).

3.7 Zambia's trade with DRC and Tanzania

Trade in food staples between Zambia and her neighbours has a long history. Originally a relationship built on the struggle for independence, trade relations with Tanzania particularly strengthened with the establishment of the Zambia-Tanzania (TAZARA) railways. Although country proximity would seem an obvious reason for trading, Bese *et al.* (2009) observed that the poor connection of some regions to their high value domestic markets render cross-border markets more attractive but regional economic integration has also played a role.

DRC and Tanzania are Zambia's key dry bean trading partners. Zambia is a net importer of beans from Tanzania while DRC is a net importer of beans from Zambia. The southern highland, Mbeya and Rukwa in particular supply the most beans to Zambia (Bese *et al.*, 2009) while Katanga province in DRC is Zambia's key export destination. However, official statistics overshadow this fact and understate the importance of intra ESA bean trade because a significant volume of dry bean trade is unrecorded in official statistics. Table 3.3 shows the proportion of formal dried beans (HS 071339) from Tanzania, Zambia and DRC traded within

Stellenbosch University https://scholar.sun.ac.za

ESA over the past six years. Although Tanzania's formal trade partners appear to be more diversified beyond the region, ESA is the main trading region for Zambia and DRC. Despite this, however, nearly none of these volumes are traded between the three countries (ITC, 2016b). Yet, Table 3.4 indicates significant quantities crossing Zambian borders though informal channels.

In 2015, for instance, Zambia's informal exports (3,792 tons) to Tanzania and DRC alone represented over 50% of what was recorded as formal export (7, 263 tons) to the rest of the world, while informal imports from Tanzania alone (9,799 tons) was about 24 times higher than the formal figure (410 tons) to the rest of the world. This means that informal trade with the selected partners constituted about 35% and 96% of total (formal plus informal) exports and imports respectively (Table 3.4). Similarly, the proportion of informal export and import to Zambia alone accounted for 29% and 85% of total trade (Table 3.4). This gap is expected to be substantially higher if all countries were strictly monitored for informal regional trade flows. Burke and Myers (2014) observed similar patterns in maize trade between Zambia and DRC. As discussed in section 3.3.2, the three countries are all members of SADC, while DRC and Zambia are also members of COMESA. Although bean trade among SADC member countries is practically free (zero tariff), trading through informal channels appear to be a more attractive option.

Stellenbosch University https://scholar.sun.ac.za

Table 3.2: Share of Intra-SEA in Tanzania, DRC and Zambia's dried beans (*HS 071339*), exports and import (2010-2015)

Country	year	Total import		Imports from SEA		
		Quantity(tons)	Value (US$)	Quantity(tons)	Share in total (%)	Value (US$)
Tanzania	2010	126	24	114	90	12
	2011	214	156	102	48	35
	2012	40	14	21	53	5
	2013	12	8	0	0	0
	2014	636	446	626	98	400
	2015	514	226	0	0	0
Zambia	2010	238	155	238	100	155
	2011	46	49	46	100	49
	2012	116	174	115	99	174
	2013	57	156	56	98	153
	2014	97	397	97	100	395
	2015	314	519	314	100	518
DRC*	2010	4,985	1,707	1,287	26	702
	2011	184	94	123	67	70
	2012	1,162	840	1,156	99	834
	2013	2,421	1,551	2,155	89	1,245
	2014	2,101	1,187	2,095	100	1,176
	2015	4,115	1,959	4,113	100	1,955

Country	year	Total export		Exports to SEA		
		Quantity(tons)	Value (US$)	Quantity(tons)	Share in total(%)	Value (US$)
Tanzania	2010	11,127	6,373	3,516	32	635
	2011	6,842	2,353	788	12	225
	2012	11,592	6,594	3,714	32	1,038
	2013	3,727	1,426	1,857	50	398
	2014	39,558	6,263	35,509	90	4,471
	2015	2,540	1,561	181	7	24
Zambia	2010	261	223	261	100	223
	2011	37	33	37	100	33
	2012	235	52	235	100	52
	2013	261	210	151	58	162
	2014	2,530	1,483	2,270	90	1,338
	2015	426	221	26	6	4
DRC*	2010	0	2	0	-	0
	2011	8	3	8	100	1
	2012	289	96	0	0	0
	2013	1	1	1	100	0
	2014	0	0	0	-	0
	2015	1	1	0	0	0

Source: ITC (2016b)

Stellenbosch University https://scholar.sun.ac.za

Table 3.3: Formal versus informal annual dry bean trade (2011-2015)

	Formal[1] trade					
	Zambia		Tanzania		DRC	
	Export (tons)	Import (tons)	Export (tons)	Import (tons)	Export (tons)	Import (tons)
2011	55	334	5,101	832	8	16,173
2012	537	148	4,626	1,426	289	10,514
2013	667	66	2,439	4,135	1	5,619
2014	4,226	111	45,099	3,223	-	18,268
2015	7,263	410	24,011	136	1	7,222

	Informal[2] trade					
	Zambia		Tanzania[3]		DRC[4]	
	Export (tons)	Import (tons)	Export (tons)	Import (tons)	Export (tons)	Import (tons)
2011	3,210*	2,694[c]	5,225	260	NA	2,692
2012	3,313*	3,245[c]	5,138	409	NA	2,688
2013	4,494*	9,033[c]	13,059	414	NA	3,894
2014	3,836*	11,081*	10,236	343	NA	3,301
2015	3,792[E]	9,799[G]	9,799[K]	656	NA	3,136

Source: 1 = ITC (2016b) 2 = FEWSNET

Note: 1 = official trade to the rest of the world, 2 = informal trade to selected neighbours, 3 = trade with Malawi and Zambia only, 4 = trade with Zambia only. * = export to DRC, Mozambique, Tanzania, Zimbabwe only. c = imports from Malawi and Tanzania only. NA = data not available. E = export to Tanzania and DRC only. G = imports from Tanzania only. K = export to Zambia only.

An analysis of ICBT trends over the last decade in Figure 3.13, indicates informal bean trade is a growing activity.

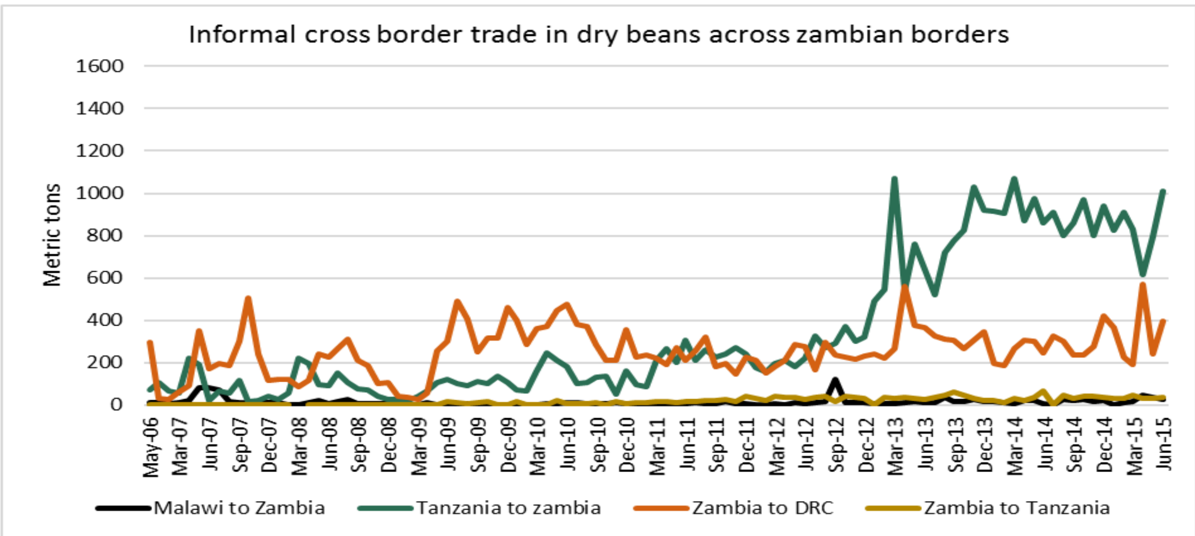


Figure 3.13: Historical patterns of informal bean trade between Zambia and her selected neighbours (2006-2015)

Source: Based on data collected by FEWSNET, Zambia.

Stellenbosch University https://scholar.sun.ac.za

3.8 An overview of Informal Cross Border Trade (ICBT)

By nature, informal cross border trade is surrounded by controversy, myths and ignorance (Little, 2010; Ogalo, 2010). Some analysts believe such trade simply reflects a normal market response to time consuming and cumbersome foreign trade procedures (Little, 2010). For others, however, it is nothing but an illegal activity, that robs government of essential revenue and a source of unfair competition in the market (Little, 2010). The latter view explains the neglect of ICBT in mainstream trade policy. Despite the negativity, however, ICBT continues to thrive in many economies constituting a significant proportion of intra-ESA trade. Understanding key aspects of this form of trade is thus crucial, when one needs to examine the functioning of markets they connect.

3.8.1 Defining ICBT

There is no universally accepted definition for ICBT. Some of the applied definitions in the literature include trade constituting transaction across national borders, that is conducted outside the formal framework. It is informal because the commodities traded are mostly sold and bought in informal markets. Little (2010) however describes this definition as a misconception arguing that, ICBT has strong ties to the formal sector. For instance, goods may be imported through informal transactions but eventually get sold through licensed domestic shops by registered traders.

ICBT has also been described as trade consisting of commodities exchanged across the border, that either pass through unofficial routes, avoiding customs checks and recording points, or goes through the official routes, gets subjected to customs control, but involve illegal practices at the customs office, such as deliberate misclassification of the goods, under invoicing, bribery and misdeclaration of the country of origin (Lesser & Moisé-Leeman, 2009).

In Ogalo's (2010) view, however it refers to exports and imports of legitimately produced goods which directly or indirectly escape all regulatory procedures set by government. Such trade usually goes unrecorded or incorrectly recorded into official statistics of the trading countries and is neither inspected nor taxed through official channels. Consequently, it is difficult to ascertain its true extent.

Other definitions attribute ICBT to trade in merchandise, which may be legal imports or export on one side of the border and illegal on the other side or *vice versa*, on account of not having been subjected to statutory border formalities such as customs clearance (Afrika & Ajumbo, 2012). This definition means the "informality" only begins at the border.

In summary, it is clear that what makes ICBT informal is the trader's failure to comply with trade related regulations and not necessarily that the products traded are in themselves illegal to use or trade. For purposes of this study, ICBT is used in line with Ogalo (2010) and refers to export and import of legitimately produced goods, usually by small scale entrepreneurs, that do not go through the legally established trade procedures and hence not taxed nor controlled by government.

Stellenbosch University https://scholar.sun.ac.za

3.8.2 History and status

Trade through informal channels is as old as the informal sector. It dates as far back as the pre-colonial era when traders moved freely to any far and near place to exchange goods through a barter system. Africa inherited arbitrary political borders at independence as national boundaries and imposed tariffs and non-tariff measures to transborder trade, declaring ICBT illegal. It is thus argued that what is seen today as ICBT could only reflect long-standing, indigenous and normal patterns of transborder trade, that makes more sense than formal trade channels (Meagher, 1997).

In essence, there is no country without an informal sector and so goes with ICBT (Ogalo, 2010). As for Africa, the informal sector is estimated to represent 43% of official GDP (Lesser & Moisé-Leeman, 2009; Ogalo, 2010). Further, only about 10% of total intra-African trade is official (Keane *et al.,* 2010). Nevertheless, studies and data collected by monitoring institution reveal a thriving ICBT (Ogalo, 2010; Brenton *et al.,* 2014; Burke & Myers, 2014). Afrika & Ajumbo (2012) reports that 30-40% of total intra-SADC trade is conducted through informal channels, valued at US$17.6 billion per annum (Southern Africa Trust, 2008; Brenton *et al*., 2014).

It is a source of income and employment to a significant number of people. Additionally, it contributes towards regional food security, women's empowerment and if properly harnessed, has potential to support the on-going efforts of poverty alleviation (Ogalo, 2010).

3.8.3 Drivers and trader characteristics

In essence, there is no single cause for the growth and prominence of ICBT between Zambia and her neighbours, as elsewhere in ESA. Some traders simply want to stick to an old practice and a trading norm that was only interrupted by the introduction of artificial borders (Ogalo, 2010; Little, 2010). Strong ethnic ties continue to unite these communities beyond borders and make informal trade with each other a much more normal form of trade. To such traders, it matters less whether the formal trading environment and procedures are attractive or not. People are just more inclined to trade across the border without government permits.

For others, however, factors perceived to increase the cost of trading formally compel them to participate in ICBT. These include improperly co-ordinated and restrictive trade policies; cumbersome, rigid and time consuming trade procedures, tariffs and wide spread corruption. At times complying to formal trade standards raise transaction costs that may not be covered by profit margins (Ogalo, 2010; Brenton *et al.,* 2014). For instance, a small scale trader dealing in beans and located near the border is often required to obtain certain documents only issued in the city centres, miles away from the border. The flexibility that comes with informal trading then provides a mean to transact on minimal costs.

Further, the SADC protocol provides for the elimination of non-tariff barriers. However, countries often impose trade bans and restrictions on food security grounds. A case in point is maize from Zambia and Tanzania (Myers & Jayne, 2012; Burke &Myers, 2014). Other barriers

Stellenbosch University https://scholar.sun.ac.za

include bureaucratic customs procedures and delays at road blocks and borders. As a consequence, some traders opt to avoid formal channels in order to facilitate faster trade.

The growth of ICBT in ESA has also been attributed to the shrinking formal economy. According to Ogalo (2010), the structural adjustment programs aimed at restructuring developing economies and eliminating government intervention, resulted in increased urbanization and a shrinking formal economy. Faced with unfair competition, low wages and lack of employment in the formal sector, informal economy and ICBT in particular offered a survival alternative. This view is in line with Kherallah *et al*., (2002) who earlier on noted that the reform policies have resulted in significant expansion of ICBT.

Essentially, ICBT is regarded as a small scale trader's activity (Little, 2010). Their small consignments make it easier to cross borders often unnoticed, sometimes just on foot or bicycle (Burke & Myers, 2014). However, this is not always the case, as large scale traders also participate in ICBT, usually by ferrying goods in smaller quantities which are later added together on the other side of the border.

These traders can also be distinguished between formal (registered) firms evading all trade related regulations and formal (registered) firms who partially evade trade regulation for instance, by only under invoicing or misdeclaring their goods. Beans across Zambian borders are nevertheless mostly traded by informal (unregistered) merchants, who operate entirely outside the spheres of policy influence (Burke & Myers, 2014). Both wholesalers and retailers participate in ICBT although women traders dominate (Brenton *et al.*, 2014).

3.8.4 Trader challenges

ICBT is generally easier, cheaper and a faster alternative to formal trade. Nevertheless, it is characterised by a number of obstacles with implications on effective market functioning. Traders usually operate under minimal resources. They cannot access credit from financial institutions since their business is in most cases unregistered. Trade is run on capital raised from traders' own resources or credit obtained on bilateral agreements from peers and family (Afrika & Ajumbo, 2012). As a consequence, lack of capital and financial instability is high and limits both business entry and expansion.

Like its formal counterpart, ICBT also suffers from poorly developed transportation, communication and other infrastructure network (Muyatwa, 2000; Little, 2010). Trade is characterized by lack of proper storage and warehouse facilities with poor access to market information (Bese *et al.*, 2009). The gravel road linking Zambia to DRC, in particular, has been in a bad state. Similarly, the road infrastructure in Rukwa region, linking Tanzania to both DRC and Zambia is very poor, with roads usually impassable during rainy season (Ackello-Ogutu & Echessah, 1998). This has implications for transportation cost and travel time. Zambia is landlocked and hence depends almost entirely on road networks for regional trade.

Further, ICBT is marred with several risk factors that have capacity to increase market costs. For instance, perpetual civil unrest, conflicts and political insecurity in DRC often creates an unsuitable business environment (FEWSNET, 2015). This depresses and at times halts cross

Stellenbosch University https://scholar.sun.ac.za

border trade activities as trader security concerns increase. Even in relatively secure border areas, traders are always under threat and at risk of confiscation of their goods by government officials (Little, 2010) or otherwise get charged exorbitantly through corruption.

However, owing to the fact that ICBT is perceived as more of a threat needing attention than a normal market pattern, policy efforts have been directed at formalizing this trade rather than improving its operating environment. In line with this, COMESA introduced a Simplified Trading Regime for selected types of goods under which consignments worth US$1,000 and less, can clear with little paperwork and without inspection by clearing agents. Nevertheless, utilization of this program has been minimal.

3.9 Conclusion

This chapter has provided a general overview of the dry bean subsector with a focus on Zambia, DRC, and Tanzania. The chapter has also discussed trade between Zambia and her neighbours DRC and Tanzania as well as the various aspects relating to informal cross border trading.

Following the analysis, it was seen that the bean trade is a growing activity at the global level. However, emerging trade patterns indicates growth in intra-regional trade. Although efforts have been made to promote formal bean trade among Eastern and Southern Africa countries, trade is still predominantly conducted through informal channels. Trends further indicate a growing pattern of informal trade. Based on the information in this chapter, a clear foundation has been laid on the importance of specifically investigating the operation of informal cross border dry bean markets in ESA. The next chapter, will discuss the methodology followed in assessing the interaction between the selected dry bean market pairs linked by informal trade.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 4: DATA AND ANALYTICAL PROCEDURE

4.1 Introduction

This chapter describes the data and analytical procedure followed in this study. The methodology was selected to achieve the goal of investigating the nature and extent of price transmission in two cross border dry bean market pairs, (1) Lubumbashi in the Democratic Republic of Congo (DRC) and Kitwe in Zambia; and (2) Mbeya in Tanzania and Kitwe in Zambia. It is organised as follows: the next section describes the study area and the data sources. Section 4.3 provides the analytical framework followed and the time series property tests conducted. Finally, section 4.5 specifies the main econometric model employed as well as the threshold selection procedure.

4.2 Study area

The study primarily focussed on Zambia while investigating market relationships with her eastern and southern neighbours Tanzania and DRC respectively. Figure 4.1 presents the locations of the markets included in the study. The studied markets were selected based on data availability, proximity to the borders and their importance as surplus or demand markets. In the case of Tanzania for instance, (as stated in sections 3.5 and 3.7), majority of beans is produced in the Southern Highland, Great Lakes and Arusha regions. However, the Southern Highland and Mbeya town in particular, is the major source of beans informally exported to Zambia (Bese *et al*., 2009). Hence, Mbeya market was employed as a surplus (source) market for Zambia's bean imports.

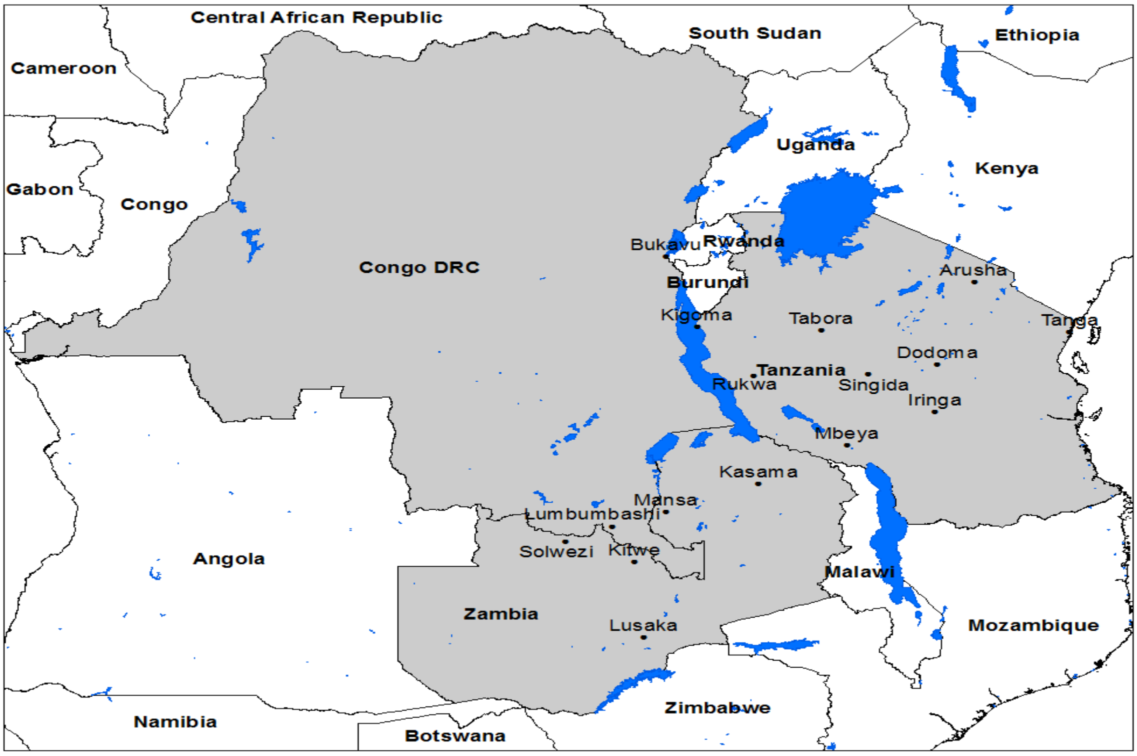


Figure 4.1: Map showing markets included in the study
Source: Adapted from google maps

Stellenbosch University https://scholar.sun.ac.za

On the other hand, Kitwe links Zambia's major bean producing regions to the DRC's deficit province, Katanga were Lubumbashi is the major high value market (refer to sections 3.5 and 5.3). Although not in itself among Zambia's top bean producing areas, it can be ranked the second most important destination for high value bean trade after Lusaka. Beans flow into DRC mostly through Kasumbalesa and Kalemie border points (FEWSNET, 2015). One can thus argue that a rational trader can only proceed to cross beans into DRC if prices in DRC are more attractive than those offered in Kitwe further suggesting that a relationship may exist between Lubumbashi and Kitwe bean prices. Lubumbashi and Kitwe were therefore studied as destination and source markets respectively.

4.3 Data sources and treatment

The study was based on secondary data and combined dry bean prices, trade volume and diesel price data obtained from several sources.

Dry bean prices: were obtained from specific national price reporting institutions in Zambia, the DRC and Tanzania. In particular, the Central Statistics Office (CSO) in Zambia supplied the Kitwe and Kasama market prices in Zambian Kwacha. Getting a good time series data for the DRC markets was however a challenge due to lapses in data collection in most markets. Nevertheless, complete data for the Lubumbashi market in Congolese Franc was available and was supplied by the Institute National de la Statistique (INS), in the DRC. Finally, price series for Mbeya market was obtained from the National Bureau of Statistics (NBS) in Tanzania.

All prices were converted from their local currencies to their US dollar equivalent using the monthly average exchange rates obtained from OANDA (2016). Price data used are monthly, retail prices for the period January, 2006 to June, 2016, with the period selected to match up the period for which informal trade data was available.

Informal trade volumes: The Famine Early Warning Systems Network (FEWSNET) in Zambia supplied data on informal cross border trade volumes. This data is collected by means of border monitoring. In particular, FEWSNET staff monitors placed at various border posts across Sub-Saharan African countries and beyond[9], take a daily count of the quantities of beans informally crossed between countries. Crossing into Zambia from Tanzania, 4 border posts, Lumi, Mpulungu, Nakonde and Zombe are currently monitored while Kasumbalesa and Kipushi are monitored for flows between Zambia and DRC. The recorded quantities (usually in small amounts) are then aggregated over a month for all monitored border posts and published in a monthly bulletin called "Informal Cross Border Food Trade in Southern Africa" for Southern Africa. This implies that trade flow values are national averages crossing the two countries rather than the volume flowing between specific markets as employed in Burke and Myers (2014).

As discussed in section 1.8 and 3.7, FEWSNET has in recent past attempted to capture formal trade flow. Statistics and data collected however indicate that such trade seldom occur between

[9] See Burke and Myers (2014) for a detailed outline of how FEWSNET collects and reports this data and the various challenges associated with capturing accurate data

Stellenbosch University https://scholar.sun.ac.za

Zambia and her neighbours. The study therefore exclusively employed informal trade data captured by FEWSNET in order to achieve the core objective of this research (to examine cross border integration under unregulated cross border trade).

Diesel data: Given the important role transfer costs play in market integration and the reality that such data are not only difficult to measure but also rarely available, the two datasets above were complemented with diesel price data employed as a proxy for transfer cost. The study utilised retail (per litre) prices for the importing market in each market pair following Ndibongo *et al.* (2010), Burke (2012) and Burke and Myers (2014). Prices for Tanzania were obtained from the Energy and Water Utilities Regulation Authority (EWURA) in Tanzania, while Zambian prices were supplied by the Energy Regulation Board (ERB) in Zambia. It is important to note that fuel prices in the two countries are regulated differently. While a common price is set across all towns regardless of location in Zambia, petroleum cap prices in Tanzania are revised monthly and are town specific. The study therefore used diesel prices specific to Mbeya district.

4.3.1 Data Analysis

The study employed a combination of Microsoft Excel 2016, Eviews 9 and Stata 14 to analyse this data. Generally, all analysis including threshold selection and final model estimation was done in Stata 14. However, Eviews 9 became handy in providing comparative results for time series property tests (unit root) and cointegration analysis. Finally, data cleaning, organization and descriptive statistics were done in Microsoft Excel 2016.

4.4 Analytical framework

The study employed the Myers and Jayne (2012) extension of the Threshold Autoregressive (TAR) and as reapplied by Burke (2012) and Burke and Myers (2014). This method follows through a series of steps as outline in Figure 4.2.

Stellenbosch University https://scholar.sun.ac.za

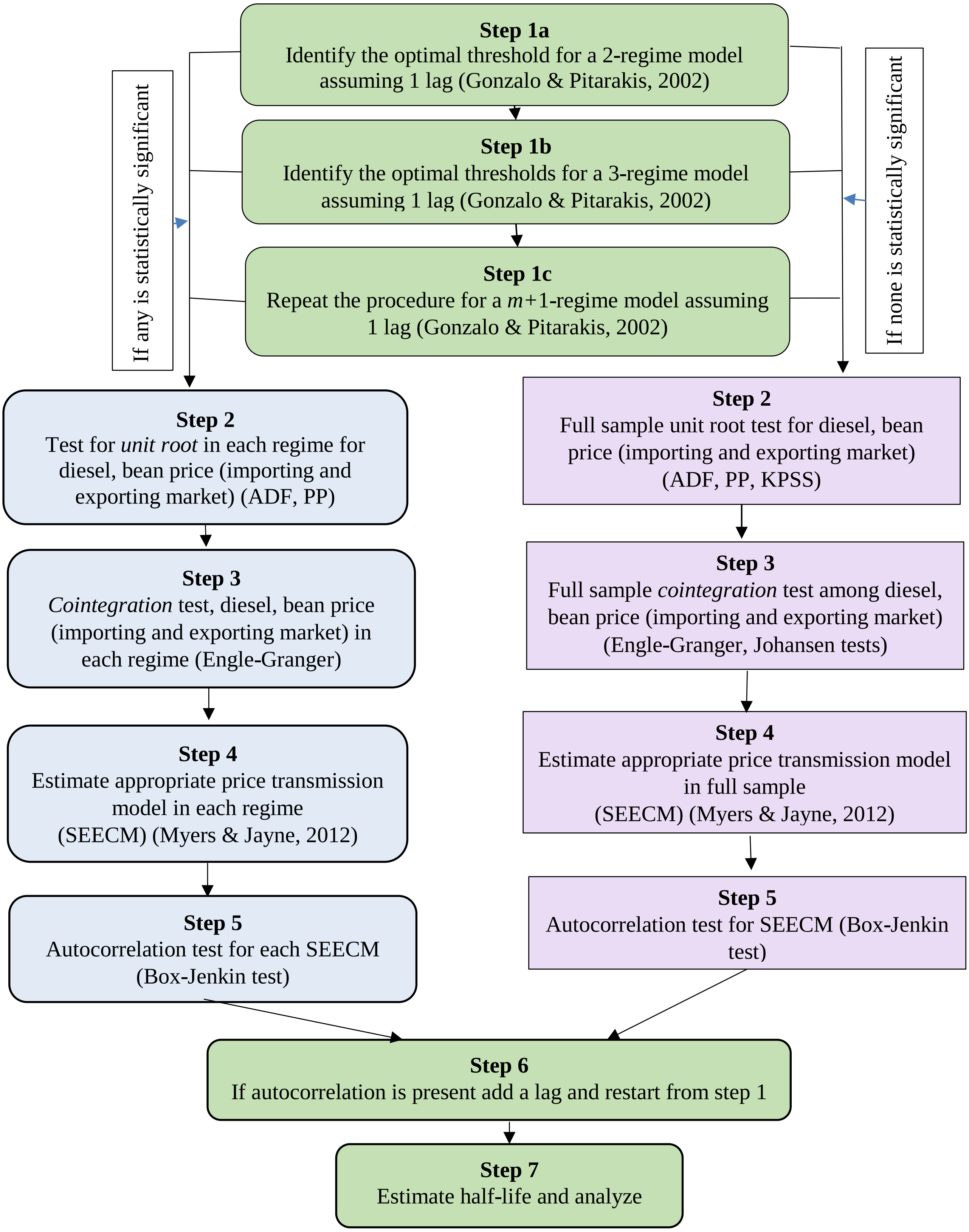


Figure 4.2: Multiple regime price transmission analytical framework
Source: Adapted from Burke (2012) and Myers and Jayne (2012)

Ideally, unit root testing should come first in market integration analysis. However, this step comes second in Figure 4.2 because the TAR model employed in this study requires that these

Stellenbosch University https://scholar.sun.ac.za

tests are also conducted in multiple regimes if trade based thresholds are statistically significant. Conducting these tests therefore required a conclusion on whether a model with trade based thresholds was optimal or not. Results in Chapter 5 are also presented as depicted in the graph.

Nevertheless, the next section explains how each of these steps was employed in this study beginning with step 2 (unit root tests), since outlining the threshold estimation procedure requires some information from the price transmission model. Step 1 is instead outlined after step 4.

4.4.1 Unit root tests (step 2)

Unit root tests were specifically conducted on bean price series for individual markets Kasama, Lubumbashi, Kitwe and Mbeya as well as diesel price series, in order to determine the level of integration[10] in each series.

The study employed three of the most widely used methods, the Augmented Dickey Fuller (ADF), Phillips-Perron (PP) and Kwiatkowski-Philips-Schmidt-Shin (KPSS). This follows a general trend in literature to validate unit root tests (see for instance Muyatwa, 2000, Ndibongo *et al*. 2010, Myers & Jayne, 2012) mainly because of their generally low statistical power to reject the null hypothesis (Gujarati, 2004). In case of conflicting results, a decision was made based on the sum of evidence.

4.4.1.1 Augmented Dickey-Fuller (ADF) test

Depending on the assumptions made about the underlying series, the ADF test can be computed in three functional forms. The general model is fitted as (Gujarati, 2004):

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \quad (4.1)$$

where ΔY_t is the first difference of the time series variable under investigation (Y_t), β_1 is the constant, β_2 is the coefficient of the linear time trend t, δ is the parameter of interest, Y_{t-1}, which is the lagged values of order one of Y_t, m is the number of lags to be included in the model, α_i is the coefficient of the ΔY_{t-i}, which represents the changes in the lagged values of (Y_t) and ε_t is the white noise error term. If Y_t follows a pure random walk[11], then β_1= 0 and β_2=0, reducing equation 4.1 to:

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \quad (4.2)$$

If, on the other hand, the series exhibits a random walk with a drift (constant), only the time trend drops (β_2=0), leaving equation 4.1 as:

[10] Level of integration refers to the number of times a series must be differenced in order to become stationary.

[11] A series is said to be a random walk if its mean value changes in a non-predictable pattern overtime.

Stellenbosch University https://scholar.sun.ac.za

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \qquad (4.3)$$

The full equation (4.1) is estimated if the series exhibits both a drift and a time trend (Gujarati, 2004).

This study employed two tests, the ADF with a constant (equation 4.3) and ADF test with both constant and a time trend (equation 4.1) to test for a unit root in all bean and diesel price series. The number of lags included for each series was based on the lag length at which no evidence of autocorrelation was found (this test is discussed in section 4.5.4).

The ADF test examined the null hypothesis that series (Y_t) contained a unit root (H_0:δ=0) against the alternative hypothesis of no unit root (H_1:δ<0). Given that the ADF test is a lower tailed test, if the test statistic automatically computed as:

$$Z_t = \frac{\hat{\delta}}{\psi_\delta} \qquad (4.4)$$

where ψ_δ is the standard error of $\hat{\delta}$, was greater than the critical values (based on Fuller, 1996), the study failed to reject the null hypothesis of a unit root and concluded that the underlying variable contained a unit root and hence was non-stationary. Alternatively, the MacKinnon (1994) approximate p-value was examined. The null hypothesis was rejected if the p-*v*alue was less than the level of significance. In addition, to determine the order of integration in non-stationary series, the procedure was repeated on differenced series. Rejecting the null hypothesis at first difference, meant the underlying series was integrated of order 1 or non-stationary (*I(1)*). These two decision rules do not conflict with each other.

4.4.1.2 The Phillips Perron (PP) test

The Phillips and Perron (1988) test for each series was based on the model equation (Gujarati, 2004):

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \varepsilon_t \qquad (4.5)$$

where ΔY_t is the first difference of the time series variable under investigation (Y_t), β_1 is the constant, β_2 is the coefficient of the linear time trend t, δ is the coefficient of Y_{t-1}, which is the lagged values of order one of Y_t, and ε_t is the white noise error term. Notice, however, that this equation is basically equation 4.1 less the changes in the lagged values of (Y_t) ($\sum_{i=1}^{m} \alpha_i \Delta Y_{t-i}$). This is the only modification the ADF test has made to the standard DF test. The PP test therefore is basically a modified DF test. Instead of using lags as in the ADF test however, it employs non-parametric statistical methods to control for autocorrelation in the error term (Gujarati, 2004).

The three functional forms described in the ADF test apply in the PP test. Each series was therefore tested under the two functional forms selected for the ADF test namely, computing equation 4.5 without the trend (*t*=0) term and one with a constant and trend (basically equation 4.5 as it is). The study tested the null hypothesis that the variable, Y_t, contained a unit root (H_0:δ=0) against the alternative hypothesis that the variable, Y_t, did not contain a unit root

Stellenbosch University https://scholar.sun.ac.za

(H_1:δ<0). If the calculated test statistic was greater than critical values, the study failed to reject the null hypothesis, meaning that series was non-stationary. In general, the rest of the decision rule remain as explained in the ADF test above.

However, the PP and ADF tests are both unit root tests procedures (tests which have the unit root under the null hypothesis), criticized for their low statistical power to reject the null hypothesis (Kwiatkowski *et al.*, 1992; Gujarati, 2004). Conclusions drawn are also usually similar.

4.4.1.3 Kwiatkowski, Phillips Schmidt and Shin (KPSS) test

The study therefore included the KPSS procedure, which tests the null hypothesis of stationarity (variable does not contain a unit root) against the alternative of non-stationarity. The test, proposed by Kwiatkowski *et al.* (1992), is a langrage multiplier procedure for testing $\sigma_u^2 = 0$. The analysis in this study was based on the model equation (Kwiatkowski *et al.*, 1992):

$$y_t = \xi t + r_t + \varepsilon_t \quad (4.6)$$

$$\text{where } r_t = r_{t-1} + u_t,\, u_t \sim iid(0, \sigma_u^2) \quad (4.7)$$

denotes a pure random walk, y_t is the time series variable under investigation, ξ is the coefficient of t, t is a parameter containing a deterministic term, that is either a constant or constant and trend and ε_t is error term assumed to be stationary and may be heteroskedastic. Like the ADF and the PP test, the KPSS test (equation 4.6) was computed with a constant only and with a constant and trend.

The study tested the null hypothesis that the time series variable (y_t) was stationary ($H_0: \sigma_u^2 = 0$), against the alternative hypothesis of non-stationarity ($H_1: \sigma_u^2 > 0$), using the one-sided langrage multiplier test statistic estimated as (Kwiatkowski *et al.*, 1992):

$$LM = \sum_{t=1}^{T} \frac{S_t{}^2}{\hat{\sigma}_\varepsilon^2} \quad (4.8)$$

$$\text{where } s_t^2 = \sum_i^t \hat{\varepsilon}_i \quad (4.9)$$

$\hat{\sigma}_t^2$ is the estimator for the error variance and $\hat{\varepsilon}_t$ is the residual of a regression of y_t. The null hypothesis was rejected at specific significant levels if the KPSS test statistic (equation 4.8) was greater than the critical value, implying that the variable (y_t) was non-stationary.

Unfortunately, like the PP and the ADP, this test also suffers from the low power limitation. Nevertheless, the test is uniquely useful for confirmatory analysis in conflicting results between the ADF and the PP tests (Konya, 2004) and was used as such in this study. In case all the three tests contradicted, the decision was made based on the sum of evidence.

4.4.2 Cointegration analysis (step 3)

The next step, after unit root testing, was cointegration analysis. A review of methods used to measure market integration in Chapter 2 (this is discussed in details in section 2.5.2.4)

Stellenbosch University https://scholar.sun.ac.za

established the importance of cointegration analysis as a pre-test to market integration analysis. This step was therefore employed as such in this study.

The study used two of the most widely used methods, the Engle and Granger (1987) and the Johansen (1988) cointegration methods to test whether bean prices in the importing and exporting markets and diesel cost, shared a common long run relationship (cointegrated).

4.4.2.1 Engle-Granger test

The Engle and Granger (1987) test, hereafter EG, is a residual based unit root test. The study employed a three-step procedure, in which step 1 examined the unit root test results from section 4.4.1. This is because cointegration analysis is only valid if all series under study are integrated of order 1 (I(1)). If this condition was met, step 2 estimated residuals from the cointegration regression of the form (Brooks, 2008):

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t \qquad (4.10)$$

where y_t is the bean price variable for the importing market at time *t*, β_1 is the constant, x_{2t} is the bean price series for the exporting market at time *t*, x_{3t} is the diesel cost and $u_t$is the error term. The ADF test with constant only was then employed to test for the presence of a unit root in the estimated residuals.

The study tested the null hypothesis of no cointegration (among bean price series for the exporting and importing markets respectively and the diesel cost) $(H_0: \psi = 0)$ in each market pair, against the alternative hypothesis of cointegration $(H_1: \psi < 0)$. As mentioned earlier, the EG test is a residual unit root test and requires stationary residuals if cointegration is present. Therefore, rejecting the null hypothesis of no cointegration, which is simply equivalent to rejecting the null hypothesis of unit root in the residuals, implies that the two price variables, y_t and x_{2t}, and the diesel variable x_{3t} are cointegrated.

4.4.2.2 Johansen cointegration test

Results from the EG test were complemented with the Johansen (1998) cointegration approach in order to detect any possibility of more than one cointegrating relationship.

Unlike the residual based EG test, the Johansen (1988) approach uses a maximum likelihood estimator to detect multiple cointegrating relationships among the examined variables. Like the EG test, however, all series under study must be non-stationary (*I(1)*).

In order to test for cointegration, the study employed a vector error correction model (VECM) specified as (Johansen & Juselius, 1990; Muyatwa, 2000):

$$\Delta x_t = \Gamma_0 + \pi x_{t-p} + \Gamma_1 \Delta x_{t-1} + \cdots . + \Gamma_{p-1} \Delta x_{t-p+1} + V_t \qquad (4.11)$$

where $\Gamma_0 = A_0$

$$\pi = I - A_1 - A_2 - \cdots - A_p$$

Stellenbosch University https://scholar.sun.ac.za

$$\Gamma_1 = A_1 - I, \Gamma_2 = A_2 - \Gamma_1, \ldots, \Gamma_p = A_p - \Gamma_{p-1}$$

and Δx_t is the vector of the first difference of endogenous variables (x_t), (bean price series for the exporting and importing markets and the transfer cost variable), Δx_{t-1} are vectors of the lagged values of the first of difference values and V_t is the normally distributed error term, A_0 is the constant and $A_1 \ldots A_p$ are (*n* x *n)* coefficient matrices of unknown parameters. The parameter of interest in this test was the long run impact matrix, π, also known as the rank of the VECM matrix (Muyatwa, 2000). If none of the variables contained in x_t, as defined above are cointegrated, the matrix π returns a rank of zero, implying a lack of a long run relationship between the variables. On the other hand, a matrix π rank of one and above, signals the presence of cointegration in the components of x_t. The question therefore is how to determine the number of cointegrating vectors (rank) as estimated by π.

Johansen and Juselius (1990) proposed the two test statistics that were applied in this study:

The trace statistic calculated as (Muyatwa, 2000; Brooks, 2008):

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \quad (4.12)$$

and the maximum eigen statistic given as (Muyatwa, 2000; Brooks, 2008):

$$\lambda_{max}(r, r+1) = -T \ \ln(1 - \hat{\lambda}_{r+1}) \quad (4.13)$$

In both equations, r is the number of cointegrating vectors, T is the number of observations, "$\hat{\lambda}_i$ is the estimated value for the *i*th ordered eigenvalue from the π matrix" (Brooks, 2008, p 351).

The trace statistic investigates the null hypothesis that the components of x_t have less than or equal to r cointegrating vectors ($H_1: r \leq 1$) against the general alternative hypothesis that the number of cointegrating vectors are more than r. On the other hand, the maximal eigen statistic tests the null hypothesis of exactly r cointegrating vectors against the alternative of $r + 1$ (Brooks, 2008). The study rejected the null hypothesis in both tests in favor of the alternative hypothesis if the test static was greater than the critical value.

Ideally, this decision should be straight forward. However, the trace and the maximum eigen statistics most often yield conflicting results, especially in small samples, such as in this study. The study relied on the trace statistics as recommended by Lüutkepohl and Saikkonen (2001). Also, the study employed the information criterion approach to determine the appropriate number of lags for the VECM model.

In conclusion, all unit root tests and cointegration analysis were employed in a single regime (full sample data) in the absence of a significant threshold. If on the other hand, one or more

Stellenbosch University https://scholar.sun.ac.za

trade based thresholds (step 1 in figure 4.2) was significant, the analysis required implementing these tests in each regime[12].

4.5 Price transmission with trade based thresholds

The fourth step in the analysis measured spatial market integration with trade based thresholds in each market pairing. An in-depth review of various methods used to test for market integration presented in section 2.5 of Chapter 2, guided the selection of the appropriate model to implement in this study. As highlighted in the same chapter, market integration has been conceptualised from different perspectives by different people and hence a wide selection of analytical methods has been developed (refer to section 2.5 for details). The choice of the appropriate model in any given case, however, is determined by the underlying study objective but to a large extent data availability.

This study considered the strengths and weaknesses of the various methods discussed in Chapter 2, and settled for the multiple regime TAR model introduced by Myers and Jayne (2012). This method is a variant of the standard TAR model originally introduced by Balky and Fomby (1997). It is the most appropriate measure for this study since it explicitly incorporates trade volume, price, and transfer cost data in the market integration analysis as required for this study. Also, the Myers and Jayne (2012) approach allows for market integration dynamics to vary across potentially multiple trade regimes, a key objective of this study.

4.5.1 Multiple regime price transmission model (step 4)

The Myers and Jayne (2012) multiple regime TAR model is estimated in a Single Equation Error Correction Model (SEECM) of the general form (Burke, 2012; Burke & Myers, 2014):

$$\Delta p_t^A = \mu_i + \beta_{1i}\Delta p_t^B + \beta_{2i}\Delta k_t + \lambda_i(p_{t-1}^A - \beta_{1i}p_{t-1}^B - \beta_{2i}k_{t-1}) + \theta_{1i}p_{t-1}^B + \theta_{2i}k_{t-1} + \sum_{j=1}^{n} b_{ji}\left(\Delta p_{t-j}^A - \beta_{1i}\Delta p_{t-j}^B - \beta_{2i}\Delta k_{t-j}\right) + \rho_{1i}\,\Delta p_t^B + \sum_{j=1}^{n} c_{ji}\left(\Delta p_{t-j}^B\right) + \rho_{2i}\Delta k_t + \sum_{j=1}^{n} d_{ji}\,\Delta k_{t-j} + u_{it} \qquad (4.14)$$

where, Δ represents the first difference operator (e.g. $\Delta p_t^A = p_t^A - p_{t-1}^A$)

p_t^A is the bean price series for the importing market (A)

μ_i is the constant

p_t^B is the bean price series for the exporting market (B)

[12] This means testing for unit roots and cointegration in observations falling below and above the threshold values separately.

Stellenbosch University https://scholar.sun.ac.za

β_{1i} is the long run equilibrium relationship between bean prices in the two markets (A and B)

k_t is the transfer cost (diesel price)

λ is the speed of price transmission

β_{2i} is the long run relationship between bean prices in the importing market (A) and the transfer cost

$\theta_{1i}, \theta_{2i}, b_{ji}, c_{ji}, d_{ji}, \rho_{1i}$ and ρ_{2i} are functions of the parameters in the error structure

$p_{t-j}^A, p_{t-j}^B, k_{t-j}$ are lagged values of order *j* of p_t^A, p_t^B and k_t respectively

n is the number of lags

u_{it} is the error term

Equation 4.14 takes on various functional forms under appropriate restrictions (Myers & Jayne, 2012). These restrictions are based on stochastic properties of the data series under study. Table 4.1 presents the three modelling assumptions considered in this study.

If for instance, based on the unit root tests outline in section 4.4.1, price series in each market pair and transfer costs were found to be stationary in levels and p_t^K *and* k_t are assumed exogenous, then ρ_1 and ρ_2 were restricted to zero in which case equation 4.14 was reduced to a stationary model.

Table 4.1: Alternative modelling assumptions

Estimate	Stationary Variable	Non-stationary Variable	Exogenous Variable	Parameter Assumption	Model Name
1	None	p_t^A, p_t^B, k_t	None	$\theta_{1i} = \theta_{2i} = 0$	Cointegration
2	p_t^A, p_t^B, k_t	None	p_t^B, k_t	$\rho_{1i} = \rho_{2i} = 0$	Stationary
3	p_t^A	p_t^B, k_t	p_t^B	$\beta_{1i} = 0$	Partial Cointegration

Source**:** Burke (2012) and summaries from Ndibongo *et al.* (2010) and Myers and Jayne (2012)

If, on the other hand, the bean price series for both markets in a given pair and the transfer cost variable were non-stationary (*I(1)*), $\theta_{1i} = \theta_{2i} = 0$, reducing equation 4.14 to a cointegration model of the form:

$$\Delta p_t^A = \mu_i + \beta_{1i}\Delta p_t^B + \beta_{2i}\Delta k_t + \lambda_i(p_{t-1}^A - \beta_{1i}p_{t-1}^B - \beta_{2i}k_{t-1}) + \sum_{j=1}^n b_{ji}\left(\Delta p_{t-j}^A - \beta_{1i}\Delta p_{t-j}^B - \beta_{2i}\Delta k_{t-j}\right) + \rho_{1i}\,\Delta p_t^B + \sum_{j=1}^n c_{ji}\left(\Delta p_{t-j}^B\right) + \rho_{2i}\Delta k_t + \sum_{j=1}^n d_{ji}\,\Delta k_{t-j} + u_{it} \quad (4.15)$$

Stellenbosch University https://scholar.sun.ac.za

This model is sufficient to measure market integration. However, a key objective of this thesis was to investigate cross border dry bean market integration with trade based thresholds. This cannot be accomplished by a mere application of any model equation from 4.14. To incorporate trade based thresholds therefore, equation 4.14 is presented in a simplified form given as:

$$\Delta p_t^A = f(X_t, \theta_i) \text{ if } \tau_{i-1} < q_t \leq \tau_i \quad \text{for all } i\text{=(1...I)} \tag{4.16}$$

where X_t are explanatory variables in equation 4.14, q_t represents the threshold variable (inter-market dry bean trade volume in this study), τ_i are the threshold values, and θ_i denotes the relevant regime dependent parameter vectors (Burke & Myers, 2014). All parameters in this model remain constant in one trade regime (a sub sample of observations below, in between or above a trade based threshold value) but changes as regime switches. Also, the appropriate model (from Table 4.1) for estimation in each regime, is selected based on regime specific stochastic properties tests (unit roots and cointegration).

Finally, for robust results from equation 4.16, an appropriate number of lags, *n*, must be included in each estimated model. The study utilised the number of lags that eliminated autocorrelation in the residual of each estimated model (the autocorrelation test is described in section 4.5.3 below).

4.5.2 Selecting the number and location of thresholds (Step 1)

By now the measurement of market integration with trade based threshold would be complete by simply estimating the appropriate models with the form 4.16. Notice, however, that in both equations, the threshold variable (q_t) is known but the threshold value (τ_i) is still unknown. As defined in section 2.6, a *threshold variable* is a variable that triggers the regime switch in a TAR model (in this study this variable is informal trade volume) while a *threshold value* is a critical observation (level) of the threshold variable at which the equilibrium and market integration dynamics change (e.g. 50 tons). On the other hand, the *number of regimes* refer to the number of sub-samples into which a full sample data can be split into[13].

Various methods used to locate threshold values and determine the number of regimes discussed in Chapter 2, section 2.6, guided the choice of the method employed in this study. As highlighted in that section, economic theory does not specify the full model structure of TAR models and hence, determining threshold values requires a separate method.

This study considered the limitations and strengths of available methods and chose the Gonzalo and Pitarakis (2002) (hereafter GP) approach. This method is the most appropriate since it is able to detect multiple threshold values and hence multiple regimes, a core component of this study. Also, the use of the penalty term instead of test statistics in the GP approach helps avoid complications resulting from nuisance parameters (Mann, 2012).

[13] e.g. a TAR model without a threshold value has 1 regime, a model with 1 threshold value has 2 regimes (below and above the threshold) while a TAR model with 2 threshold values has 3 regimes and so on.

Stellenbosch University https://scholar.sun.ac.za

Threshold selection was based on optimizing the GP objective function estimated as (Burke & Myers, 2014):

$$Q_T(i-1) = \max_{\tau} \ln\left[\frac{SS_T}{SS_T(\tau)}\right] - \frac{g(T)}{T} K.(i-1) \quad (4.17)$$

where

$Q_T(i-1)$ is a model with $i-1$ thresholds

SS_T is the residual sum of squares for a model without thresholds (single regime)

$SS_T(\tau)$ is the residual sum of squares for a model from the i -regime model

K is the number of parameters for estimation in a given single regime model

$g(T)$ is the penalty function

The GP approach uses five alternative specification for the penalty function $g(T)$ in equation 4.17, used in hypothesis testing in this study. These are; the Akaike Information Criteria (AIC) where $g(T)$ is simply 2, the three Bayesian Information Criteria (BIC) distinguished as BIC ($\ln T$), BIC2 ($2\ln T$) and BIC3 ($3\ln T$) and the Hannah-Quinone (HQ) criteria with $g(T)$ as $2\ln\ln T$.

To employ the GP criterion function (equation 4.17), however, one needs to compute $SS_T(\tau)$ first. This was generated from the "cointegration" SEECM specified in equation 4.15. Even if this model was not the appropriate model for estimation in a particular regime, in a given market pair, the value of $SS_T(\tau)$ from any of the models described in table 4.1 is the same hence, any model can be used to generate $SS_T(\tau)$ (Burke, 2012; Myers & Jayne, 2012).

Myers and Jayne (2012) then present equation 4.17 in a format analogous to the BIC criterion function as:

$$I = arg \max_{0 \le I \le I^*} Q_T(I-1) \quad (4.18)$$

See Burke (2012) for the derivation of this equation, but note that $I*$ is the maximum number of regimes that the sample could be split into and $I-1$ is the number of thresholds to be estimated. $Q_T(I-1)$ is the GP criterion function presented in equation 4.17 as recommended by Gonzalo and Pitarakis (2002) (Burke & Myers, 2014).

In practice this method can be implemented to identify thresholds sequentially or jointly. Using the joint threshold estimation approach, the following steps were followed in this study (Burke, 2012):

Step 1: Find the optimal threshold value for a single threshold (2 regime) model, assuming an initial lag of 1.

Stellenbosch University https://scholar.sun.ac.za

Step 2: Disregard results from step 1 and find the optimal threshold values for a two-threshold (3 regime) model. The two thresholds represent the optimal threshold value for low and upper volumes traded respectively.

Step 3: Disregard results from step 1 and step 2 and find the optimal threshold values for a three-threshold (4 regime) model. This procedure continues with an addition of a potential threshold (*n*) in every subsequent test until the first rejection of the null hypothesis of the presence of an additional threshold in each subsample.

Step 4: Examine the GP criterion values and compare them across examined steps 1 to 3. The appropriate model is one with the highest GP criterion values.

The study tested the null hypothesis of a linear model, $(H_0\colon \tau = 0)$ in step 1, against the alternative hypothesis of a single threshold model $(H_0\colon \tau = 1)$. In step two, the analysis tested the null hypothesis of a linear model, $(H_0\colon \tau = 0)$ against the alternative hypothesis of a two-threshold model $(H_0\colon \tau = 2)$. If the GP criterion values (AIC, BIC, BIC2, BIC3, HQ) returned negative values in any step, the study failed to reject the null hypothesis of a linear model and concluded that that there was no evidence of informal trade based threshold effect. This ultimately implied estimating the appropriate SEECM selected from table 4.1, in a single regime (full sample). If on the other hand, the analysis in a given market pair returned positive values in step 1 for instance, the study rejected the null hypothesis of a linear model in favor of a model with one threshold (2 regime model). In case of conflicting evidence from criterion values, the study based the decision on BIC and to a very small extent BIC2 and BIC3 as recommended by Gonzalo and Pitarakis (2002).

Burke (2012) argues for the superiority of the joint GP approach in giving consistent results in very small samples. However, the only major difference between the sequential and joint estimation procedure is how they treat the previously identified thresholds. While the sequential procedure maintains the first GP optimal threshold (previous thresholds) while examining the possibility of the second threshold (subsequent thresholds), the joint approach disregards the information learnt in step 1 (the previous step) when identifying the possibility of two thresholds (subsequent thresholds).

This study employed the joint estimation approach but also compared results with the sequential procedure. In the sequential approach, the study eliminated the smallest and largest 10% of the values in selecting thresholds and each of the remaining 80% of the observation in each regime was considered a potential threshold value.

4.5.3 Testing for autocorrelation in SEECM (Step 6)

Step 6 in the analysis tested for autocorrelation in the residuals of estimated SEECM for each market pair in order to ensure robust results. The study used the Ljung-Box test and tested the null hypothesis of no autocorrelation up to order *m* against the alternative hypothesis of autocorrelation up to order *m*.

Stellenbosch University https://scholar.sun.ac.za

The test statistic as proposed by Ljung and Box (1978, cited by Gujarati, 2004) was computed as:

$$LB = n(n+2)\sum_{k=1}^{m}\frac{\tau_k^2}{n-k} \sim \chi_m^2 \qquad (4.19)$$

where n is the sample size, m is the highest order of autocorrelation for which to test and τ_k^2 is the k^{th} autocorrelation.

The null hypothesis was rejected if the p-value was less than the level of significance concluding that autocorrelation was present.

To correct for autocorrelation, lags were added to the model one by one until no evidence of autocorrelation was found. This lag length also applied to the SEECM used in locating thresholds.

In addition, this test was also applied to univariate series to determine the appropriate number of lags for the ADF test and the EG cointegration test.

4.5.4 Calculating half-life (Step 7)

Once the appropriate SEECM passed the autocorrelation test, the final step, 7, calculated half-life of a price shock for each market pair. Half-life is defined as time taken for a market to eliminate half of the deviations from long run equilibrium, caused by a shock. Based on information from the SEECM model (step 5), half-life was calculated as:

$$hl = \ln(0.5)/\ln(1+\lambda) \qquad (4.20)$$

where λ is the speed of price transmission value from the estimated SEECM model. Half-life calculation is also regime dependent. If the calculated half-life estimate fell below 3 months, it was concluded that dry bean price transmission between the observed markets was rapid. Values between 4 and 6 months, and above 6 months however, were considered as moderate and slow transmission respectively.

4.6 Conclusion

This chapter focused on the discussion of the data and empirical models utilised in the study. The study used a combination of dry bean price, diesel price and informal trade data to analyse cross border bean market integration in this study. The data was analysed using the Myers and Jayne (2012) threshold autoregressive model estimated in a single equation error correction model structure. However, prior to the analysis, a series of analysis were conducted beginning from, threshold estimation and location, using the Gonzalo and Pitarakis (2002) approach, Unit root tests and cointegration analysis. Results from all these procedures are reported in the next chapter.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 5: RESULTS AND DISCUSSIONS

5.1 Introduction

This chapter presents and discusses results for the empirical analysis of market integration, discussed in Chapter 4. The chapter first examines the full sample descriptively. Thereafter, each cross-border market pair is examined in its own sub-section, before summarising the results.

5.2 Price variability and trade flow in selected market pairs

Table 5.1 reports summary statistics for nominal bean prices, diesel cost and trade flow for the studied markets. Based on literature and trade flow data, Lubumbashi and Kasama are reported as net importing markets throughout this study while Mbeya and Kitwe are exporting markets.

Table 5.1: Data summary statistics

	Number of Observations	Mean	Median	Standard Deviation	Minimum	Maximum
Zambia-DRC						
Kitwe (Zambia) price (USD/kg)	126	1.62	1.67	0.31	1.07	2.35
Lubumbashi (DRC) price (USD/kg)	126	1.96	1.97	0.36	1.21	2.96
Diesel cost Zambia (USD/L)	126	1.43	1.47	0.33	0.72	2.46
Trade flow Zambia to DRC (MT)	126	248.74	228.83	160.42	12.00	1373.00
Zambia-Tanzania						
Mbeya (Tanzania) price (USD/kg)	126	0.86	0.84	0.17	0.54	1.56
Kasama (Zambia) price (USD/kg)	126	1.07	1.07	0.29	0.55	1.78
Diesel cost Tanzania (USD/L)	126	1.15	1.19	0.24	0.58	1.75
Trade flow Tanzania to Zambia (MT)	126	343.56	199.38	336.42	0.00	1071.32
Trade flow Zambia to Tanzania (MT)	126	18.08	11.58	21.58	0.00	111.18
Net trade Tanzania- Zambia (MT)	126	325.48	188.04	323.20	-2.00	1041.22

Source: Based on data collected by FESNET, Zambia, CSO, Zambia, NIS, DRC and NBS, Tanzania.

Across Lubumbashi and Kitwe, the highest nominal price was observed in the Lubumbashi market, as expected, fluctuating between a minimum of $1.21 per kg of beans to a high of $2.96 per kg. However, Lubumbashi prices had a larger standard deviation when compared to those of Kitwe. A similar pattern was also observed in the Mbeya Kasama market pairs. Nevertheless, statistics revealed a noticeable difference in nominal diesel prices across the two country pairs. Zambia's diesel costs over the study period was on average 1.24 times higher than that of Tanzania. This observation suggests that everything else being equal, it is more expensive to transport beans from Zambia to DRC than it is to move it into Zambia from Tanzania. The high diesel prices in Zambia can be attributed to the fact that unlike Tanzania, Zambia being landlocked, has no direct access to a sea port, a condition known for high transport costs. Despite this observation, data records show unidirectional flow between Zambia and DRC. The volumes fluctuated from a high of 1,373MT to a low of 12 MT during the study period. This is not surprising given the large bean deficit in the DRC (FEWSNET, 2015).

Stellenbosch University https://scholar.sun.ac.za

Trade flow between Zambia and Tanzania, on the other hand, is bidirectional. Nevertheless, there are very small amounts of beans flowing to Tanzania. The two country pairs may be expected to therefore portray similar market relationships. Additionally, it is important to also note that, although trade volumes flowing between Lubumbashi and Kitwe displayed a wider range in terms of minimum and maximum (Table 5.1), on average Tanzania's exports to Zambia (343.56) is 40% more than the volumes flowing to DRC from Zambia. In other words, there are more beans flowing between Tanzania and Zambia than between DRC and Zambia.

To augment the general description of data above, a graphical behaviour of nominal price series is depicted in Figure 5.1 while their seasonal variations are presented in Figures 5.2 and 5.3. As can be seen in Figure 5.1, the price data indicates that there are some spatial variations in the selected cross border markets as well as some seasonal variations. Prices generally appear to follow an upward trend although this pattern is clearer when the graph is depicted in their original currency as shown in Appendix B1.

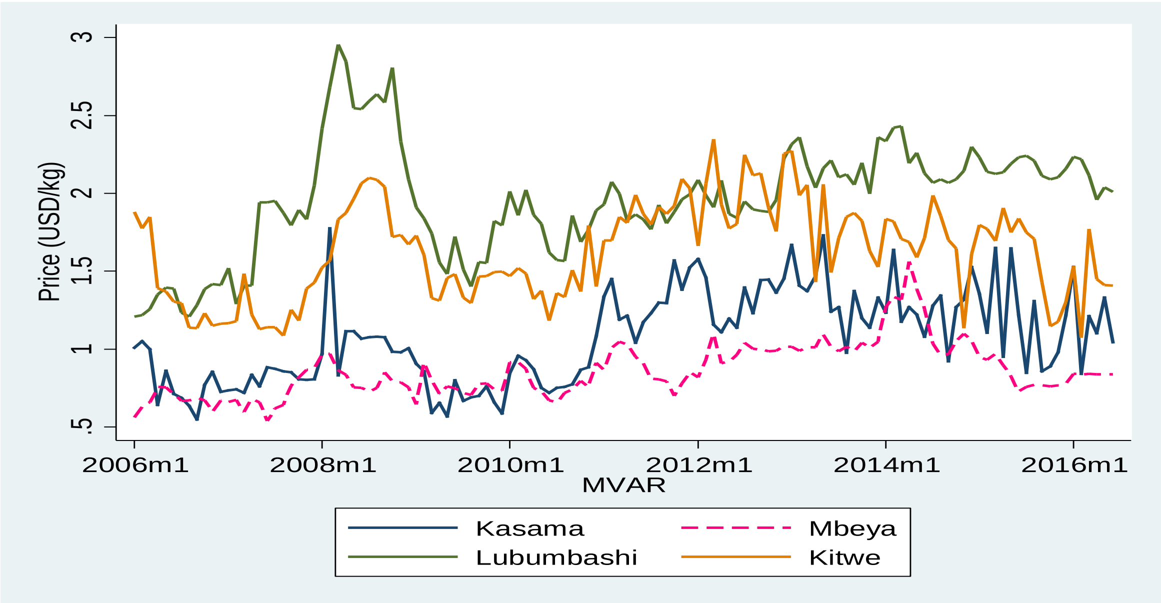


Figure 5.1: Nominal dry bean prices by market (2006-2016)
Source: Based on data collected by CSO, Zambia, NIS, DRC and NBS, Tanzania.

The seasonal variations depicted in Figure 5.2 suggest that prices in the studied markets are lowest in June, July and August. They start to rise above their annual average from November, reaches a peak in December to February and begin to fall below their annual average in April and May. This pattern is not surprising as they simply reflect normal patterns of agricultural price variability resulting from production timelines.

Stellenbosch University https://scholar.sun.ac.za

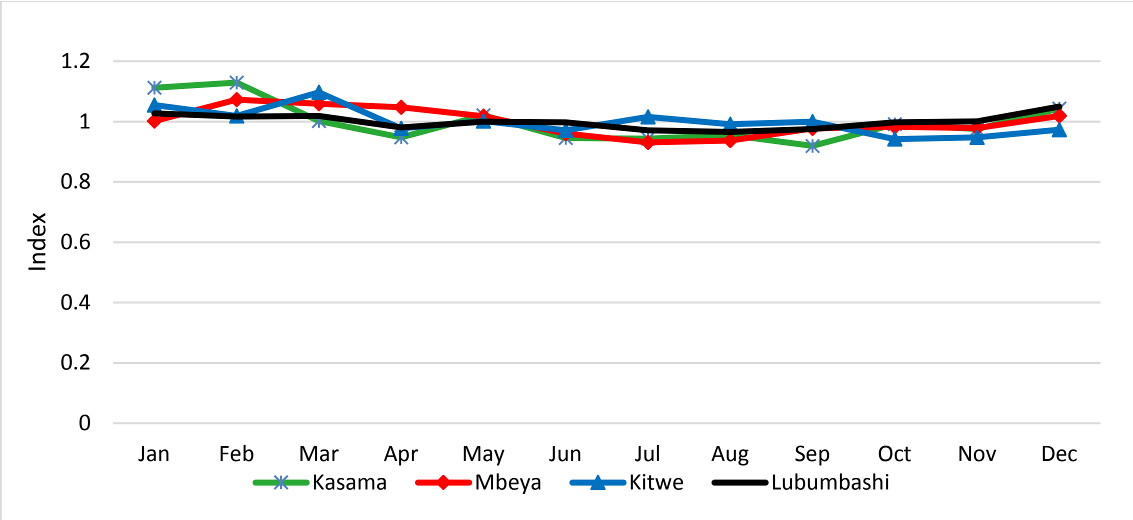


Figure 5.2: Seasonal indices of dry bean prices
Source: Based on data collected by CSO, Zambia, NIS, DRC and NBS, Tanzania.

Additionally, Figure 5.3 shows that the highest seasonal price observed between Kasama and Mbeya occurred in February, at 13 percent above the annual average price while the lowest was in July, at 8 % below the annual average price. Both prices were recorded in Kasama. On the other hand, seasonal variations were much lower across Kitwe and Lubumbashi. The highest price, on average 9.8% above the annual average price, was observed in March while October recorded the lowest price, 5.7% below the annual average price. Overall however, market price data indicates that bean prices are volatile. In the next section, the chapter examines the two market pairings separately and in detail.

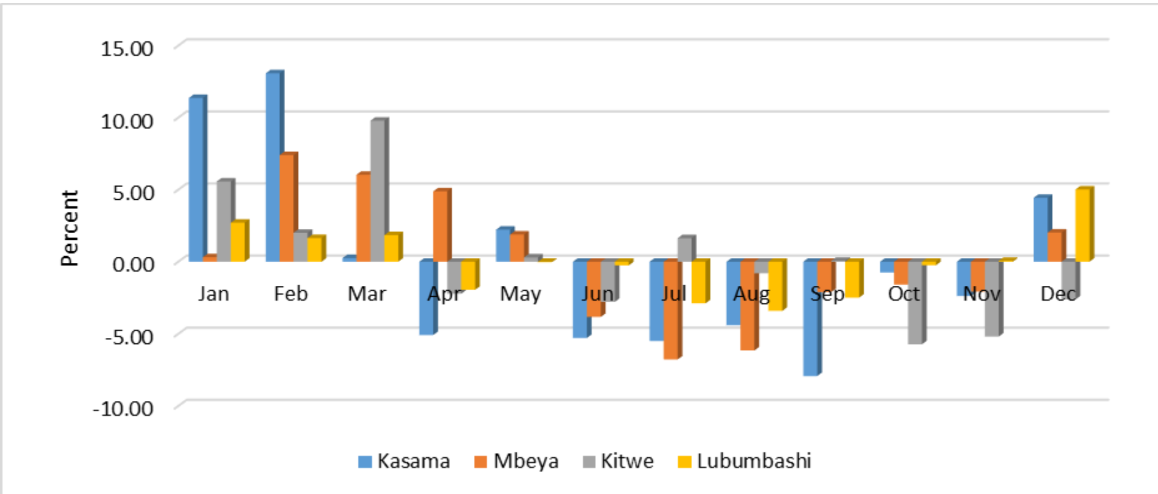


Figure 5.3: Percentage of average seasonal variations in bean prices
Source: Based on data collected by CSO, Zambia, INS, DRC and NBS, Tanzania.

5.3 Results for Kitwe, Zambia and Lubumbashi, DRC

Zambia is linked to DRC through one of her busiest borders, Kasumbalesa. The border is barely 98 km from Kitwe with an added distance of 90 km to reach Lubumbashi, the capital town of the mining province, Katanga in DRC. Lubumbashi is also DRC's second largest city after the

Stellenbosch University https://scholar.sun.ac.za

capital Kinshasa. Although Kitwe is not in itself among Zambia's top bean producing areas, it can be ranked the second most important destination for high value bean trade. Furthermore, the town links Zambia's major bean producing regions to the DRC markets. One can thus argue that a rational trader can only proceed to cross beans into DRC if prices in DRC are more attractive than those offered in Kitwe. This suggests that there may be a relationship between Lubumbashi and Kitwe bean markets. Everything else being equal (such as availability of sufficient bean inventory and perfect information), it is expected that more beans should flow to DRC whenever positive price differences between the two markets is largest.

5.3.1 Graphical relationship between spatial price difference and trade flow

Figure 5.3 illustrates the relationship between the quantities of beans informally exported to DRC and the price difference between Kitwe and Lubumbashi. Generally, price differences have been positive for the most part of the study period. The lowest gap was observed in the first quarter of 2006. From then, it rose to just slightly above zero before dropping and rising to a larger positive value towards the end of 2006. During this period, trade volumes appear to follow similar spikes as those in price differentials up until end of 2007. Then around January to April 2008, price differences hit their first highest increase for the study period. On the contrary, trade volumes decline. This could simply be missed arbitrage opportunity by traders given that this period coincides with Zambia's growing period when bean inventories are at their lowest. Trade volumes however picked up in May 2008, although they fluctuated and dropped even further around January to April 2009. In essence, price differentials do not appear to co-move with trade volumes for the most part of 2008 and 2009.

Towards the end of 2009, some co-movement begins to show but for the most part of 2010 the mixed pattern returns. Price differential drops below zero around May 2011. Surprisingly, trade volume fluctuates, but still maintains an average high even at negative price gaps. This may be explained by price data aggregation problems. Price data employed in this study is monthly averages while trade volumes are monthly totals. It is possible therefore that although the monthly average price recorded a negative price differential, there could have been days or weeks within these months when the price differential was positive, and hence created incentives for traders to export beans to Lubumbashi. Beyond data related justification, this observation may be explained by the fact that trade was conducted under imperfect price information. In the period from 2013 to 2016, the price gap is positively high. Trade appears to co-move with price gaps in some periods but not in other periods.

Overall however, it is difficult to ascertain a clear picture of the informal trade and price gap relationship graphically, as it generally appears mixed throughout the study period. To that end, the study employed statistical methods to establish whether informal trade levels have any effect on the long run price relationships between Lubumbashi and Kitwe.

Stellenbosch University https://scholar.sun.ac.za

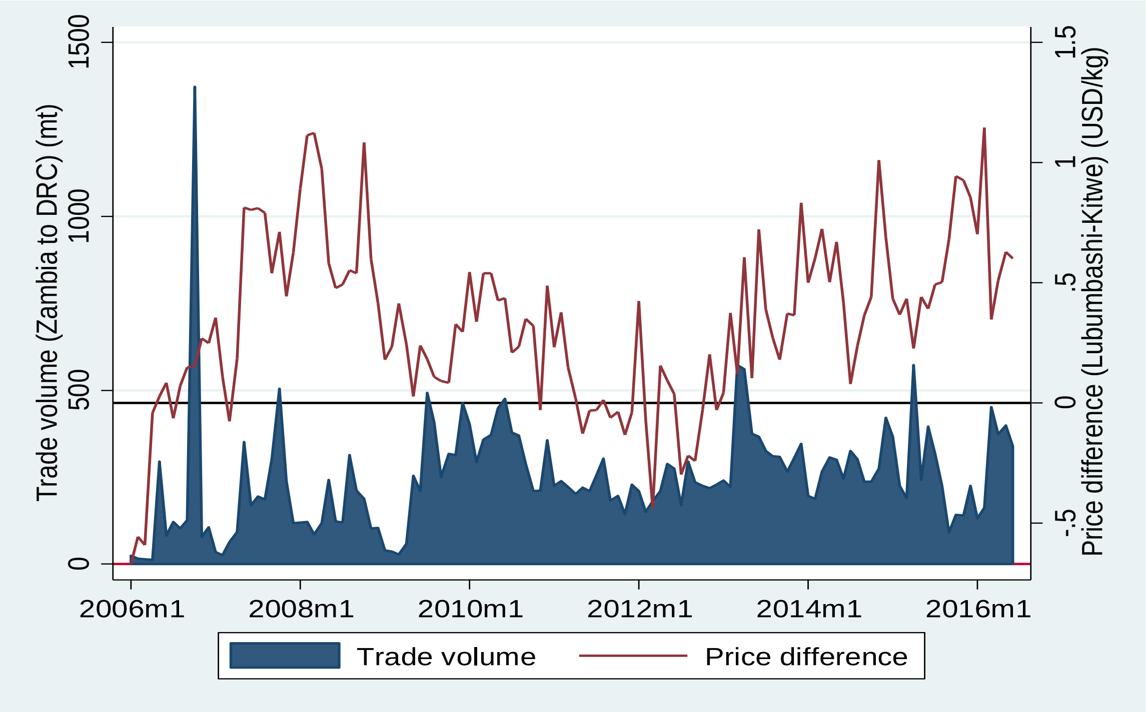


Figure 5.4: Informal trade and price difference between Lubumbashi and Kitwe
Source: Based on data collected by FEWSNET, Zambia, CSO, Zambia, INS, DRC and ERB, Zambia.

5.3.2 Threshold estimation and selection

The study specifically employed the Gonzalo and Pitarakis (2002) method as discussed in section 4.5.2. The GP values reported here, are estimates from equation 4.17 appropriately adjusted to estimate each value (GP) as specified in section 4.5.2. In addition, threshold selection utilised the SEECM model equation 4.15 and proceeded through the joint approach (non-sequential). There were 126 trade volume observations with each value a potential threshold.

Tables 5.2 reports the optimal threshold values selected and their GP criterion values.

Table 5.2: Jointly estimated threshold values (Kitwe and Lubumbashi)

	Threshold Value (MT of bear		Penalty Criterion Function (GP Values)				
Model	**Threshold 1**	**Threshold 2**	**BIC**	**AIC**	**HQ**	**BIC2**	**BIC3**
Two regime	121		-0.1688	0.0359	-0.0472	-0.5186	-0.8685
Three regime	295	359	-0.3559	0.0535	-0.1128	-1.0556	-1.7553

Source: Based on data collected by FEWSNET, Zambia, CSO, Zambia, INS, DRC and ERB, Zambia.

The first threshold defining two regimes was identified at 121 metric tons (Table 5.2). The AIC displayed a positive GP value of 0.04. This suggests that AIC rejects the null hypothesis of a linear SEECM model in favour of a model with one threshold. However, evidence from BIC

Stellenbosch University https://scholar.sun.ac.za

(-0.16), HQ (-0.05), BIC2(-0.52) and BIC3 (-0.87) all point to a stronger rejection of the null hypothesis of a threshold model in favour of a linear model, since they all return negative GP values. When the procedure is repeated to search for the possibility of a three regime (two threshold) model a similar pattern emerges. The AIC (0.05) still favour a model with two thresholds, suggesting that two trade based thresholds are in fact significant in influencing market integration dynamics between Lubumbashi and Kitwe. Once again, BIC (-0.36), HQ (-0.11), BIC2 (-1.06) and BIC3 (-1.76) returns negative values, indicating a failure to reject the null of a linear model (no thresholds) against the alternative hypothesis of a three-regime model (two thresholds)[14].

Based on the sum of evidence from Table 5.2 therefore, one can clearly conclude that the existence of an informal trade based threshold is rejected. Nevertheless, the study still turned to Gonzalo and Pitarakis (2002) for a more justified decision. In their Monte Carlo simulated results, Gonzalo and Pitarakis (2002) illustrate that on overall BIC followed by BIC2 criterions by far display the best performance in finite samples such as used in this study. In the same simulations, AIC and HQ criterions were found to incorrectly point to a threshold model more than 50% and 30% of the times respectively. Given that all BIC criterions unanimously favour the rejection of the null hypothesis of two regimes against a single regime, as well as the null of three regimes against a single regime, the study concluded that there is no evidence strong enough to support informal trade based threshold in the Lubumbashi Kitwe market pairing. This means that the quantities of beans traded informally between these markets, as well their fluctuations in volumes, do not significantly influence price transmission between the two markets.

An attempt was also made to estimate thresholds sequentially. As stated in section 4.5.2, the only major difference between the sequential and joint estimation procedures is how they treat the previously identified thresholds. While the sequential procedure maintains all previously identified thresholds and build on them when examining the possibility of subsequent thresholds, the joint approach disregards all information learnt from the previous step when investigating the possibility of subsequent thresholds. Results from the GP sequential procedure are presented in Table 5.3. They too strongly support the rejection of the presence of informal trade based thresholds (i.e. GP BIC values are all negative) in favour of a linear model. Based on this evidence, the analysis concluded that the appropriate model to apply in measuring market integration between Lubumbashi and Kitwe, was a model without informal trade based thresholds (linear model). As such, the analysis proceeded by estimating a linear SEECM model.

---

[14] Notice the anomaly, however, from Table 5.2. The three-regime model selects 295 MT and 359 MT as the first and second optimal thresholds respectively and yet none of these thresholds was identified under the one threshold model, which selected 121 MT as the optimal threshold. Burke (2016) confirms that this is not a modelling problem but a common case in small samples. The author further adds that the modelling assumption for the one threshold and two threshold models are different and hence there is no reason for them to find the exact same results.

Stellenbosch University https://scholar.sun.ac.za

Table 5.3: Sequentially estimated threshold values (Kitwe and Lubumbashi)

Variable	**Threshold value (MT of beans)**	**No. of obs. (below threshold)**	**No. of obs. (above threshold)**	**GP BIC value**
First threshold	121	26	100	-0.1708
Second threshold (over q > 121)	359.0	78	22	-0.4140
Third threshold (over 121 < q <= 359.0)	275.2	54	24	-1.1875

Source: Based on data collected by FEWSNET, Zambia, CSO, Zambia, INS, DRC and ERB, Zambia.

5.3.3 Unit root tests

Before estimating the model, however, the procedure (outlined in Figure 4.2) requires that the optimal model to be estimated is also selected based on stochastic properties of data series under study. This is because the SEECM model presented in equation 4.14, is a general model that can either be estimated as a "co-integration", "stationary" or a "partial co-integration" model (assumptions for these models are indicated in Table 4.1), depending on the stochastic properties of the Kitwe and Lubumbashi bean price series and the diesel cost series. The study used the ADF, PP and KPSS tests specified in equations 4.1, 4.5 and 4.6 respectively and each series was tested under two conditions, first with a constant and second with a constant and time trend. Results are summarised in Table 5.4.

For the Kitwe bean prices, the ADF test fails to reject the null hypothesis that series contain a unit root with or without adding a trend (Table 5.4). This means Kitwe bean prices are non-stationary. The PP test results, on the contrary, rejects the non-stationary null hypothesis suggesting that the series is in fact stationary. To break the contradictory tie, the KPSS test was employed. Unlike ADF and PP, KPSS tests the null hypothesis that series are stationary (*i.e.* does not contain a unit root). The test results support the ADF results rejecting the stationary and trend stationary null hypothesis at 10% and 2.5% significance levels respectively. In addition, all the three tests indicate that Kitwe bean price series are stationary at first difference. The study therefore concluded Kitwe bean price series are integrated of order 1.

For Lubumbashi, the results are far more mixed and non-mutually reinforcing. The ADF and PP tests without including a trend rejects the null hypothesis that Lubumbashi bean price series is non-stationary at 5% and 7% significance levels respectively. But when a trend is included, the same tests strongly fail to reject the non-stationary null hypothesis. KPSS too rejects the stationary null and fails to reject the trend stationarity null hypothesis. Literature suggest the ADF test gives superior results compared to the PP test in finite samples (Gujarati, 2004) and KPSS offers confirmatory results in case of conflicting results (Konya, 2004). Even under this rule however, the conclusion is still not definite, but since the ADF test without a trend term, only rejects non-stationarity at 7%, the study considered Lubumbashi bean prices as non-stationary.

For the diesel price series, five of the six tests suggest diesel prices are non-stationary. Particularly, the ADF and PP unit root tests fail to reject the null hypothesis that diesel prices in levels, contains a unit root with or without a trend. KPSS also rejects the stationary null but

Stellenbosch University https://scholar.sun.ac.za

fails to reject the trend stationary null hypothesis. Also, all tests fail to reject the hypothesis that diesel price series are stationary at first difference. Diesel price series are therefore no-stationary (I(1)).

Table 5.4: Unit root test results (Kitwe and Lubumbashi)

	In levels			At first difference		
Test	**Kitwe**	**Lubumbashi**	**Diesel**	**Kitwe**	**Lubumbashi**	**Diesel**
Unit root (NH: Non-stationary						
ADF	0.1043***	0.0709	0.1017***	0.00	0.00	0.00
ADF trend	0.3086***	0.2418***	0.1673***	0.00	0.00	0.00
PP	0.0014	0.0469	0.2248***	0.00	0.00	0.00
PP trend	0.0024	0.1709***	0.3473***	0.00	0.00	0.00
Unit root (NH: Stationary						
KPSS	<0.10*	<0.10*	<0.101	<0.101	<0.101	<0.101
KPSS trend	<0.03**	<0.101	<0.10*	<0.101	<0.101	<0.101
Cointegration						
Engle-Granger		0.0282				

Source: Based on data collected by CSO, Zambia, INS, DRC and ERB, Zambia.

Note: *, ** and *** denotes series is non-stationary at 10%, 5% and 1% respectively. The values presented for ADF and PP are p-values from MacKinnon (1994). P values for KPSS are based on Kwiatkowski *et al.* (1992).

5.3.4 Cointegration analysis

Having established that series are non-stationary and integrated of the same order (I(1)), Engle-Granger two step analysis was conducted and reveal evidence of a co-integrating relationship (Table 5.4). The ADF test (equation 4.3) results based on residuals from the co-integrating regression, equation 4.10, returned a p-value of 0.0282, indicating that the null hypothesis of no cointegrating relationship among bean prices in Kitwe, bean prices in Lubumbashi and the diesel prices, should be rejected in favour of the alternative hypothesis of cointegration.

In the Johansen test, cointegration was tested based on two test statistics trace and maximum eigenvalue, specified in equations 4.12 and 4.13. The optimal number of lags for estimating the VECM specified in equation 4.11 was 1 and 2, determined by the BIC and HQ, and AIC criteria respectively. Results are reported in Table 5.5. They suggest that at both lag lengths, the Maximum Eigenvalue statistics returns a rank of zero while trace statistics return a matrix rank of 1 cointegrating relationship when VECM is estimated with one lag, and zero under VECM with two lags. This implies that, based on maximum eigenvalue as well as the trace statistic for the two lag model, bean prices in Kitwe and Lubumbashi, and diesel prices are not cointegrated. However, the trace test at 1 lag indicates one cointegrating relationship exists. Following, Lüutkepohl and Saikkonen (2001) who recommend relying on trace statistics under such contradictions and recognizing that the trace statistic indicates at least one cointegrating relationship at one lag, bean prices in Kitwe and Lubumbashi, and diesel prices are concluded to be cointegrated and hence move together in the long run.

Stellenbosch University https://scholar.sun.ac.za

Table 5.5: Johansen cointegration results (Kitwe and Kasumbalesa)

	VECM with one lag			VECM with two lags		
Null hypothesis	**Maximum Eigenvalue (λ_{max})**	**Trace statistic (λ_{trace})**	**5% Critical value**	**Maximum Eigenvalue (λ_{max})**	**Trace statistic ((λ_{trace})**	**5% Critical value**
r = 0	-	39.52	29.68	-	25.67*	29.68
r = 1	0.21	10.76*	15.41	0.11	11.32	15.41
r = 2	0.07	2.20	3.76	0.06	4.31	3.76
r = 3	0.02	-	-	0.03	-	-

Source: Based on data collected by CSO, Zambia, INS, DRC and ERB, Zambia.

Note: * denotes rejection of null hypothesis at 5% significance level

Overall, test results appear to suggest Lubumbashi and Kitwe bean prices are non-stationary in levels and that there is some evidence of cointegration among them. The study therefore concluded that the "cointegration" SEECM model from Table 4.1 (characterised by non-stationary and cointegrated series), is the appropriate model to be estimated in the Lubumbashi and Kitwe market pair. The absence of statistically significant evidence to support trade based thresholds between these markets further implied that a linear SEECM model was the optimal model for estimation.

5.3.5 Price transmission estimation results

As discussed in Chapter 4, this is the main model applied to analyse integration between Lubumbashi and Kitwe. Analysis was carried out by estimating equation 4.15. The model had nine parameters to estimate and 126 observations. Prior to its estimation, however, diagnostic tests involving autocorrelation was conducted using the Ljung-Box test specified in equation 4.19. The test found evidence of autocorrelation at one lag, which was sufficiently eliminated with an addition of one more lag as indicated in the test results at the bottom of Table 5.6. The parameter estimate results reported in Table 5.6 therefore, are based on a linear SEECM model equation 4.15 estimated with two lags.

The results indicate the rejection of the null hypothesis of a significant long run price equilibrium relationship between Lubumbashi and Kitwe. In this regard, the coefficient estimate for the long run relationship between the two-price series (β_1), is 0.24 against the null hypothesis that β_1 is 0. Unexpectedly, its p-value is 0.59, indicating that β_1 is not statistically significant at any meaningful significance level. This suggests that there is no evidence of a long run price relationship between bean prices in Kitwe market and those of Lubumbashi. This unexpected finding partially explains the mixed results displayed by Johansen cointegration test above (Table 5.5). The results could be due to the several bottlenecks relating to the operating environment of these markets. In particular, a good portion of the road connecting Kitwe to Lubumbashi is gravel and in poor state as discussed in section 3.6 and could be hindering effective price transmission between the two markets. Also, trader insecurities caused by perpetual civil unrest in the DRC could also be limiting spatial arbitrage, leaving DRC markets isolated from the Zambian market influence.

Stellenbosch University https://scholar.sun.ac.za

Table 5.6: Parameter estimates of price transmission model for Kitwe and Lubumbashi

Parameter	Value	Std. error	95% confidence interval	
α: Constant	0.151	0.087	[-0.021	0.322]
β_1: Long run price relationship	0.243	0.449	[-0.645	1.132]
β_2: Relationship with diesel cost	0.275	0.390	[-0.489	1.047]
λ: Speed of adjustment	-0.121***	0.045	[-0.209	-0.032]
b_1	-0.015	0.092	[-0.197	0.167]
ρ_1	-0.170	0.431	[-1.024	0.684]
c_1	0.033	0.071	[-0.107	0.173]
ρ_2	-0.076	0.386	[-0.84	0.688]
d_1	0.136	0.124	[-0.109	0.381]
Half-life (Months)	5.37			
Goodness of fit				
R^2	0.13			
Adjusted R^2	0.07			
Residual autocorrelation				
Q(1)	0.99			
Q(5)	0.33			
Q(10)	0.44			

Source: Based on data collected by CSO, Zambia, INS, DRC and ERB, Zambia.

Note: *, ** and ***denotes statistically significance at 10%, 5% and 1% respectively. Residual autocorrelation (Q(j)) results are p-values from the Ljung-Box test of no residual autocorrelation against the alternative hypothesis of residual autocorrelation. Insignificant results imply no autocorrelation.

The second insight regards the extent of inter-market price influence. The speed of price transmission parameter (λ) estimate is -0.121, with a p-value of 0.008 and 95% confidence interval of -0.209 and -0.03. This suggests λ is statistically significant. The coefficient estimate translates into a half-life of 5.37 months as estimated using equation 4.20, with a 95% confidence interval of 2.95 months to 21.01 months. The results suggest that any price shocks from Kitwe to the Lubumbashi market die out in about 5.4 months. Prices then adjust back to their long run equilibrium.

However, although the parameter estimate β_2 has a correct sign with a value of 0.28, the p-value was found to be 0.48, suggesting that diesel prices are statistically insignificant in explaining the existing price relationship between Lubumbashi and Kitwe. This surprising finding contradicts with conventional economic theory, which suggests that transfer costs have a bearing on spatial price transmission (see for instance Baulch, 1997a). Burke and Myers (2014) made a similar discovery in their analysis of maize price transmission between Kitwe, Zambia and Kasumbalesa, DRC. The authors argue this could be due to the short distance between markets which implies that changes in diesel prices do not have a large impact on the overall costs of trade. Alternatively, the common use of other means of transport such as bicycle, in informal cross border trade could justify this finding. These costs, not directly associated with the price of diesel, could explain the price differences between the two markets.

Stellenbosch University https://scholar.sun.ac.za

The remaining parameters in the table do not have any meaningful economic interpretation individually (Myers & Jayne, 2012). Overall, the results are consistent with Teka and Azeze (2002) and Teka *et al*. (1999). The next section examines the relationship between prices in Mbeya, Tanzania and Kasama, Zambia and compares the findings with the price transmission results estimated here.

5.4 Results for Kasama, Zambia and Mbeya, Tanzania

Mbeya located in the Southern Highlands of Tanzania, as discussed in Chapter 3, has favourable bean growing conditions and a surplus producer of main food staples even beyond beans (Bese *et al.,* 2009). It is among Tanzania's top bean producing areas and a key source of beans informally and formally exported to Zambia (Bese *et al.*, 2009). Kasama, on the other hand, is a rural town situated in the Northern province of Zambia. It lies between Tanzania's Southern Highland and Zambia's major consuming regions, copper belt and Lusaka. Beans from Tanzania crosses into Zambia through the Northern province. Travelling by road, Kasama is 348 km from Mbeya and 241 km from the Zambia-Tanzania border (Nakonde), while Lusaka is 1,125km away. Given that the majority of informal traders are small by nature, they are likely to utilise a more nearby market as long as spatial arbitrage opportunities exist. One would thus expect a relationship between prices in Mbeya and Kasama if markets are working well.

5.4.1 Graphical relationship between spatial price difference and trade flow

Figure 5.4 depicts the relationship between informal trade volumes and price differences between Mbeya and Kasama. Similar to the case with Kitwe and Lubumbashi, trade and price gap relationship is mixed throughout the study period. For instance, the most parts of 2006 to 2008 as well as 2012 Tanzania's trade flow to Zambia appears to comove with the price differentials. But around the first quarter of 2014, price difference hit their bottom lowest and below zero and yet trade flow to Zambia nearly hits its highest point. From 2015, a mixed picture is again observed.

Notice however that unlike the case with Zambia and DRC, informal trade flow between Tanzania and Zambia is bidirectional. Zambia occasionally informally crosses beans into Tanzania even though in comparison, the volumes are minute. This was observed in 65% of the study period. Basic economic theory suggests this is problematic since trade flow is based on spatial arbitrage. It is therefore not possible for such opportunity to arise at the same time between the same markets. However, Pugel (2012) justifies reverse trade also known as intra industry trade, a common phenomenon in international trade, on the basis of comparative advantage, product differentiation and consumer preferences. Dry beans as considered in this study is a mixture of several varieties with varying consumer preferences (Sichilima *et al.,* 2016). Some buyers may have strong preference for foreign varieties even when local varieties are available. Moreover, the seasonal difference in comparative advantage tied to Tanzania's ability to produce beans twice in a year (Katungi *et al.,* 2009), may create incentive for reverse trade.

Stellenbosch University https://scholar.sun.ac.za

Outside of economic theory, Burke (2012) and Burke and Myers (2014) justifies reverse trade on the basis of data aggregation problems. Price differences are monthly averages while trade volumes are monthly aggregates. It is possible therefore that trade flow could have switched on different dates or weeks within a given month. Alternatively, traders may have simply operated on imperfect information and price expectations resulting in reverse trade. That said, the chapter proceeds to examine statistically any possibility for informal trade based thresholds and later on measure price transmission.

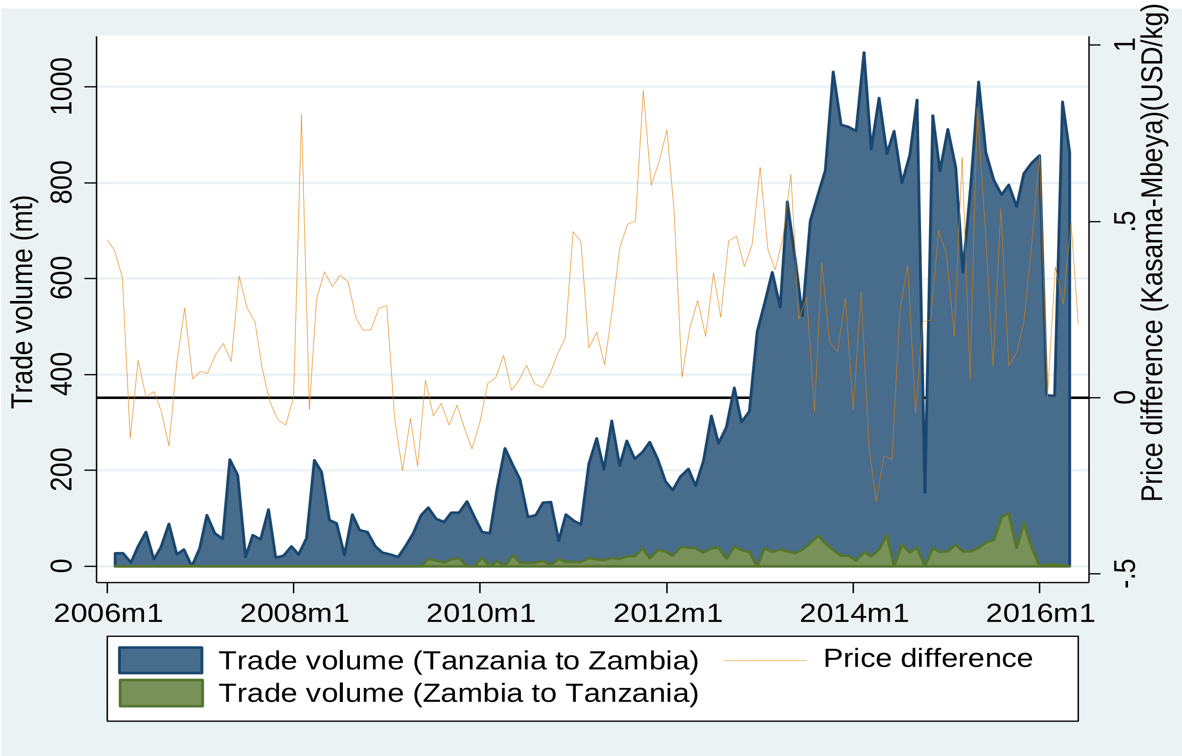


Figure 5.5: Informal trade and price difference between Mbeya and Kasama
Source: Based on data collected by FEWSNET, Zambia, CSO, Zambia and NBS Tanzania.

5.4.2 Threshold estimation and selection

Once again, unlike the case with Zambia and DRC, the presence of reverse trade presents a modelling problem when one needs to examine possibility of trade based thresholds. In fact, the application of threshold models with trade volume as a threshold variable has been mostly applied on markets with unidirectional trade flow (see for example, Ndibong *et al.*, 2010; Myers & Jayne, 2012; Stephens *et al.*, 2012). The question of which trade volume to use as a threshold variable, and much more of whether the model is appropriate for this case therefore becomes pertinent. The analysis in this market was based on net trade following Burke and Myers (2014)[15]. The model employed for Kitwe and Lubumbashi pair, is also appropriate for this case since the autocorrelation structure within the model (equation 4.14) allows for price

[15] The analysis also explored total trade as a possible threshold variable as in Burke (2012). But this led to the same conclusion regarding the evidence of thresholds (see results in Appendix B2).

Stellenbosch University https://scholar.sun.ac.za

transmission modelling without making any exogenous assumptions (Burke, 2012). The rest of the analysis thus proceeded as in the Zambia DRC case.

The SEECM model equation 4.15 was used for threshold selection and the GP values were calculated from equation 4.17 (appropriately adjusted to estimate each value (GP) as specified in section 4.5.2). Selection proceeded through the joint approach, searching through 126 net trade volume observations.

Tables 5.7 presents the optimal threshold values selected and their GP criterion values. The first threshold, defining two regimes is identified at 764.77 metric tonnes (Table 5.7). Unlike the Kitwe Lubumbashi case, HQ in addition to AIC returns positive GP values of 0.06 and 0.14 respectively, suggesting a rejection of the null hypothesis of a linear model in favour of a two threshold SEECM model. However, evidence from BIC (-0.06), BIC2 (-0.41) and BIC3 (-0.76) all point to the rejection of the null hypothesis of one threshold (two regime model) against the alternative of no threshold (single regime model). When the procedure is repeated to search for the possibility of a second threshold (three regime model) a similar pattern emerges. Both AIC (0.28) and HQ (0.12) support the presence of two statistically significant thresholds at 203.50 and 764.77 metric tons but BIC (-0.12), BIC2 (-0.82) and BIC3 (-1.52) rejects this conclusion.

Table 5.7: Jointly estimated threshold values estimated (Kasama and Mbeya)

	Threshold value		**Pernalty criterion function (GP values)**				
Optimal model	Threshold 1	Threshold 2	BIC	AIC	HQ	BIC2	BIC3
Two regime	764.66		-0.0598	0.1449	0.0618	-0.4096	-0.7595
Three regime	203.50	764.66	-0.1246	0.2848	0.1185	-0.8243	-1.5240

Source: Based on data collected by FEWSNET, Zambia, CSO, Zambia, EWURA, Tanzania and NBS, Tanzania.

Given that the BIC, as established in section 5.3.2, give superior results in finite samples, the study concluded that there was no evidence strong enough to support informal net trade based threshold in the Mbeya-Kasama market pair. Once again, attempts to estimate thresholds sequentially led to a similar conclusion regarding the presence of thresholds. Results reported in Table 5.8 still strongly support the rejection of the presence of trade based thresholds (*i.e.* GP BIC values are still negative). Given this finding, the analysis concluded that there was no evidence to support informal trade based thresholds between Kasama and Mbeya dry bean markets. A linear SEECM is therefore estimated in the next section.

Table 5.8: Sequentially estimated threshold values (Kasama and Mbeya)

	Threshold value	**No. of obs (below threshold)**	**No. of obs (above threshold)**	**GP BIC value**
First threshold	764.66	104	22	-0.1701
Second threshold (over q<=764.66)	203.50	72	31	-0.2155
Third threshold (q<=203.50)	136.37	56	15	-0.7809

Source: Based on data collected by CSO, Zambia, FEWSNET, Zambia, EWURA, Tanzania and NBS, Tanzania.

Stellenbosch University https://scholar.sun.ac.za

5.4.3 Unit root tests

Before estimating the model however, Mbeya and Kasama dry bean and diesel price series were tested for unit roots using the ADF, PP and KPSS tests specified in equations 4.1, 4.5 and 4.6 respectively. Each series was tested under two conditions, first with a constant and second with a constant and time trend. The results are summarised in Table 5.9.

For Mbeya and Kasama bean prices, the ADF test results for both a constant only and the addition of a time trend fail to reject the null hypothesis of a unit root in each series. The PP test on the other hand, weakly rejects the non-stationarity hypothesis in Mbeya price and strongly reject non-stationarity in the Kasama prices. To break the contradictory tie, KPSS test was employed. The test results support the ADF test and reject the stationary and trend stationary null hypothesis in both Mbeya and Kasama. In addition, all tests except KPSS test without a trend, fail to reject the hypothesis that diesel prices are non-stationary in levels but stationary at first difference. The three series are therefore non-stationary (I(1)) and the analysis proceeded to examine cointegration among them.

Table 5.9: Unit root test results (Kasama and Mbeya)

	In Levels			At First Difference		
Test	**Mbeya**	**Kasama**	**Diesel**	**Mbeya**	**Kasama**	**Diesel**
Unit root (NH: Non-stationary						
ADF	0.100***	0.1471***	0.2313***	0.00	0.00	0.00
ADF trend	0.147***	0.1421***	0.6006***	0.00	0.00	0.00
PP	0.0589	0.000	0.2131***	0.00	0.00	0.00
PP trend	0.0947	0.000	0.6133***	0.00	0.00	0.00
Unit root (NH: Stationary						
KPSS	<0.01***	<0.01***	<0.101	<0.101	<0.101	<0.101
KPSS trend	<0.10*	<0.05**	<0.10*	<0.101	<0.101	<0.101
Cointegration						
Engle-Granger		0.0206				

Source: Based on data collected by CSO, Zambia and NBS, Tanzania, EWURA, Tanzania.

Note: *,**,*** denotes series is non-stationary at 10%, 5% and 1% respectively. The values presented for ADF and PP are p-values from MacKinnon (1994). P-values for KPSS are based on Kwiatkowski *et al.* (1992).

5.4.4 Cointegration analysis

Cointegration analysis with the Engle-Granger test examined the stationarity of residuals from the cointegrating equation 4.10, using the ADF test specified in equation 4.3. The test results, reported in Table 5.9, indicate a p-value of 0.0206, suggesting that the residuals obtained from the cointegrating regression (Mbeya, Kasama and Diesel) are stationary. This implies the null hypothesis of no cointegration is rejected at 5% significant level and hence the three series are cointegrated.

Stellenbosch University https://scholar.sun.ac.za

In the Johansen test to examine possibility of more than one cointegrating relationship, cointegration was based on two test statistics trace and maximum eigenvalue, specified in equations 4.12 and 4.13. The optimal number of lags for estimating the VECM specified in equation 4.11 was 1 and 2, determined by the BIC and HQ, and AIC criterions respectively. The results presented in Table 5.10 supports the conclusion from the Engle granger test given that the trace statistic returns a maximum rank of at least 1 cointegrating relationship in both cases. The discussion is limited to the trace statistics since, as explained in section 5.3.4, the maximum eigenvalues are unreliable. The three series are therefore cointegrated and share a common long run relationship.

Table 5.10: Johansen cointegration test results (Mbeya and Kasama)

	VECM with one lag			VECM with two lags		
Null Hypothesis	**Maximum Eigenvalue** (λ_{max})	**Trace Statistic** (λ_{trace})	**5% Critical Value**	**Maximum Eigenvalue** (λ_{max})	**Trace statistic** (λ_{trace})	**5% Critical Value**
r = 0	-	52.35	29.68	-	29.95	29.68
r = 1	0.28	12.85*	15.41	0.14	12.21*	15.41
r = 2	0.08	3.22	3.75	0.07	3.49	3.76
r = 3	0.02	-	-	0.03	-	-

Source: Based on data collected by CSO, Zambia, EWURA, Tanzania and NBS, Tanzania.

Note: *indicate rejection of the null hypothesis at 5% significance level.

Overall, evidence from unit root tests and cointegration analysis suggests the "cointegration" SEECM model from Table 4.1, is the appropriate model for analysing the nature and extent of market integration between Mbeya and Kasama. This model was therefore estimated and is discussed in the next section.

5.4.5 Price transmission estimation result

As applied in the Kitwe Lubumbashi markets, analysis was based on estimating equation 4.15. The model had nine parameters to estimate and 126 observations. Prior to its estimation, the model was tested for the presence of autocorrelation using Ljung-Box test specified in equation 4.19. The test results presented at the bottom of Table 5.11, returned insignificant p-values, indicating that one lag was sufficient to eliminate residual autocorrelation. The parameter estimate results reported in Table 5.6 therefore, are based on a linear SEECM model equation 4.15 estimated with one lag.

Unlike the Kitwe Lubumbashi market pairing, there is a long run equilibrium relationship between Mbeya and Kasama dry bean prices. In this case the estimate of the long run price transmission parameter, β_1, was 0.899 against the null hypothesis that β_1 is 0 (*i.e.* Mbeya prices have no effect on Kasama prices). The p-value is 0.029, which indicates that Mbeya prices are statistically significant (at 5% level) in explaining Kasama prices. The estimated 95% confidence interval returns 0.09 and 1.71. These results are consistent with the evidence obtained from cointegration tests in Tables 5.9 and 5.10, which unanimously supported the evidence of a long run relationship between Mbeya and Kasama markets.

Stellenbosch University https://scholar.sun.ac.za

The speed of price transmission parameter (λ) is also statistically significant. Its estimated value is -0.331 with a p-value of 0.00. This translates into a half-life of 1.72 months. The 95% confidence interval is given by -0.51 and -0.15, indicating a half-life interval of 0.98 months to 4.12 months. The results suggest that any price shocks from Mbeya to the Kasama market die in about 1.7 months. Prices then adjust back to their long run equilibrium.

Table 5.11: Parameter estimates of price transmission model for Kasama and Mbeya

Parameter	**Value**	**Std. error**	**95% confidence interval**	
α: Constant	0.031	0.11	[-0.180	0.242]
β_1: Long run price relationship	0.899**	0.41	[0.092	1.171]
β_2: Relationship with diesel cost	0.177	0.29	[-0.406	0.759]
λ: Speed of adjustment	-0.331***	0.09	[-0.508	-0.155]
b_1	0.37***	0.09	[-0.543	-0.196]
ρ_1	-0.465	0.41	[-1.279	0.348]
c_1	0.308	0.32	[-0.941	0.324]
ρ_2	-0.378	0.35	[-1.070	0.315]
d_1	0.104	0.30	[-0.497	0.70]
Half-life (Months)	1.72			
Goodness of fit				
R^2	0.38			
Adjusted R^2	0.34			
Residual autocorrelation				
Q(1)	0.55			
Q(5)	0.55			
Q(10)	0.34			

Source: Based on data collected by CSO, Zambia, EWURA, Tanzania and NBS, Tanzania.

Note: *, ** and ***denotes statistically significant at 10%, 5% and 1% respectively. Residual autocorrelation results are p-values from the Ljung-Box test of no residual autocorrelation against the alternative hypothesis of residual autocorrelation. Insignificant results imply no autocorrelation.

However, the parameter estimate β_2 (relationship between the diesel cost and Kasama price) returns a positive value of 0.177. Surprisingly, the p-value is 0.55, suggesting that diesel prices are not statistically significant in explaining the long run price relationship between the two markets at any meaningful significance level. A similar observation was made in the Lubumbashi Kitwe analysis. It is quite difficult, at this point, to explain this finding given that Kasama is 348km from Mbeya. However, since diesel cost only reflects a proportion of transportation costs, other aspects of transfer costs such as the cost per unit are likely to influence the price relationship.

The discussion is once again limited to the above parameters since the remaining parameters do not have any significant economic interpretation individually (Myers & Jayne, 2012). These results are consistent with Burke (2012), Burke and Myers (2014).

Stellenbosch University https://scholar.sun.ac.za

5.5 Conclusion

This chapter focussed on presenting and discussing the results of the empirical procedures employed in investigating market integration in two cross border market pairs; (1) Kitwe, Zambia and Lubumbashi, DRC, (2) Kasama, Zambia and Mbeya, Tanzania. The chapter first examined the possibility of changing patterns of price transmission depending on the inter-market bean trade volumes. After a careful and thorough search for thresholds, the results showed no evidence to support informal trade based thresholds in both market pairs. This means, specific differences (small or large monthly volumes) in informal beans trade volumes between these markets do not significantly influence the extent to which price changes in one market influence price changes in another. From the analytical point of view, this invalidated the estimation of price transmission with in multiple regime for both pairs.

Preliminary analysis involving unit root tests and cointegration indicated that price and diesel series in each market was non-stationary and integrated. However, the single equation error correction model results for the Kitwe, Lubumbashi pair revealed that the two markets are segmented with a moderate speed of price transmission (5.37 months). On the contrary, Mbeya and Kasama are integrated with rapid price transmission.

In light of the above, this chapter concludes that variations in informal trade volumes, do not significantly influence the degree to which dry bean price changes in one market cause a change in another. The chapter also concludes that although a free cross border trading environment has great potential to drive well-functioning markets, it is not an outright guarantee that markets will be integrated.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.1 Introduction

This chapter presents conclusions of this study at four levels. The next section provides a summary of the thesis, procedures followed as well as the major findings drawn from the study. A comment on the questions and hypothesis set out in chapter one then follows. Section three discusses the implications of the main findings and makes recommendation relevant for policy makers. The chapter concludes by making suggestions useful for applied researchers focussing on market integration.

6.2 Thesis overview

Spatial market integration occurs when there is a smooth transfer of price signal across geographically separated markets. For an agricultural commodity like beans with spatial production concentration, spatial market integration is important to ensure a regional balance among deficit and surplus regions. Literature examining cross country market integration in Eastern and Southern Africa (ESA) suggest agricultural markets are relatively isolated from outside price changes partly because of restrictive and prohibitive trade policies and high transfer costs. Yet, little evidence has been gathered to examine how markets free from direct political influence may perform. The behavioural functioning of markets such as beans connected by informal cross border trade can therefore help shed light on the subject matter.

The main question that this thesis attempted to answer, therefore, was whether cross border dry bean markets and particularly those dominated by informal trade in ESA are integrated. To answer this question, the study examined two cross border market pairs, Kitwe in Zambia and Lubumbashi in DRC, and Kasama in Zambia and Mbeya in Tanzania.

The theoretical framework to direct the analysis was set out in Chapter 2, from which, the appropriate analytical model to apply in this study was selected. The study employed the Myers and Jayne (2012) extension of the Threshold Autoregressive (TAR) model, which as outlined in Chapter 4, was operationalised through a number of reinforcing steps.

Firstly, the relative importance of informal trade volumes on price transmission dynamics (e.g. speed, response time path) in each market pair, was investigated using the Gonzalo and Pitarikis (2002) approach. Although not the only method available for threshold identification, it is better suited to detect multiple thresholds and their specific values. After a careful and thorough search for thresholds, the results showed no evidence strong enough to support informal trade based thresholds in both market pairs. In this regard, all reliable GP values (BIC, BIC2 and BIC3) returned negative values. This means, specific differences (small or large monthly volumes) in informal bean trade volumes between these markets do not significantly influence the extent to which price changes in one market influence price changes in another. Although unexpected, the results suggest something positive, which is that the functioning of informal cross border markets is independent of exogenous limitation to trade such as heavy trade restrictions. From the analytical point of view, this finding invalidated the estimation of

Stellenbosch University https://scholar.sun.ac.za

price transmission model (SEECM) with trade based thresholds in each market pair. A linear model was therefore estimated.

Secondly, the existence of a long run price equilibrium relationship in each market pair was examined in order to determine whether markets were integrated or not. Some previous studies (for instance, Burke, 2012) have demonstrated that international markets relatively unimpeded by interventionist trade policies not only perform well but also that they perform better than their opposite counterparts. Preliminary analysis involving unit root tests and cointegration analysis indicated that all bean and diesel price series in each market pair were non-stationary at order 1 and that they were cointegrated. However, using the Myers and Jayne (2012) single equation error correction model (SEECM), the study found no evidence of a common long run price equilibrium relationship between Lubumbashi and Kitwe, suggesting that the two markets are not integrated. Any significant price deviations above transfer cost between Lubumbashi and Kitwe therefore may continue to grow without any tendency to equilibrium.

On the contrary, analysis found strong evidence to support the presence of a long run relationship between Mbeya and Kasama, implying that the two markets are integrated. This means that price changes in the Mbeya market are smoothly transferred to the Kasama market. It was further established that in the long run, every dollar unit ($1 per kg) increase in bean prices in Mbeya, stimulates a corresponding $0.90 increase in bean prices in Kasama.

In the final stage of the analysis, the speed of price transmission between each market pair was estimated based on the results from the SEECM. Considerable differences between the studied markets once again emerged. The long run equilibrium adjustment coefficient for the Lubumbashi and Kitwe market pair was statistically significant with an estimated half-life of 5.3 months. This finding suggests that 5.3 months are required for Lubumbashi to fully adjust to bean price changes in the Kitwe markets, clearly indicating moderately slow response. On the other hand, price transmission between Kasama and Mbeya is rapid. The estimated half-life is 1.7 month indicates that it takes Kasama only 1.7 months to fully adjust to bean price changes in Mbeya. The two informal trading markets are therefore linked by competitive arbitrage.

6.3 Reporting on the research questions and validating the hypothesis

Based on the findings above, the stated questions in chapter 1, “Is there a common long run price equilibrium relationship between each studied market pair?” “How long does it take for a price change (shock) in one market to be fully transmited to the other market in each pair?” and “ Do variations in inter-market informal trade flow volume have any effect on the market integration dynamics in each pair” have been adequately adressed.

The first stated hypothesis, “there is a long run price relationship between each selected market pair and therefore markets are well integrated with each other” is clearly rejected for the Lubumbashi and Kitwe pair and accepted as valid for the Kasama and Mbeya paring.

The second hypothesis, “price transmission between the selected markets is rapid. This is because exploiting arbitrage opportunities between markets is both time and cost saving given

Stellenbosch University https://scholar.sun.ac.za

that traders do not follow formal international trade procedures and government control" is also only accepted as valid in the Mbeya and Kasama bean market pair and rejected for Lubumbahsi and Kitwe.
The third hypothesis, "The level of intermarket informal trade has a significant effect on market integration" is rejected in both market pairs.

6.4 Implications and policy recommendations

The empirical results from this study suggests that even under minimal government intervention in trade and minimal transfer costs (relative to the costs of trading with countries not in border proximity), markets in ESA may not always perform as textbook economic theory predicts. This conclusion directly counters (oppose) the main conclusion reached by Burke (2012) and Burke and Myers (2014), in their analysis of Southern African cross border maize markets connected by informal trade. It also contradicts the common recommendations given in studies examining markets affected by trade policy regulations (see for example, Korir *et al.* 2003, Mauyo *et al.* 2007, Myers and Jayne, 2012, Ndibongo *et al*. 2010).

The results however concur with one of the conclusions reached by Burke and Myers (2014) namely, that markets with these characteristics if integrated, can be expected to have rapid price transmission between each other, with fairly competitive spatial arbitrage. It is important, at this point, to note that this is not the first time such markets have been found to fall short of their theoretical expectations in terms of integration. Livestock markets between Ethiopia and her neighbours Djibouti and Somali (Teka and Azeze, 2002), and Ethiopia and Kenya (Teka *et al.*, 1999), with cross border trade virtually uncontrolled by government, displayed very weak integration and no integration respectively.

That said, the results have two stories to tell and hence warrant two sided recommendations. Firstly, the high degree of integration between Mbeya and Kasama confirm the potential of politically unimpeded markets to perform better than their trade restricted counterparts. An agricultural marketing and regional trade policy that focus on limiting cross border market controls should therefore be promoted. A policy of liberal regional trade will most likely stimulate integration in cross border markets which in turn, is likely to facilitate regional food deficit and food surplus balance. This recommendation directly speaks to the ongoing efforts to open regional trade in ESA through creation of free trade areas both under SADC, COMESA and the Tripartite FTA. However, beyond a policy focus on eliminating tariffs, policies aimed at reducing the time consuming and cumbersome procedures in formal trade should be promoted.

On the other hand, the lack of market integration in informal trading markets such as reflected between Lubumbashi and Kitwe, can be due to a number of bottlenecks related to non-price factors. This study did not establish empirically the exact bottlenecks behind this finding, but safely ruled out government intervention and trade flow levels. However, other studies (for instance Muyatwa, 2000; Amikuzuno & Donkoh, 2012) have shown that market operating environment aspects such as storage, transportation and road infrastructure, access to market

Stellenbosch University https://scholar.sun.ac.za

information, lack of finance can greatly hamper integration of markets and therefore some recommendations from this study may be warranted.

Firstly, policy efforts aimed at improving marketing infrastructure should be prioritised in a bid to improve the functioning of regional markets. Chapter 3, particularly revealed a poor road infrastructure connecting Zambia to DRC as well as in country rural road networks, features that increase transaction costs and impede integration. Increased investment in road construction and maintenance is thus required.

Secondly, authorities can foster market integration through provision of effective market information. This is particularly critical for effective spatial arbitrage. Although various efforts have been made to set up agricultural information systems in specific countries, access to such information by market actors is limited especially in remote areas. Moreover, such systems are often only useful for traders dealing in nationally recognised products like Maize. A market information system involving country networking would thus improve regional arbitrage.

In summary, governments in the region needs to provide an enabling environment that promotes effective participation of market actors. This does not only require liberalising food markets but also concerted efforts to improve other market aspects such as infrastructure.

6.5 Critiques and suggestions for future research

This thesis, in an attempt to investigate market integration between informal trading partners, employed a combination of bean price, trade, and diesel price (proxy for transfer cost) data, an approach highly recommended in market integration literature. Nevertheless, the study was subject to some limitations. The principle limitation, also common in majority of related studies, was that of the evaluation of spatial integration purely based on secondary data. The conclusion on the relationship between markets based on this dataset only provides evidence on whether integration exists or not. Although this study attempted to eliminate government influence and trade volume related factors from a secondary data perspective, future studies should combine secondary data with primary data so as to comprehensively and empirically disentangle the effect of other factors unexamined in this study on price transmission. Also to the extent possible, transfer costs that goes beyond diesel prices should be included. In practice, it is almost impossible to find such data, however, it could be investigated under primary data.

This study has also identified aspects that needs to be researched further. In particular, the overall results reveal that informality in trade does not always guarantee market integration. This theme needs to be revisited with research that expands into other agriculture commodities and trading markets (countries). This will provide a conclusion from a broader perspective and better feed into regional integration policies.

Future studies should also consider exploring regional dry bean market integration between markets dominated by formal trade in order to establish any comparative patterns.

During the course of the analysis, the author of this thesis also observed that converting datasets into another currency by use of exchange rates can have significant implications on price

Stellenbosch University https://scholar.sun.ac.za

transmission findings. It is recommended therefore that exchange rates are employed as minimally as possible in such analysis. In light of this, analysis conducted at domestic level should strictly use domestic currency prices.

Stellenbosch University https://scholar.sun.ac.za

REFERENCES

Ackello-Ogutu, C. & Echessah, P.N. 1998. Unrecorded cross-border trade between Tanzania and her neighbours. Implication for food security. SD publication series 89. Bureau for Africa.

Acosta, A. 2012. Measuring spatial transmission of white maize prices between South Africa and Mozambique: An asymmetric error correction model approach. *African Journal of Agricultural and Resource Economics*, 7(1):1-13.

African Centre for Biodiversity. 2016. *N2Africa, the Gates foundation and legume commercialisation in Africa*. Johannesburg: African Centre for Biodiversity.

Afrika, J.G. & Ajumbo, G. 2012. Informal cross-border trade in Africa: implications and policy recommendations. *Africa Economic Brief*, 3(10):1-13.

Akhter, S. 2016. Market integration between surplus and deficit rice markets during global food crisis period. *Australian Journal of Agricultural and Resource Economics*, 59:1-17.

Akibode, S. & Maredia, M. 2011. Global and regional trends in production, trade and consumption of food legume crops. [Online]. Available at: http://impact.cgiar.org/sites/default/files/images/Legumetrendsv2.pdf (Accessed 14 August, 2015)

Alexander, C. & Wyeth, J. 1994. Cointegration and market integration: An application to the Indonesian rice market. *The Journal of Development Studies*, 30(2): 303-334.

Ama, N.O., Mangadi, K.T., Okurut, F.N. & Ama, H.A. 2013. Profitability of the informal cross-border trade: A case study of four selected borders of Botswana. *African Journal of Business Management*, 7(40):4221-4232.

Amikuzuno, J. 2009. Spatial price transmission and market integration in agricultural markets after liberalization in Ghana: Evidence from fresh tomato markets. Unpublished doctoral dissertation. Göttingen: Georg-August-University

Amikuzuno, J. & Donkoh, S.A. 2012. Border effects on spatial price transmission between fresh tomato markets in Ghana and Burkina-Faso: Any case for promoting trans-border trade in West Africa? *Journal of International Agricultural Trade and Development*, 8(1):81-97.

Amikuzuno, J., Setsoafia, E. & Seini, A.Y. 2015. Regional integration in West Africa: Tradeoffs and integration between Ghanaian and Burkinabe fresh tomato markets. *International Journal of Development*,1(1):42-58.

Ardeni, P.G. 1989. Does the law of one price really hold for commodity prices? *American Journal of Agricultural Economics*, 71(3):661-669.

Artingi-Ego, M., Opoloti, J. & Drale, A.S. 2006. *Can the benefits of developed country agricultural trade reforms trickle down to the rural agricultural households in least developed countries: Analysis via price transmission in selected agricultural markets in Uganda*. IIIS Discussion Paper Series No.159.

Stellenbosch University https://scholar.sun.ac.za

Baffes, J. 1991. Some further evidence on the law of one price: The law of one price still holds. *American Journal of Agricultural Economics*, 73(4):1264-1278.

Balcombe, K., Bailey, A. & Brooks, J. 2007. Threshold effects in price transmission: the case of Brazilian wheat, maize, and soya prices. *American Journal of Agricultural Economics*, 89(2):308-323.

Balke, N. & Fomby, T. 1997. Threshold cointegration. *International Economic Review*, 38(3):627-645.

Barrett, C.B. 1996. Market analysis methods: are our enriched toolkits well suited to enlivened markets? *American journal of agricultural economics*, 78(3):825-829.

Barrett, C.B. 1997. Food marketing liberalization and trader entry: evidence from Madagascar. *World Development*, 25(5):763-777.

Barrett, C.B. 2001. Measuring integration and efficiency in international agricultural markets. *Review of Agricultural Economics*, 23(1):19-32.

Barrett, C.B. 2005. Spatial market integration. In L.E. Blume & S.N. Durlauf (eds.). *The New Palgrave Dictionary of Economics*. 2nd edition. London: Palgrave Macmillan.

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.

Baulch, R.J. 1994. Spatial price equilibrium and food market integration. Unpublished doctoral dissertation. Stanford: Stanford University.

Baulch, B. 1997a. Transfer costs, spatial arbitrage and testing for food market integration. *American Journal of Agriculture economics*, 79:477-487.

Baulch, B. 1997b. Testing for food market integration revisited. *The Journal of Development Studies*, 33(4): 512-534.

Beebe, S., Skroch, P.W., Tohme, J., Duque, M.C., Pedraza, F. & Nienhuis, J. 2000. Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Science*, 40(1):264-273.

Bese, D., Düchting, A., Gebauer, H., Gerken, A., Maeda, C., Manyong, V.M., Rupschus, C. & Starosta, S. 2009. Promoting regional trade to enhance food security: A case study on the border region of Tanzania and Zambia. SLE Publication Series-S239. Berlin: SLE.

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., Logozzo, G., Stougaard, J., McClean, P., Attene, G. & Papa, R. 2012. Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is revealed by sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, 109(14): E788-E796.

Brenton, P., Dihel, N., Hoppe, M. & Soprano, C. 2014. Improving behaviour at borders to promote trade formalization: The charter for cross-border traders. *Policy Note* No. 41. [Online]. Available at: http://documents.worldbank.org/curated/en/894601468346146304/pdf/894730BRI0Char0Box0385291B00PUBLIC0.pdf (Accessed 20 March, 2016).

Stellenbosch University https://scholar.sun.ac.za

Brooks, C. 2008. *Introductory econometrics for finance*. 2nd edition. New York: Cambridge University Press.

Broughton, W.J., Hernandez, G., Blair, M., Beebe, S., Gepts, P. & Vanderleyden, J. 2003. Beans (Phaseolus Spp.)-model food legumes. *Plant and Soil Journal,* 252(1):55-128.

Burke, W.J. 2012. Maize production in Zambia and regional marketing: Input productivity and output price transmission. Unpublished doctoral dissertation. Michigan: Michigan State University.

Burke, W.J. 2016. Personal communication (12/10/2016). Lecturer, Stanford University, CA.

Burke, W.J. & Myers, R.J. 2014. Spatial equilibrium and price transmission between Southern African maize markets connected by informal trade. *Food Policy,* 49:59-70.

Buruchara, R., Chirwa, R., Sperling, L., Mukankusi, C., Rubyogo, J.C., Mutonhi, R. & Abang, M.M. 2011. Development and delivery of bean varieties in Africa: The Pan-Africa Bean Research Alliance (PABRA) model. *African crop science journal*, 19(4):227-245.

Cassel, G. 1918. Abnormal deviations in international exchanges. *The Economic Journal*, *28*(112):413-415.

Chan, K.S. 1993. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *The Annals of Statistics*, (1): 520-533.

Chan, N.G., Yau, C.Y. & Zhang, R. 2015. LASSO estimation threshold autoregressive models. *Journal of Econometrics*.189:285-296.

Chemonics International. 2009. Staple foods value chain analysis country report – Zambia. USAID

Chen, H., Chong, T.T.L. & Bai, J. 2012. Theory and applications of TAR model with two threshold variables. *Econometric Reviews,* 31(2):142-170.

Chen, Z. & Knez, P.J. 1995. Measurement of market integration and arbitrage. *Review of financial studies*, 8(2):287-325.

Cirera, X. & Arndt, C. 2008. Measuring the impact of road rehabilitation on spatial market efficiency in maize markets in Mozambique. *Agricultural Economics*, 39(1):17-28.

Coleman, A. 2009. A model of spatial arbitrage with transport capacity constraints and endogenous transport prices. *American Journal of Agricultural Economics*, 91:42–56.

Cranfield, J., Preckel, P. & Hertel, T. 2007. Poverty analysis using an international cross-country demand system. Policy Research Working Paper 4285. Washington, DC. World Bank.

Delgado, C.L. 1986. A variance components approach to food grain market integration in northern Nigeria. *American Journal of Agricultural Economics,* 68(4): 970-979.

Dercon, S. 1995. On market integration and liberalisation: method and application to Ethiopia. *The Journal of Development Studies*, 32(1):112-143.

Dickey, D.A. & Fuller, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427-431.

Stellenbosch University https://scholar.sun.ac.za

Dickey, D.A. & Fuller, W.A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 49(4):1057-1072.

Dougherty, C. 2007. *Introduction to econometrics*. USA: Oxford university press.

Enders, W. & Siklos, P.L. 2001. Cointegration and threshold adjustment. *Journal of Business and Economic Statistics*, 19(2):166-176.

Engel, C. & Rogers, J.H. 1996. How Wide Is the Border? *The American Economic Review*, 86(5):1112-1125.

Engle, R.F. & Granger, C.W. 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55(2):251-276.

Engle, R.F. & Yoo, B.S. 1987. Forecasting and testing in co-integrated systems. *Journal of econometrics*, *35*(1):143-159.

Enke, S. 1951. Equilibrium among spatially separated markets: solution by electric analogue. *Econometrica*, 19(1):40–47.

Fackler, P. L. 1996. Spatial price analysis: A methodological review. Proceedings of the NCR-134 conference on applied commodity price analysis, forecasting, and market risk management. Chicago, IL.

Fackler, P.L. & Goodwin, B.K. 2001. Spatial price analysis. In B.L. Gardner & G.C. Rausser (eds.). *Handbook of agricultural economics.* Volume 1B. Amsterdam: Elsevier. 971-1024

FEWSNET (Famine Early Warning Systems Network). 2015. Democratic Republic of the Congo: Staple food market fundamentals. [Online]. Available: https://www.usaid.gov/sites/default/files/documents/1866/DRC_MarketFundamentals_20151026.pdf [Accessed 7 June, 2016]

Faminow, M.D. & Benson, B.L. 1990. Integration of spatial markets. American *Journal of Agriculture Economics*, 72:49-62.

FAO (Food and Agriculture Organization). 2012. *Foreign agricultural investment country profile: Democratic Republic of the Congo (DRC).* Rome: Food and Agriculture Organization.

FAO (Food and Agriculture Organization). 2016. FAOSTAT. [online]. Available: www.fao.org [Retrieved in May to August, 2016].

Gepts, P. 1998. Origin and evolution of common bean: past events and recent trends. *HortScience*, 33(7):1124-1130.

Goletti, F. & Babu, S. 1994. Market liberalization and integration of maize markets in Malawi. *Agricultural Economics*, 11(2): 311-324.

Goletti, F., Ahmed, R. & Farid, N. 1995. Structural determinants of market integration: The case of rice markets in Bangladesh. *The Developing Economies*, 33(2):196-198.

Gonzalez-Rivera, G. & Helfand, S.M. 2001a. 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.

Stellenbosch University https://scholar.sun.ac.za

González-Rivera, G. & Helfand, S.M. 2001b. Economic development and the determinants of spatial integration in agricultural markets. Working Paper 01-28. Riverside: University of California.

Gonzalo, J. & Pitarakis, J.Y. 2002. Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics*, 110(2):319-352.

Goodwin, B.K. & Piggott, N.E. 2001. Spatial market integration in the presence of threshold effects. *American Journal of Agricultural Economics*, 83(2):302-317.

Goodwin, B.K. & Schroeder, T.C. 1990. Testing perfect spatial market integration: An application to regional United States cattle markets. *North Central Journal of Agriculture Economics*, 12(2):173-186.

Goodwin, B.K. & Schroeder, T.C. 1991. Cointegration tests and spatial price linkages in regional cattle markets. *American Journal of Agricultural Economics*, 73(2):452-464.

Goodwin, B.K. 1992a. Multivariate cointegration tests and the law of one price in international wheat markets. *Review of Agricultural Economics*, 14(1):117-124.

Goodwin, B.K. 1992b. Multivariate cointegration tests and the law of one price: A clarification and correction. *Review of Agricultural Economics*, 14(2): 337-338.

Götz, L., Glauben, T. & Brümmer, B. 2013. Wheat export restrictions and domestic market effects in Russia and Ukraine during the food crisis. *Food Policy*, 38(1):214-226.

Graham, P.H. & Ranalli, P. 1997. Common bean (Phaseolus vulgaris L.). *Field Crops Researc*h, 53(1-3):131-146.

Granger, C.W.J. & Newbold, P. 1974. Spurious regressions in econometrics. *Journal of econometrics*, 2(2):111-120.

Granger, C.W.J. & Weiss, A.A. 1983. Time series analysis of error correcting models. In S. Karlin, T. Amemiya & L. A. Goodman (eds.). *Studies in Econometrics, Time Series and Multivariate Statistics*. New York: Academic Press. 255-278.

Granger, C.W.J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424-438.

Granger, C.W.J. 1981. Some properties of time series data and their use in econometric model specification. *Journal of econometrics*, 16(1):121-130.

Greb, F., von Cramon-Taubadel, S., Krivobokova, T. & Munk, A. 2013. The estimation of threshold models in price transmission analysis. *American Journal of Agricultural Economics*, 95(4):900-916.

Gujarati, D.N. 2004. *Basic econometrics*. 4th edition. New York: McGraw-Hill.

Gupta, S. & Mueller, R.A. 1982. Analysing the pricing efficiency in spatial markets: Concept and application. *European Review of Agricultural Economics*, 9(3):301-312.

Hamilton, J.D. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 57(2):357-384.

Hamilton, J.D. 1994. *Time series analysis.* Volume 2. Princeton: Princeton university press.

Stellenbosch University https://scholar.sun.ac.za

Hamilton, J.D. 1995. Rational expectations and the economic consequences of changes in regime. In K.D. Hoover (ed.). *Macroeconometrics*: *Development, tension and prospects*. New York: Springer. 325-349.

Hamilton, J.D. 2001. A parametric approach to flexible nonlinear inference. *Econometrica*, 69(3):537-573.

Hamzakaza, P., Katungi, E., Reyes, B., Maredia, M., Muimui, K. & Ojara, M. 2014. Assessing access and adoption of common bean improved varieties in Zambia. Research Technical Report. Pan African Bean Research Alliance (PABRA).

Hansen, B.E. & Seo, B. 2002. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of Econometrics*, 110(2):293-318.

Hansen, B.E. 1996. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64(2):413-430.

Hansen, B.E. 1999. Testing for Linearity. *Journal of Economic Surveys*, 13(5):551-576.

Harriss, B. 1979. There is method in my madness: or is it vice versa? Measuring agricultural market performance. *Food Research Institute Studies*, XVII(2):197-218.

Hillocks, R.J., Madata, C.S., Chirwa, R., Minja, E.M. & Msolla, S. 2006. Phaseolus bean improvement in Tanzania, 1959-2005. *Euphytica*, 150(1-2):215-231.

Hossain, M.I. & Verbeke, W. 2010. Evaluation of rice market integration in Bangladesh. *The Lahore Journal of Economics*, 15(2):77-96.

Ihle, R. 2009. Models for analyzing nonlinearities in price transmission. Unpublished doctoral dissertation. Göttingen: Georg-August-University.

Ihle, R., von Cramon-Taubadel, S. & Zorya, S. 2009. Markov-switching estimation of spatial maize price transmission processes between Tanzania and Kenya. *American Journal of Agricultural Economics*, 91(5):1432-1439.

INS (Institut National de la Statistique). 2015. *Annuaire statistique 2014*. Institut National de la Statistique. Kinshasa.

ITC (International Trade Centre). 2016a. United Republic of Tanzania value chain roadmap for pulses 2016-2020. [online]. Available at: http://www.mit.go.tz/uploads/files/Tanzania%20Roadmap%20for%20Pulses.pdf [Accessed 7 May, 2016]

ITC (International Trade Centre). 2016b. Trade map data. [Online]. Available at: www.trademap.org [Retrieved in May to August, 2016].

Jayne, T.S. & Argwings-Kodhek, G. 1997. Consumer response to maize market liberalization in urban Kenya. *Food Policy*, 22(5):447-458.

Jensen, R. 2007. The digital provide: Information (technology), market performance and welfare in the South Indian fisheries sector. *The quarterly journal of economics*, 122(3):879-924.

Stellenbosch University https://scholar.sun.ac.za

Johansen, S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamic Control*, 12(2-3):231-254.

Johansen, S. & Juselius, K. 1990. Maximum likelihood estimation and inference on cointegration-with application to the demand money. *Oxford Bulletin of Economics and Statistics*, 52(2):169-210.

Kabbiri, R., Dora, M., Elepu, G. & Gellynck, X. 2016. A Global perspective of food market integration: A review. *Agrekon*, 55(1-2):62-80.

Kambewa, P.S. 1997. The bean subsector in Malawi: Historical developments, current status and policy issues. Unpublished master's thesis. Michigan: Michigan State University.

Kami, J., Velásquez, V.B., Debouck, D.G. & Gepts, P. 1995. Identification of presumed ancestral DNA sequences of phaseolin in Phaseolus vulgaris. *Proceedings of the National Academy of Sciences of the United States of America*, 92(4):1101-1104.

Katungi, E., Farrow, A., Chianu, J., Sperling, L. & Beebe, S. 2009. Common bean in Eastern and Southern Africa: A situation and outlook analysis. *International Centre for Tropical Agriculture*.

Keats, S., Wiggins, S., Compton, J. & Vigneri, M. 2010. Food price transmission: Rising international cereals prices and domestic markets. Project briefing, No. 48. London: ODI.

Kherallah, M., Delgado, C., Gabre-Madhin, E., Minot, N. & Johnson, M. 2002. The road half travelled: agricultural market reform in Sub-Saharan Africa. *Oléagineux, Corps gras, Lipides*, 9(6):416-425.

Klaes, M. 2000. The history of the concept of transaction costs: neglected aspects. *Journal of the History of Economic Thought*, 22(02):191-216.

Konya, L. 2004. Unit root, cointegration and Granger causality test results for export and growth in OECD countries. *International Journal of applied Econometrics and Quantitative Studies,*1-2:67-94.

Korir, M.K., Odhiambo, M.O., Kimani, P.M., Mukishi, P. & Iruria, D.M. 2003. Spatial price integration: A cointegration approach to regional bean markets in Kenya and Tanzania. *African Crop Science Conference Proceedings*, 6:609-612.

Kouyaté, C. & von Cramon-Taubadel, S. 2016. Distance and border effects on price transmission: A meta-analysis. *Journal of Agricultural Economics*, 67(3):255-271.

Krolzig, H.M. 1997. *Markov-switching vector autoregressions: Modelling, statistical inference and application to business cycle analysis*. Berlin: Springer-Verlag.

Krolzig, H.M. 2004. MSVAR - An ox package designed for the econometric modelling of univariate and multiple time series subject to shifts in regime, Version 1.31k. [Online]. Available at: http://www.economics.ox.ac.uk/research/hendry/krolzig/msvar.html [Accessed 14 February, 2016)

Kwiatkowski, D.P., Phillips, C. B., Schmidt, P. & Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54:159-178.

Stellenbosch University https://scholar.sun.ac.za

Lesser, C. & Moisé-Leeman, E. 2009. Informal cross-border trade and trade facilitation reform in Sub-Saharan Africa, final report. OECD Trade Policy Working Paper No. 86, Organisation for Economic Co-operation and Development. [online]. Available at: https://www.oecd.org/tad/42222094.pdf (Accessed 10 August, 2016)

Listorti, G. & Esposti, R. 2012. Horizontal price transmission in agricultural markets: fundamental concepts and open empirical issues. *Bio-based and Applied Economics*, 1(1): 81-108.

Listorti, G. 2009. Testing international price transmission under policy intervention: An application to the soft wheat market. Unpublished doctoral dissertation. Ancona: Università Politecnica delle Marche.

Little, P.D. 2010. Unofficial cross-border trade in Eastern Africa. In A. Sarris & J. Morrison (eds.). *Food security in Africa: Market and trade policy for staple foods in Eastern and Southern Africa*. Cheltenham: Edward Elgar Publishing. 158-181.

Loveridge, S. 1991. Marketing in Rwanda imports and infrastructure. *Food Policy*, 16:95–104.

Lüutkepohl, H., Saikkonen, P. & Trenkler, C. 2001. Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process. *The Econometrics Journal*, 4(2):287-310.

Mahadeva, L. & Robinson, P. 2004. Unit root testing to help model building. In A. Blake & G. Hammond (eds.). *Handbooks in central banking*, No. 22. London: Bank of England.1-53.

Mann, J.M. 2012. Threshold cointegration with applications to the oil and gasoline industry. Unpublished doctoral dissertation. Kingston: Queen's University.

Marshall, A. 1890. *Principles of economics*. 8th edition. London: Macmillan and Co.

Mauyo, L.W., Chianu, J.N., Nassiuma, B.K. & Musebe, R.O. 2010. Cross-Border bean market performance in Western Kenya and Eastern Uganda. *Journal of Service Science and Management*, 3:501-511.

Mauyo, L.W., Okalebo, J.R., Kirkby, R.A., Buruchara, R., Ugen, M., Mengist, C.T., Anjichi, V.E. & Musebe, R.O. 2007. Technical efficiency and regional market integration of cross-border bean marketing in western Kenya and Eastern Uganda. *African Journal of Business Management*, 1(4):77-84.

Mayaka, V.K. 2013. An assessment of dry beans market integration in selected markets in Kenya. Unpublished master's thesis. Njoro: Egerton University.

McNew, K. & Fackler, P.L. 1997. Testing market equilibrium: Is cointegration informative? *Journal of Agricultural and Resource Economics*, 22(2):191-207.

McNew, K. 1996. Spatial market integration: Definition, theory, and evidence. *Agricultural and Resource Economics Review*, 25(1):1-11.

Meyer, J. 2004. Measuring market integration in the presence of transaction costs-A threshold vector error correction approach. *Agricultural Economics*, 31(2-3):327-334.

Minten, B. & Kyle, S. 1999. The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. *Journal of Development Economics*, *60*(2):467-495.

Stellenbosch University https://scholar.sun.ac.za

Mtumbuka, W.S., Mapemba, L., Maonga, B. & Mwabumba, M. 2014. Spatial price integration among selected bean markets in Malawi: A threshold autoregressive model approach. Working paper 07. International Food Policy Institute. [online]. Available at: http://cdm15738.contentdm.oclc.org/utils/getfile/collection/p15738coll2/id/128214/filename/128425.pdf (Accessed 2 December, 2015)

Muimui, K.K., Chirwa, R.M., Zulu, R.M. & Maereka, E.K. 2016. Identifying acceptable improved beans varieties for income and nutrition in Zambia. Poster presentation at the joint Pan-African grain legume and world cowpea conference. 27 February-4 March, Livingstone, Zambia.

Mushtaq, K., Gafoor, A. & Dad, M. 2008. Apple market integration: Implications for sustainable agricultural development. *Lahore Journal of Economics*, *13*(1):129-138.

Mutambatsere, E. & Christy, R.D. 2008. Market integration and efficiency in the presence of cross-border trade restrictions: Evidence from selected maize markets in Southern Africa. Paper prepared for presentation at the American Agricultural Economics Association annual meetings. 27-29 July, Orlando, United States of America.

Mutambatsere, E., Mabaya, E. & Christy, R.D. 2006. Integration and equilibrium of maize markets in Southern Africa: A SADC sub regional assessment. Working paper 2006-18. Ithaca, USA: Cornell University.

Muyatwa, P.V. 2000. The liberalization and integration of regional maize markets in Zambia. Unpublished doctoral dissertation. Winnipeg: University of Manitoba.

Myers, R.J. 1994. Time series econometrics and commodity price analysis: A review. *Review of Marketing and Agricultural Economics*, *62*(2):167-181.

Myers, R.J. 2013. Evaluating the effectiveness of inter-regional trade and storage in Malawi's private sector maize markets. *Food Policy*, *41*:75-84.

Myers, R.J. & Jayne, T.S. 2012. Multiple-regime spatial price transmission with an application to maize markets in Southern Africa. *American Journal of Agricultural Economics*, 94(1):174-188.

Ndibongo, L.T., Myers, R.J., Jayne, T.S. & Meyer, F.H. 2010. Measuring integration and efficiency in maize grain markets: The case of South Africa and Mozambique. Paper presented at the Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference. 19-23 September, Cape Town, South Africa.

Negassa, A. & Myers, R.J. 2007. Estimating policy effects on spatial market efficiency: An extension to the parity bounds model. *American Journal of Agricultural Economics*, 89(2): 338-352.

Negassa, A., Myers, R. & Gabre-madhin, E. 2003. *Analyzing the grain market efficiency in developing countries: Review of existing methods and extensions to the parity bound model*. MTID Discussion Paper, No. 63. Washington: International Food Policy Research Institute.

Ng, S. & Perron, P. 1995. Unit root tests in ARMA models with data-dependant methods for the selection of truncated lag. *Journal of the American Statistical Association*, 90(429):268-281.

Stellenbosch University https://scholar.sun.ac.za

OANDA. 2016. Average exchange rates. [Online]. Available at: www.oanda.com [Retrieved in May to July, 2016].

Officer, L.H. 1986. The law of one price cannot be rejected: Two tests based on the tradable/nontradable price ratio. *Journal of Macroeconomics*, 8(2):159-182.

Ogalo, V. 2010. *Informal cross-border trade in EAC: Implications for regional integration and development.* Research paper. Nairobi: CUTS African Resource Centre.

Park, A., Jin, H., Rozelle, S. & Huang, J. 2002. Market emergence and transition: arbitrage, transaction costs, and autarky in China's grain markets. *American Journal of Agricultural Economics*, 84(1):67-82.

Phillips, P.C. & Ouliaris, S. 1988. Testing for cointegration using principal components methods. *Journal of Economic Dynamics and Control*, 12(2):205-230.

Phillips, P.C. & Perron, P. 1988. Testing for a unit root in time series regression. *Biometrika*, 75(2):335-346.

Rapsomanikis, G., Hallam, D. & Conforti, P. 2003. Market integration and price transmission in selected food and cash crop markets of developing countries: Review and applications. In FAO (ed.). *Commodity market review 2003-2004*. Rome: Food and Agriculture Organization. 51-75.

Rashid, S., Minot, N. 2010. Are staple food markets efficient in Africa? Spatial price analyses and beyond. Paper prepared for the COMESA policy seminar on food price variability: Causes, consequence and policy options, 25–26 January, Maputo, Mozambique.

Ravallion, M. 1986. Testing market integration. *American Journal of Agricultural Economics,* 68(1):102-109

Rogoff, K. 1996. The purchasing power parity puzzle. *Journal of Economic literature*, 34(2): 647-668.

Samuelson, P. 1952. Spatial price equilibrium and linear programming. American Economic Review, 42(3):283-303.

Serra, T., Gil, J.M. & Goodwin, B.K. 2006. Local polynomial fitting and spatial price relationships: price transmission in EU pork markets. *European Review of Agricultural Economics*, 33(3):415-436.

Sexton, R.J., King, C.L. & Carman, H.F. 1991. Market integration, efficiency of arbitrage and imperfect competition: Methodology and application to U.S. celery. *American Journal of Agricultural Economics*, 73(3):568-580

Sichilima, T., Mapemba, L. & Tembo, G. 2016. Drivers of dry common beans trade in Lusaka, Zambia: A trader's perspective. *Sustainable Agriculture Research,* 5(2):15-26.

Siddiq, M. & Uebersax, M.A. 2013. Dry beans and pulse production and consumption: An overview. In M. Siddiq & M.A. Uebersax (eds.). *Dry beans and pulses production, processing and nutrition*. Iowa: John Wiley and sons. 3-22.

Siegel, P. & Alwang, J. 2005. Poverty reducing potential of smallholder agriculture in Zambia: Opportunities and constraints. World Bank Africa Region Working Paper Series No. 85.

Stellenbosch University https://scholar.sun.ac.za

Silvapulle, P. & Jayasuriya, S. 1994. Testing for Philippines rice market integration: A multiple cointegration approach. *Journal of Agricultural Economics*, 45(3):369-380.

Sitko, N.J., Chapoto, A., Kabwe, S., Tembo, S., Hichaambwa, M., Lubinda, R., Chiwawa, H., Mataa, M., Heck, S. & Nthani, D. 2011. *Technical compendium: Descriptive agricultural statistics and analysis for Zambia in support of the USAID mission's Feed the Future Strategic Review*. Working paper no. 52. Lusaka: Food Security Research Project.

Somda, J., Tollens, E. & Kamuanga, M. 2005. Transaction costs and the marketable surplus of milk in smallholder farming systems of The Gambia. *Outlook on Agriculture*, 34(3):189-195.

Spiller, P.T. & Huang, C.J. 1986. On the extent of the market: wholesale gasoline in the north eastern United States. *The Journal of Industrial Economics*, 35(2):131-145.

Spiller, P.T. & Wood, R.O. 1988. The estimation of transaction costs in arbitrage models. *Journal of Econometrics*, 39(3):309-326.

Stephens, E.C., Mabaya, E., von Cramon-Taubadel, S. & Barrett, B.C. 2012. Spatial price adjustment with and without trade. *Oxford Bulletin of Economics and Statistics*, 74(3):453-469.

Strikholm, B. & Teräsvirta, T. 2006. A Sequential procedure for determining the number of regimes in a threshold autoregressive model. *Econometrics Journal*, 9(3):472-491.

Takayama, T. & Judge, G. 1964. Spatial equilibrium and quadratic programming. *Journal of Farm Economics*, 46(1), February: 67-93.

Takayama, T. & Judge, G. 1971. *Spatial and temporal price and allocation models*. Amsterdam: North-Holland Publishing Compan.

Teka, T. & Azeze, A. 2002. *Cross-border livestock trade, and food security in the Ethiopia-Djibouti, and Ethiopia-Somalia borderlands*. OSSREA Development Research Report Series, no. 4. Addis Ababa: Organization for Social Science Research in Eastern and Southern Africa.

Teka, T., Azeze, A. & Gebremariam, A. 1999. *Cross-border livestock trade and food security in southern and South Eastern borderlands of Ethiopia*. OSSREA Development Research Report series no. 1. Addis Ababa: Organization for Social Science Research in Eastern and Southern Africa.

Tembo, S. & Sitko, N. 2013. Technical compendium: Descriptive agricultural statistics and analysis for Zambia. Working Paper 76. Lusaka. Indaba Agricultural Policy Research Institute.

Timmer, C.P. 1974. A model of rice marketing margins in Indonesia. *Food research Institute studies*, 12(2):309-326.

Timmer, C.P. 1987. Corn marketing. In C.P Timmer (ed.). *The corn economy of Indonesia*. Ithaca, NY: Cornell University Press.

Tong, H. & Lim, K. S. 1980. Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society Series B*, 42(3):245-292.

Stellenbosch University https://scholar.sun.ac.za

Tsay, R.S. 1989. Testing and modelling threshold autoregressive processes. *Journal of the American Statistical Association*, 84(405):231-240.

UNDP (United Nations Development Programme). 2015. *Human development report 2015*. New York: United Nations Development Programme.

Van Campenhout, B. 2007. Modelling trends in food market integration: Method and an application to Tanzanian maize markets. *Food Policy*, 32:112–127.

Van Dijk, D., Teräsvirta, T. & Franses, P.H. 2002. Smooth transition autoregressive model: A survey of recent developments. *Econometric reviews*, 21(1):1-47.

Van Rooyen, J. 2000. Regional food security and agricultural policy in Southern Africa: A challenge of policy conversion in diverse settings. *Development Southern Africa*, 17(1):7-22.

Verbeek, M. 2004. *A guide to modern econometrics.* 2nd edition. West Sussex: John Wiley and Sons.

Versailles, B. 2012. Market integration and border effects in Eastern Africa. CSAE Working Paper WPS/2012-01. Centre for the Study of African Economies, University of Oxford.

Williamson, J. 1986. Target Zones and the Management of the Dollar. *Brookings Papers on Economic Activity*, 17(1):165-174.

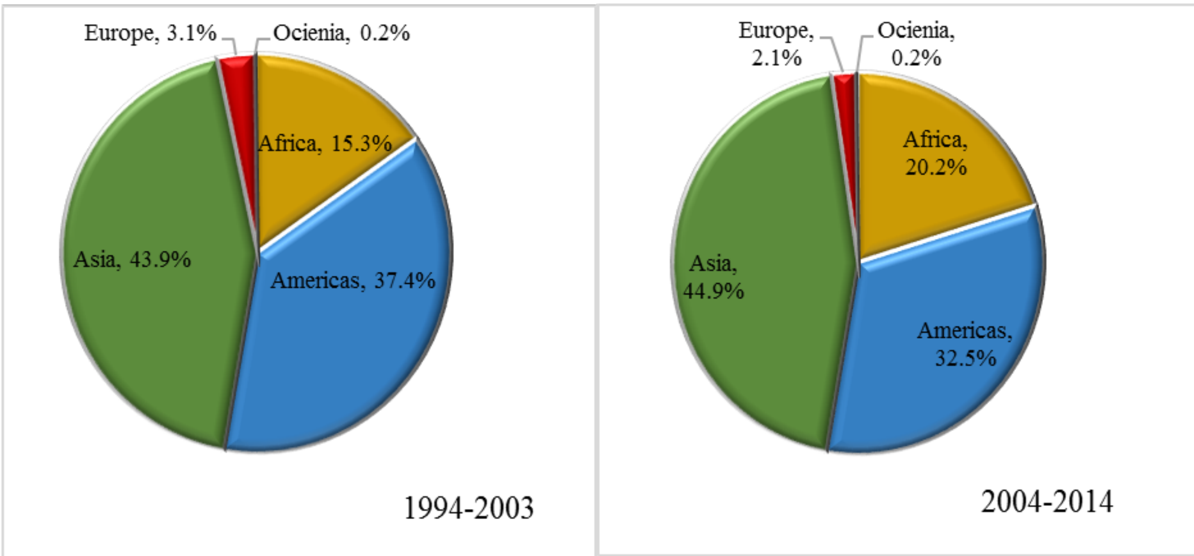
Wortman, C., Kirkby, R., Aledu, C. & Allen, D. 1998. Atlas of common bean (Phaseolus vulgaris L.) production in Africa. CIAT Publication No. 297, Cali, Colombia.

Zant, W. 2012. How is the liberalization of food markets progressing? Market integration and transaction costs in subsistence economies. *The World Bank Economic Review*, 27: 28–54.

Zhao, J., Goodwin, B. K. & Pelletier, D. 2012. A New approach to investigate market integration: A markov-switching autoregressive model with time-varying transition probabilities. Paper prepared for presentation at the Agricultural and Applied Economics Association annual meeting. 12-14 August, Seattle, United States of America.

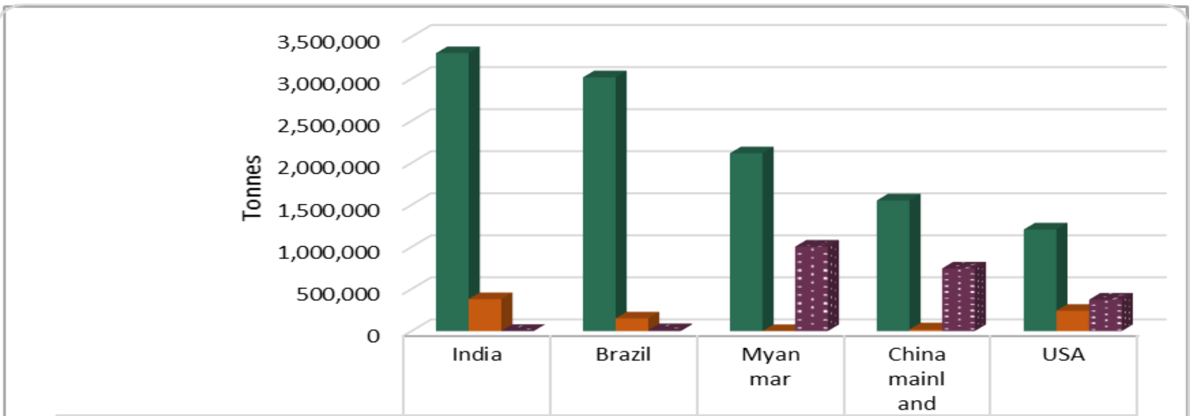
Zulu, B., Nijhoff, J.J., Jayne, T.S. & Negassa, A. 2000. *Is the glass half-empty or half full? An analysis of agricultural Production trends in Zambia*. Working paper no. 3. Lusaka: Food Security Research Project.

APPENDIX A1: Regional share in world beans production (1994-2003) and (2004-2014)



Source: FAO (2016)

APPENDIX A2: Annual average export and imports in the top 5 producing countries (1994-2013)

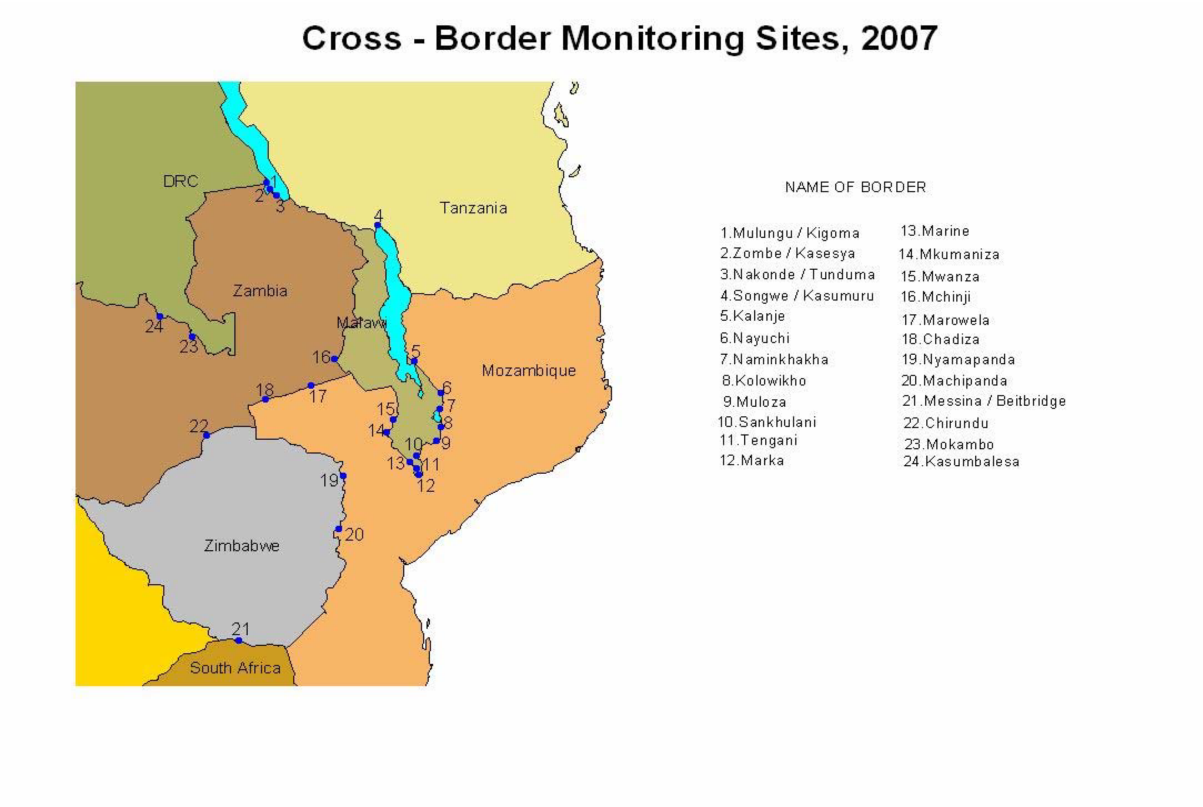


	India	Brazil	Myan mar	China mainl and	USA
Production quantity (average)	3,304,605	3,016,546	2,113,893	1,553,645	1,207,780
import quantity (average)	379,755	150,039	302	19,675	240,984
Export quantity (average)	1,959	10,796	1,004,527	744,654	373,546

Source: FAO (2016)

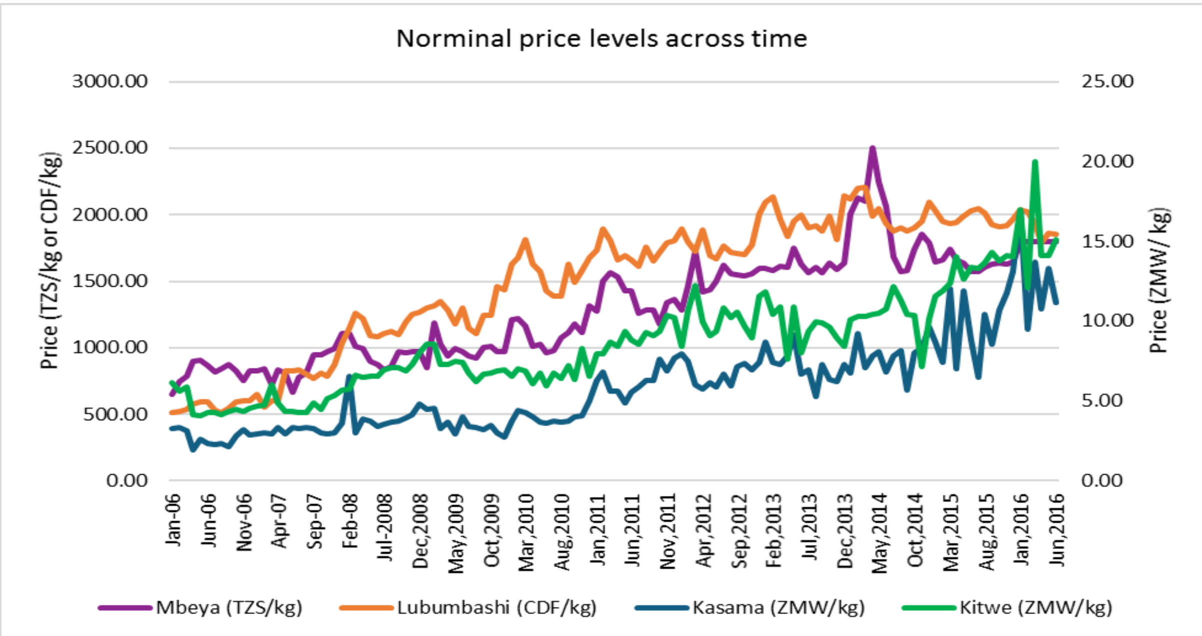
Stellenbosch University https://scholar.sun.ac.za

APPENDIX A3: Location of FEWSNET border monitors



Source: fewsnet.net

APPENDIX B1 : Nominal prices of dry beans in their local currencies



Source: Based on data collected by CSO, Zambia, NBS, Tanzania and INS, DRC.

Stellenbosch University https://scholar.sun.ac.za

APPENDIX B2: Jointly estimated threshold values using total trade (Mbeya and Kasama)

	Threshold value		Pernalty criterion function (GP values)				
Model	**Threshold 1**	**Threshold 2**	**BIC**	**AIC**	**HQ**	**BIC2**	**BIC3**
Two regime	209.125		-0.108	0.097	0.013	-0.458	-0.808
Three regime	209.125	878.06	-0.242	-0.168	-0.001	-0.941	-1.640

Source: Based on data collected by CSO, Zambia, EWURA, Tanzania and NBS, Tanzania.