CESifo Working Papers
www.cesifo.org/wp

Measuring the Integration of Staple Food Markets in Sub-Saharan Africa: Heterogeneous Infrastructure and Cross Border Trade in the East African Community

Rico Ihle
Stephan von Cramon-Taubadel
Sergiy Zorya

CESIFO WORKING PAPER NO. 3413
CATEGORY 8: TRADE POLICY
APRIL 2011

CESifo *Working Paper No. 3413*

Measuring the Integration of Staple Food Markets in Sub-Saharan Africa: Heterogeneous Infrastructure and Cross Border Trade in the East African Community

Abstract

This analysis employs cointegration methods and semiparametric regression in order to assess the integration of maize markets and the factors determining national and cross-national transmission of price signals in Sub-Saharan Africa. We use a rich dataset of 16 series of wholesale maize prices between 2000 and 2008 for Kenya, Tanzanian and Uganda. Distance is shown to have a significant nonlinear impact on the transmission of information - modelled using a semi-parametric partially linear model. Border effects are found to be heterogeneous. The empirical results provide strong evidence that the Tanzanian market is isolated from the rest of East Africa and internally fragmented.

JEL-Code: C320, Q110, Q130, Q170, Q180.

Keywords: border effect, spatial market integration, cointegration, semi-parametric regression, partially linear model, Eastern Africa, maize.

Rico Ihle
Institute of Agricultural Economics and Rural Development
Georg-August-University Göttingen
Platz der Göttinger Sieben 5
Germany - 37073 Göttingen
rihle@gwdg.de

Stephan von Cramon-Taubadel
Institute of Agricultural Economics and Rural Development
Georg-August-University Göttingen
Göttingen / Germany
scramon@gwdg.de

Sergiy Zorya
Agricultural and Rural Development Dept.
World Bank
Washington, D.C. / USA
szorya@worldbank.org

Rico Ihle thanks the federal state of Niedersachsen and the Centre for Statistics at the Georg-August-Universität Göttingen for the provision of a Georg-Christoph-Lichtenberg-PhD scholarship.

Measuring the Integration of Staple Food Markets in Sub-Saharan Africa: Heterogeneous Infrastructure and Cross Border Trade in the East African Community

1 Introduction

Maize is the most important staple food in Eastern Africa as well as in Sub-Saharan Africa (see, for example, Awuor 2007). It represents the major food staple in the region as it is the main source of calories in the average diet accounting for more than one third of caloric intake in Kenya and Tanzania and one tenth in Uganda (FAOSTAT 2009). Consequently, the crop plays an important role in production and intra-regional trade. Tanzania and Kenya are the largest producers in the region, the former being largely self-sufficient, however the latter being by far the largest importer in the region. The demand of the net consumption areas of Kenya is largely met by maize flowing in mainly from the country's central highlands, eastern Uganda and northern Tanzania (World Bank 2008a), with Uganda being the largest exporting country in the region (for a more detailed account of the regional maize trade flows, see, e.g., Awuor 2007, Michigan State University 2008 or FEWSNET 2009). According to the UN COMTRADE database (United Nations 2009b), approximately 60% on average of total maize exports from Uganda and Tanzania are shipped to Kenya, and the crop is one of the five most important commodities exported by Uganda and Tanzania to the East African market.

Although all three countries belong to the East African Community (EAC), in whose framework a customs union was agreed on commencing in January 2005, they differ in their attitudes towards agriculture and in their trade policies which partly do not follow the objectives of the union. Kenya and Uganda are largely liberalized economies in contrast to Tanzania. According to the Agricultural Distortions Project of the World Bank (World Bank 2008b), distortions of the agricultural sector have reduced in all three countries since 1980. However, while Kenya and Uganda changed their policies from taxation to a slight support of farm-gate prices, Tanzania's sector is still relatively regulated and price incentives remain strongly distorted. Uganda's political decision makers seem to be aware of the country's potential to become the region's food basket and earn considerable export revenues. Hence, no export duties, bans or other restrictions on trade of food commodities exist. Agricultural policy in Kenya aims at supporting and stabilizing prices via the National Cereals and Produce Board (Jayne et al. 2008). Tanzania's approach is characterized by the effort to ensure nation-wide food security. This target is pursued by a variety of measures from the local up to the national level, e.g., national export bans for maize (Delgado & Minot 2000; see Temu et al. 2007 for a detailed chronological account of the country's export policy; Meijerink et al. 2009).

This study focuses on the question of whether the heterogeneous national policies lead to differences in the price response to disequilibria (i.e. price transmission PT) within and between these countries. If the customs union is working effectively, the magnitude of the price reaction should not differ for arbitrage within countries from arbitrage across national borders, i.e. no border effects should be present in the data.

We first assess the question of whether maize markets in and between the three countries share a long-run price relationships (i.e. are cointegrated). Second, we analyze the factors explaining the magnitude of price transmission. In doing so, we draw on the border effects literature, which aims at explaining price volatility and evolved following the seminal paper of Engel & Rogers (1996), and the gravity model literature pioneered by McCallum (1995), which aims at explaining trade flows (see also, e.g., Feenstra 2002, Evans 2003, Morshed 2003, Chen 2004 or Anderson & van Wincoop 2004). Within the field of agricultural economics, few publications to date have investigated border effects in trade flows or in prices of agricultural products. For example, Gardner & Brooks (1994) and Berkowitz & DeJong (1999) assess, among others, the impact of distance and administrative borders on regional food prices in Russia. Furtan & van Melle (2004) study border effects in agricultural trade between Canada, Mexico and the US. Olper & Raimondi (2008a, 2008b, 2009) analyze determinants of the food trade between a number of OECD members.

In contrast to the above-mentioned studies which focus on trade flows and prices, the central focus of this article lies in the analysis of the effects of borders and distance on the magnitude of price transmission which is measured by the reaction of two prices on deviations from their long-run price equilibrium. The methodological analysis consists of two steps. First, we evaluate the cointegration properties of 85 market pairs in Eastern Africa and estimate pairwise vector error correction models (VECM). We focus on the estimation of the strength of PT in the domestic markets of Kenya, Tanzania and Uganda, respectively. Due to the region's trade pattern, i.e., Kenya's large maize imports from the other two countries, we hence also consider cross-border PT between Kenya and Tanzania and Kenya and Uganda, respectively. In the second step, we assess the influence of various determinants on the estimated strengths of PT and discuss the results in detail by comparing competing model strategies of a parametric and a semi-parametric model.

The paper is structured as follows. The next section briefly explains the concept of cointegration and the relationship between transaction costs (TC) in more detail, as well as the strength of PT. Section 3 provides a detailed account of the role of border effects as a component of TC. The following chapter discusses the problem of misspecification bias due to the unknown functional form and suggests an alternative semi-parametric estimation approach which is the one adopted in this paper. Afterwards, the data used in the analysis is briefly presented. In section 6, the results of both the parametric and the semi-parametric models are presented. Finally, section 7 discusses the results, and the last section concludes the paper.

2 Theoretical Background

The magnitude of which price signals between markets are transmitted provides insight into the degree of market integration. This allows for conclusions regarding the capability of spatial arbitrage to soften the price and implied welfare effects of shortages or gluts. A central question in this context is whether markets share a long-run equilibrium and particularly, if this is the case, how strongly they react on deviations from it - potentially induced by a number of potential either supply- or demand-driven shocks. The magnitude of this reaction on a cross-national scale is likely to depend on various factors such as whether the spatial arbitrage crosses borders or certain countries or cities might structurally differ from the region's average which is likely to be particularly relevant for the countries analyzed in this study, as described above.

The measures of the strength of PT which represent the core quantity of the analysis are the adjustment speeds α of the linear VECM for market pairs estimated by the Johansen-procedure[3]:

$$\begin{pmatrix} \Delta p_t^A \\ \Delta p_t^B \end{pmatrix} = \begin{pmatrix} \alpha^A \\ \alpha^B \end{pmatrix} ect_{t-1} + \sum_{i=1}^{k} \mathbf{\Gamma}_i \begin{pmatrix} \Delta p_{t-i}^A \\ \Delta p_{t-i}^B \end{pmatrix} + \begin{pmatrix} \varepsilon_t^A \\ \varepsilon_t^B \end{pmatrix} \tag{1}$$

$\Delta p_{t-i}^l = p_{t-i}^l - p_{t-(i+1)}^l, i = 0,1,\dots,k; l = \{A, B\}$ denote the (lagged) price differences of markets *A* and *B*, α^l the adjustment parameters measuring the speed of adjustment of the prices p_t^l to deviations from their long-run equilibrium values $p_t^{l\,equ}$ where, e.g., $p_t^{A\,equ} = \beta_0 + \beta^B p_t^B$. These deviations are measured by the error-correction term $ect_t = p_t^l - p_t^{l\,equ}$. $\mathbf{\Gamma}_i, i = 1,\dots,k$ are *(2x2)* matrices denoting the short-run coefficients, and ε_t^l are Gaussian errors terms. For the second step of the analysis we only consider the adjustment speeds α.

Spatial arbitrage does involve costs incurring for the completion of the physical commodity transactions. These transaction costs (TC) are in general neither identical to the mere costs of transportation, i.e., the costs of freight, although the latter can be expected to account for a large share of the former, nor do they necessarily increase with distance. In spatial commodity arbitrage,

[3] The Hannan-Quinn criterion is used to determine the lag length *k* (see Lütkepohl & Krätzig 2004: 111 for a discussion of asymptotic properties of lag-selection criteria.).

various other costs incur during the performance of a physical transaction from one market to another. Barrett (2001) discusses various components of TC:

$$\tau^{AB} = fr^{AB} + v^{AB} + d^{AB} + \theta^{AB} = transfer\,costs^{AB} + \theta^{AB} \tag{2}$$

where τ^{AB} are the TC per unit between markets *A* and *B*, *fr* denotes freight rates per unit, *v* are variable costs incurring for insurance, financing, contracting and satisfying formal and informal barriers to trade, such as quality standards, *d* are the average unit duties on the product, all of which can rather be classified as transfer costs, and θ denote immeasurable TC such as opportunity or search costs.

The components of TC as formulated in decomposition (3), which might not be complete, are determined by a number of quantities. The per unit freight rate itself consists of variable and fixed costs, while the former includes fuel and maintenance and the latter includes labour and capital of the trucking companies. In the countries under review, the share of the variable costs in total transport costs is estimated at 70% (Teravaninthorn and Raballand, 2009; Zorya, 2009). The per unit fright rate might depend on distance, the quality of the transport infrastructure between the markets, and quantity transported. The efficiency and the network of traders or hired transporters (as often the case in developing countries) can also play a role.

Since TC consist of a variety of components, their effects do not necessarily enhance each other, but rather, they might also neutralize each other. Hence, the relationship between distance and the magnitude of TC does not need to be proportional. In general, TC will increase with distance, however, for distant markets that are linked by good infrastructure, e.g. by a paved highway, TC can be less than for markets located closer to each other but linked by a dirt road. Similarly, a well-developed business infrastructure will lead to less TC between a market pair in comparison with an equally distant pair with a rudimentary business infrastructure.

Most of the discussed components of TC, except distance, are likely to be neither accessible nor measurable at all. Nevertheless, they can be classified into various categories which can easily be determined for each market pair. Since distance determines a large part of freight rates in Eastern Africa, a significant *distance effect* can be expected to be observable. In the case of arbitrage crossing national borders, further costs except for freight might incur. Olper & Raimondi (2008a) classify these into costs related and unrelated to policy barriers, respectively, and the substitutability between imports and domestic produce. The magnitude of variable costs can become considerable, e.g. for meeting quality standards and presenting necessary documentation for the import. Such costs which only incur at the border crossing or due to the transnational arbitrage might constitute a significant component of TC and thus represent the determinants of the so called *border effect*. Finally, several of the components of τ are country-specific, such as the quality of transport infrastructure, subsidies or taxation of fuel and transport and institutions that influence trade (e.g. contract enforcement). Hence, a significant *country effect* might also be a relevant component of TC so that equation (3) can, for example, be rewritten as:

$$\tau^{AB} = f(d^{AB}) + D^{AB}_{border} + D^{AB}_{country\ X} \tag{3}$$

where
$$\begin{aligned} D^{AB}_{border} &= \begin{cases} 1 & if\ \ trade\ between\ A\ and\ B\ crosses\ a\ border \\ 0 & otherwise \end{cases} \\ D^{AB}_{country\ X} &= \begin{cases} 1 & if\ \ A\ or\ B\ are\ located\ in\ country\ X \\ 0 & otherwise \end{cases} \end{aligned} \tag{4}$$

and d^{AB} denotes the distance between markets *A* and *B*. Hence, TC might be thought of as some function of the distance, a border effect and a country effect. Similarly, further categories in the form of dummies can be created. Due to the above-mentioned difficulties in identifying and measuring the TC components, the inclusion of dummies represents a useful way to account for common (unknown) characteristics of markets across pairs by classifying them into easily determinable categories. We consider border dummies for market pairs involving transactions across any border

(D_B). In alternative models, we differentiate between the effects of the Kenyan-Tanzanian (D_{KT}) and the Kenyan-Ugandan (D_{KU}) border, respectively, since these effects might depend on the border. Furthermore, we consider country dummies for pairs of which both markets are located inside Tanzania (D_{Tan}) and Uganda (D_{Ug}), respectively, in order to account for country heterogeneity in PT. Finally, we include a Nairobi dummy (D_{Nai}) for pairs of which one market is Nairobi since it represents the largest city in Eastern Africa surrounded by maize deficit areas and can thus be expected to play a special role in consumption and trade.

Because we are interested in the determinants of the reaction speed on disequilibria per market pair, the sum of the estimated adjustment speeds per pair $S_\alpha^{AB} = \alpha^{AB} + \alpha^{BA}$ (hereafter *pair-wise adjustment speed*) is the relevant variable to be regressed on the potential determinants.[4] This is reasonable because the strength of PT depends on the reaction of both markets to deviations from equilibrium in each period. This idea is also the reason why only estimated adjustment speeds which are significant at the 10% level are summed up[5]. Hence, the following models are of interest, first, accounting for a general border effect D_B:

$$S_\alpha^{AB} = f(d^{AB}) + \beta_0 + \beta_1 D_B^{AB} + \beta_2 D_{Tan}^{AB} + \beta_3 D_{Ug}^{AB} + \beta_4 D_{Nai}^{AB} \quad , \tag{5}$$

and, second, distinguishing the (potentially) heterogeneous border effects:

$$S_\alpha^{AB} = f(d^{AB}) + \gamma_0 + \gamma_1 D_{KT}^{AB} + \gamma_2 D_{KU}^{AB} + \gamma_3 D_{Tan}^{AB} + \gamma_4 D_{Ug}^{AB} + \gamma_5 D_{Nai}^{AB} \quad . \tag{6}$$

We expect to find a heterogeneous border effect in the pair-wise adjustment speeds since arbitrage across borders can partly by expected to involve higher TC leading to slower PT. Hence, a potential border effect is likely to have a negative sign. Distance is also expected to matter since it is likely to

[4] Note that this measure implies a price transmission elasticity of $\beta_1 = 1$ in $ect_t = p_t^A - \beta_0 - \beta_1 p_t^A$.

[5] Most of the adjustment speeds were significant at the 1% level.

constitute a major component in the Eastern Africa maize trade. Furthermore, country effects seem likely since the policies towards agricultural production and trade are very heterogeneous[6].

3 Unknown functional form and partial linear models

In equation (6), we stated that the pair-wise adjustment speed is, among others, a function of the distance d^{AB} between the two markets. The explicit form of the functional relationship of the partial influence of distance on the adjustment speed remains unspecified since economic theory cannot suggest a particular choice. However, misspecification of the functional form can lead to serious biases of the parameter estimates. Some of the border effects literature tries to cope with this problem by estimating alternative functional specifications of this partial influence, usually both a logarithmic and a quadratic form (see, e.g., Engel & Rogers 1996, or Morshed 2003). Engel & Rogers (1996) are aware of this problem and note that "the effect of distance may also be understated if the log-distance function is not the appropriate one".

This study adopts a different approach. We do not speculate about the functional relationship between distance and the pair-wise adjustment speeds since it is likely to be too restrictive and thus potentially wrong. A semi-parametric version of model (6) is used instead (Härdle et al. 2004). The partial impact of distance is thus not restricted to take a pre-specified parametric form, but is flexibly estimated using a nonparametric estimator. In particular, a partially linear model seems appropriate since it allows modelling the impact of one or more variables in a nonparametric way and the linear partial impacts in a parametric way. In the context of this paper, the variable which is very likely to have nonlinear impact on the strength of PT is the distance while the effects have a linear impact, so that the following semi-parametric models are estimated:

[6] Note that by assessing the relationships as, e.g., formulated in models (6) and (7), we are not only determine the influencing factors of the strength of PT in Eastern African maize markets but are also capable to determine relevant components of TC in maize trade as formulated in equation (4).

$$S_{\alpha}^{AB} = m(d^{AB}) + \beta_1 D_B^{AB} + \beta_2 D_{Tan}^{AB} + \beta_3 D_{Ug}^{AB} + \beta_4 D_{Nai}^{AB} \tag{7}$$

and

$$S_{\alpha}^{AB} = m(d^{AB}) + \gamma_1 D_{KT}^{AB} + \gamma_2 D_{KU}^{AB} + \gamma_3 D_{Tan}^{AB} + \gamma_4 D_{Ug}^{AB} + \gamma_5 D_{Nai}^{AB} \quad . \tag{8}$$

where $m(\bullet)$ is some smooth function which also incorporates the intercepts of (6) and (7). It is interpreted in the same way as the dummy coefficients, namely as the partial influence of distance on the pair-wise adjustment speed, with the only difference that its functional form is not restricted, but allowed to be flexibly modelled by the information in the data. Such a semi-parametric specification has both the benefits of complete flexibility of nonlinear regression and the easy interpretability of simple fixed effects multivariate regression models. It allows for the luxury of not being forced to decide on the functional form of those explanatory variables whose impact on the explained variable might be nonlinear in some way and avoids thus the potential misspecification bias. The assumption of smoothness of the functional relationship between distance and adjustment speed is clearly much less restrictive than deciding on a particular parametric specification.

4 Data

The dataset analyzed consists of a rich collection of 16 maize wholesale price series from Kenya, Tanzania and Uganda (for the markets analyzed see Table 1 and the maps in the appendix). The time series were published by different sources and all converted into US$/t as a common unit. Most of the price data were obtained from the Regional Agricultural Trade Intelligence Network of the Eastern Africa Grain Council (RATIN, 2008). Missing values were either replaced by the converted monthly data which was published in local currencies in the report of the Michigan State University (2008), or by monthly averages of weekly data which was published in local currencies by the Ministry of Industry and Trade of Tanzania (2008) and InfoTradeUganda (2008). The distances between the markets are measured for the shortest link by national trunk roads using Google maps.

Each series ranges from January 2000 to October 2008 (106 observations). After the filling of missing values, still 59 of the total 1696 observations (3.5%), all belonging to the Tanzanian and Ugandan series, were missing. Those were imputed by using the algorithm proposed by King et al. (2001). However, due to the complexity of the data we did not adopt the multiple imputation approach suggested by them. We performed 1000 imputations for each missing value instead and estimated the most likely value from this set using Parzen's nonparametric mode estimator (Poncet 2007).

Table 1: The Markets studied

Country	City	Category
Kenya	Nairobi	net consumer
	Mombasa	net consumer
	Eldoret	net producer
	Nakuru	net producer
Tanzania	Dar es Salaam	net consumer
	Iringa	net producer
	Mbeya	net producer
	Songea	net producer
	Arusha	net producer, export center to Kenya
Uganda	Kampala	net consumer
	Iganga	net producer
	Kasese	net producer
	Lira	net producer
	Masaka	net producer
	Masindi	net producer
	Mbale	net producer, export center to Kenya

Source: Own.

The series are transformed into logged values for analysis. All series are found to be integrated of first order at the 5% level of significance, only the Songea prices are I(1) at the 7% level[7]. We consider in total 85 market pairs which we test for cointegration by the Johansen cointegration test for which the Akaike Information criterion, the Hannan-Quinn criterion and the Schwarz criterion are considered for lag length selection. Eight pairs, mostly consisting of a market in Southern

[7] Detailed results either of the time series property tests or the estimation of VECM (1) are not displayed here since they are not the focus of the paper. They may be obtained from the authors upon request.

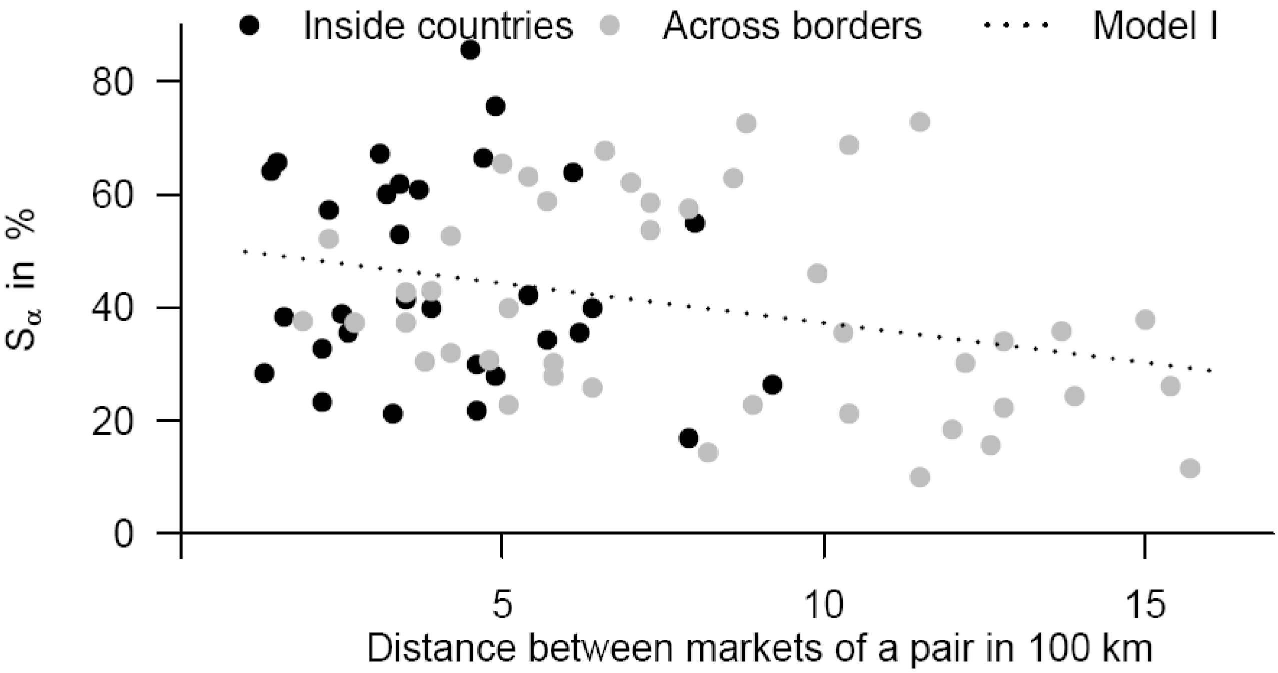
Tanzania and one of the Kenyan markets were not cointegrated at the 10% level (Iringa-Nairobi, Iringa-Eldoret, Iringa-Nakuru, Songea-Eldoret, Mbeya-Arusha, Mbeya-Nakuru, Kampala-Iganga, Kampala-Kasese). Five pairs showed mixed evidence for cointegration in the sense that the optimal lag lengths differed strongly among the criteria, and so did the test conclusions. They are considered to be cointegrated since some evidence for cointegration is found. The final regression dataset and descriptive statistics of the 77 remaining pairs are displayed in Figure 1 and Table 2, respectively. We estimate the VECM formulated in (1) for each of these pairs.[8]

Table 2: Descriptive statistics of the regression data

	Distance in 100 km	S_α^{AB} in %	D_B	D_{KT}	D_{KU}	D_{Tan}	D_{Ug}	D_{Nai}
Minimum	1.3	10.1	0	0	0	0	0	0
Median	5.4	38.1	1	0	0	0	0	0
Mean	6.5	42.2	0.58	0.20	0.38	0.08	0.26	0.19
Maximum	15.7	85.7	1	1	1	1	1	1

Source: Authors's calculations.

Figure 1: Pair-wise adjustment speed S_α vs. distance between the markets of a pair



Source: Authors' calculations.

[8] Significant adjustment parameters with the wrong sign are excluded from further analysis (3 cases).

5 Results

5.1 Parametric specifications

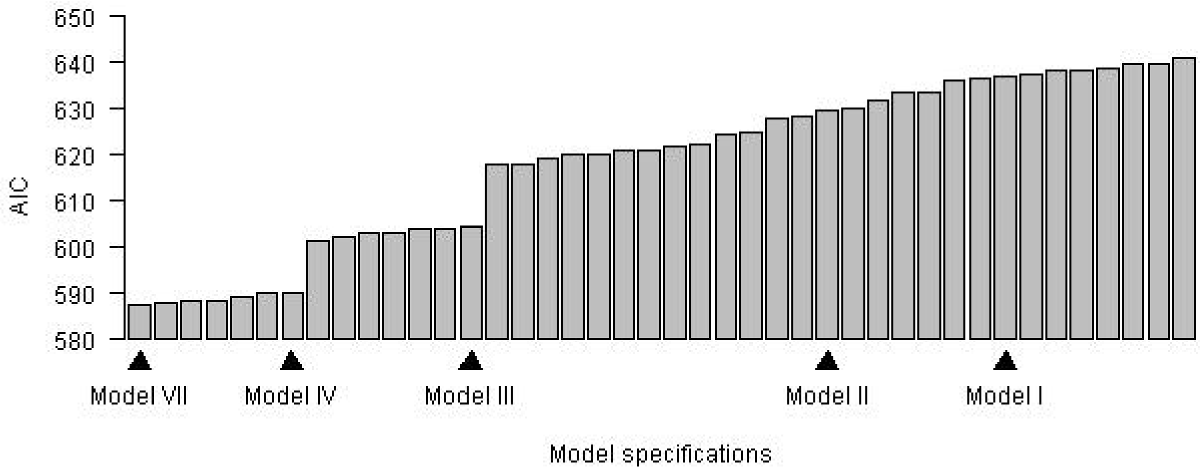
In section 4, a number of variables potentially affecting the pair-wise adjustment speed S_{α}^{AB} were identified. However, economic theory cannot a priori identify the exact set of variables which explain the par-wise adjustment best. Hence, a statistical model selection approach based on the Akaike Information Criterion (AIC) is adopted to complement the variable selection based on economic theory. This strategy enables the identification of the "best" model consisting of those variables which have the largest power in explaining S_{α}^{AB}. The "best" model is the one which explains S_{α}^{AB} with as few as possible but as many as necessary variables, i.e., which contains only the most meaningful variables but excludes all redundant variables. The AIC is in general calculated as:

$$AIC = -2\ln(L) + 2k \tag{9}$$

where L denotes the likelihood of the model and k the number of estimated parameters. The better the fit of a model, the higher the likelihood L, and thus the lower becomes the first term of the AIC. Since a better fit can simply be achieved by adding more variables, the AIC "penalizes" the number of explanatory variables by the so called penalizing term $2k$ which increases the AIC. Variables of little explanatory power lead thus to a larger AIC, i.e., a decrease in the quality of the model. Consequently, the best model is the one with the lowest AIC.

Figure 2 displays the AIC values versions of the parametric models (6) and (7) in which $f(d^{AB})$ is either assumed to be a linear, quadratic or logarithmic function. 43 model specifications, of which only a selection is presented in detail here, are estimated and ordered according to their AIC values in increasing order. The model quality increases continuously from right to left except for two noticeable jumps. The first jump occurs for model III and the six models to its left as the AIC markedly increases since either D_{KT} and D_{Nai} or D_{KT} and D_{Tan} were both, among others, included into the model. The second strong decrease of the AIC occurs with model IV and the six specifications to its left as the three effects D_{KT}, D_{Tan} and D_{Nai} are included.

Figure 2: AIC for various parametric model specifications



Source: Authors' calculations.

Table 3 displays the OLS parameter estimates, their p-values and the AIC of selected specifications. In Figure 1, a clear functional relationship between distance and S_{α}^{AB} becomes apparent (the dotted line). Not accounting for further explanatory variables, the linear influence of distance is strongly significant (column I), however, in a quadratic model of distance (results not displayed here) only the intercept turns out to be significant. Model II additionally considers a general border effect (the coefficient of D_B) and a Nairobi effect (the coefficient of D_{Nai}). Distance remains significant and the Nairobi dummy appears to have a large significant positive impact on the pair-wise adjustment speed while a significant border effect is not found.

Table 3: Estimation results of various parametric model specifications[a]

		Model						
		I	II	III	IV	V	VI	VII
Explanatory variables	Intercept	51.3 (<0.001)	49.4 (<0.001)	43.3 (<0.001)	48.1 (<0.001)	42.2 (<0.001)	48.1 (<0.001)	46.9 (<0.001)
	Distance d	-1.4 (0.001)	-1.3 (0.028)	0.5 (0.779)	-0.4 (0.446)	2.0 (0.225)		
	d^2			-0.1 (0.543)		-0.2 (0.131)		
	$\ln(d)$						-0.9 (0.720)	
	D_B		-3.1 (0.505)					
	D_{KT}			-22.5 (<0.001)	-26.6 (<0.001)	-26.0 (<0.001)	-28.0 (<0.001)	
	D_{KU}			2.5 (0.535)				
	D_{Tan}				-20.7 (<0.001)	-22.7 (<0.001)	-21.3 (<0.001)	
	D_{Ug}				1.5 (0.711)			
	D_{Nai}		16.9 (0.001)	17.2 (<0.001)	16.5 (<0.001)	15.2 (<0.001)	16.0 (<0.001)	15.2 (<0.001)
	$D_{TanGen} = D_{Tan} + D_{KT}$							-26.5 (<0.001)
	AIC	637.0	629.6	604.7	590.2	587.9	589.3	586.9

Source: Authors' calculations.

[a] The numbers in parentheses denote the p-values of the estimates.

If the general border dummy is split into one dummy for each border (model III), a strongly significant border effect for the Kenyan-Tanzanian border is found while the dummy for the Kenyan-Ugandan border is not significant, and can thus be dropped from the model. The Nairobi dummy and the intercept stay highly significant; the partial influence of distance, either modelled in linear or quadratic form, becomes insignificant. Obviously, this model is much better than the two previous ones because its AIC is more than 25 points lower than for model II (see Figure 2).

In model IV, the significance of country-specific effects is checked for. The dummy for market pairs inside Uganda is not significant while pairs in Tanzania show a strongly significant and lower pair-wise adjustment speed; the AIC further decreases by almost 15 points to 590 (see Figure 2). Models

V and VI use the significant variables of model IV and check for a quadratic and logarithmic partial impact of distance on S_{α}^{AB} showing that neither specification renders a significant partial relationship; the AIC values are slightly below 590. Distance modelled in these parametric forms thus appears to be redundant and does not have to be considered in the final model while a significant border and country effect are identified[9]. The coefficients of both dummies are negative and similar in size so that a Wald test is carried out in order to test whether the coefficients of these two effects equal, i.e., whether both variables can be summarized into a general Tanzania dummy $D_{TanGen} = D_{Tan} + D_{KT}$ (see Figure 5 in the appendix). The corresponding test statistic has a value of 1.4 corresponding to a p-value of 0.25 which is clearly not significant[10]. Hence, the two Tanzania effects can be summarized into one, and model VII appears to be the best parametric model having an AIC of 586.9[11].

5.2 Semi-parametric specifications

Although distance appeared to be insignificant in the above estimated parametric models, it is included in the semi-parametric specifications in order to evaluate whether it has a potentially nonlinear impact on the pair-wise adjustment speeds. Based on the results in Table 3, we omit all insignificant dummy variables and estimate the following partially linear model using the Speckman estimator (Speckman 1988, Härdle et al. 2004):

$$S_{\alpha}^{AB} = m(d^{AB}) + \beta_1 D_{Tan}^{AB} + \beta_2 D_{KT}^{AB} + \beta_3 D_{Nai}^{AB} \qquad \text{(VIII)}$$

and

[9] Note that distance might, however, have a significant partial impact of another functional form which might be the reason why the usual forms of modelling its partial influence do not become significant.

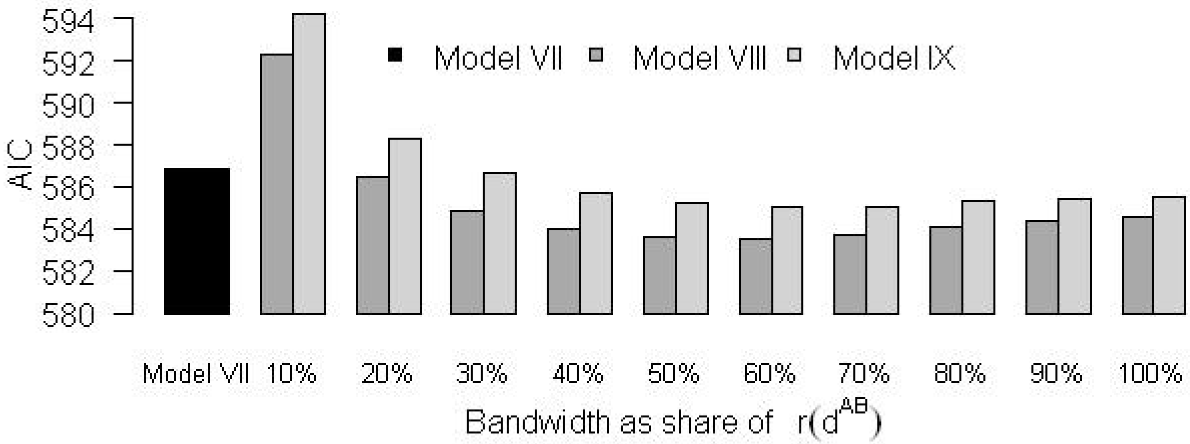
[10] A similar test for the coefficient equality of the Nairobi and the general Tanzania dummy is strongly rejected with a test statistic of 73.4 and a p-value smaller than 0.001.

[11] Note that the decrease of the AIC, indicating a significantly better model, corresponds to the result of the Wald test, which suggests that both variables should more adequately be combined since their coefficients do not significantly differ from each other. This illustrates that both model selection approaches lead to identical results.

$$S_{\alpha}^{AB} = m(d^{AB}) + \gamma_1 D_{TanGen}^{AB} + \gamma_2 D_{Nai}^{AB} \qquad \text{(IX)}$$

which correspond to models (VI) and (VII). Since the estimation results of the semi-parametric model depend on the bandwidth h which governs the degree of smoothing of its nonparametric part, $m(d^{AB})$, different bandwidths have been employed for a robustness check as illustrated in Figure 3 and Figure 4[12]. For most of the bandwidths, either specification represents a better model than the best parametric model VII (Table 3, AIC of 586.9). Furthermore, it is evident that the inclusion of the distance in a nonparametric fashion improves the quality of the model strongly. Distance is thus shown to have a significant partial nonlinear impact on the pair-wise adjustment speeds. Consideration of this variable thus markedly improves the quality of the model.

Figure 3: AIC for various bandwidths of the semi-parametric specifications vs. model VII



Source: Authors' calculations.

The AIC of both models appears to be fairly stable for bandwidths larger than one third of the range of the distance (appr. 4.3) which indicates that the choice of the bandwidth in this range does almost not matter[13]. Furthermore, for each bandwidth specification IX has a lower AIC than specification VIII and is thus to be preferred. For a bandwidth of 60% of $r(d^{AB})$, i.e., an optimal bandwidth of

[12] We use ten bandwidths for each model which are calculated as the deciles of the range of the distance data $r(d^{AB}) = \max(d^{AB}) - \min(d^{AB}) = 14.4$. The higher the bandwidth, the smoother the nonparametric estimate will appear.

[13] The graph also shows that a semi-parametric model is not a priori superior of the corresponding parametric one.

h_{opt}= 8.6, the AIC of specification IX takes the lowest value among all considered models (583.5) which is more than 3 points less than the AIC of the best parametric model (586.9).

Consequently, the partially linear model IX with a bandwidth of 8.6 turns out to be the best model to describe the data with the following estimates (p-values underneath in parentheses):

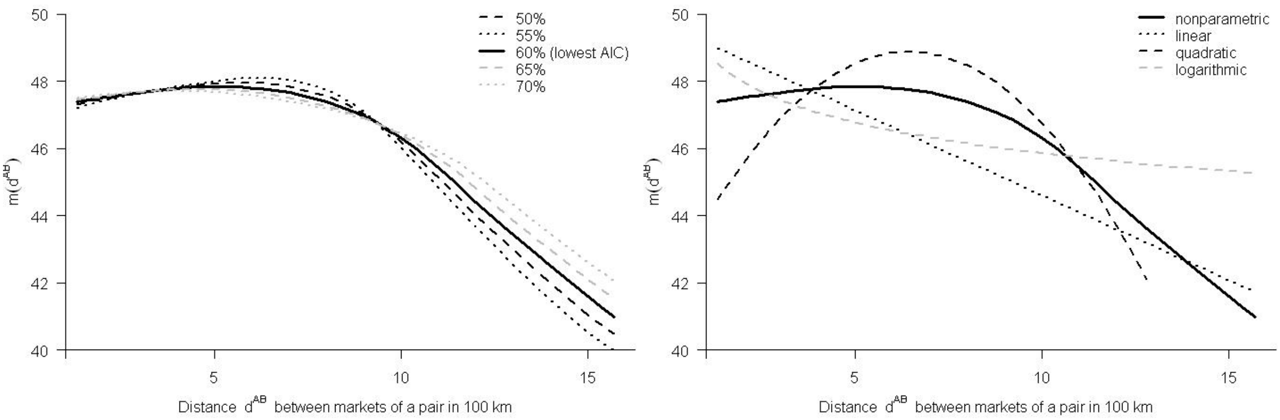
$$S_{\alpha}^{AB} = m(d^{AB}) - \underset{(<0.001)}{24.9} \cdot D_{TanGen}^{AB} + \underset{(<0.001)}{15.2} \cdot D_{Nai}^{AB} + \varepsilon^{AB} \quad . \tag{10}$$

where the estimated nonlinear impact of distance can be inspected in the left panel of Figure 4. The estimates of the parametric part are very similar to those of model VII, however they are free from potential misspecification bias due to the unknown functional relationship between distance and adjustment speed. A significant and strongly negative general Tanzania effect of almost 25 percentage points and a similarly strong, but positive Nairobi effect of approximately 15 percentage points, are detected.

The left panel of Figure 4 illustrates the estimated nonlinear relationship $m(d^{AB})$ for a range of bandwidths suggesting that the estimated relationship appears to be fairly stable[14]. The relationship can be characterized in the following way. Holding all other factors equal, pairs of markets which are located close to one another adjust on average roughly 47% of the deviation from the long-run equilibrium in the subsequent period. Interestingly, this adjustment increases slightly by one percentage point up to a distance of 600 km. The more the distance between a pair increases beyond this value, the stronger the expected pair-wise adjustment speed S_{α}^{AB} decreases. Beyond 1100 km, the change of the adjustment becomes almost constant with a rate of almost exactly one percentage point per 100km. Consequently, market pairs lying 1500 km apart from each other are expected to correct only 42% of the equilibrium deviations (see also Figure 5 in the appendix).

[14] The estimate of nonlinear relationship is plotted for the optimal bandwidth together with the data in Figure 5 in the appendix.

Figure 4: The estimated nonlinear relationship between distance and the pair-wise adjustment speeds



Source: Authors' calculations.

The right panel of Figure 4 plots the best semi-parametric versus several parametric specifications of $m(d^{AB})$. Visual inspection suggests that the estimated nonparametric function differs strongly from the parametric functions. Härdle et al. (2004) propose a bootstrapped modified likelihood ratio test which allows for the testing of this question, i.e., the null hypothesis that the function $m(\bullet)$ is a specified parametric function against the alternative that it is an arbitrary smooth function. We perform the test for three functional forms: a linear, a quadratic and a logarithmic functional relationship between d^{AB} and S_{α}^{AB} against the estimated nonlinear relationship of the optimal bandwidth of 8.6 by using 1000 bootstrap replications for each test.

Table 4: Test results various parametric specifications of the nonlinear relationship against a smooth function

Null hypothesis	Test statistic	p-value
Linear: $m(d^{AB}) = \beta_0 + \beta_1 d^{AB}$	29.3	<0.001
Quadratic: $m(d^{AB}) = \beta_0 + \beta_1 d^{AB} + \beta_2 \left(d^{AB}\right)^2$	5.4	<0.001
Logarithmic: $m(d^{AB}) = \beta_0 + \beta_1 \log\left(d^{AB}\right)$	76.5	<0.001

Source: Authors' calculations.

The test results in Table 4 show that the statistics of all tests are highly significant, i.e., the null hypothesis has to be rejected in each case. Hence we obtained strong econometric evidence that modelling the relationship between distance and the pair-wise adjustment speed by a linear, quadratic or logarithmic parametric function is inadequate, at least for the data set analyzed here.

6 Discussion

Both the parametric and the semi-parametric model find a significant strongly negative country effect for market pairs which are either located in Tanzania or across the country's border. This finding is consistent with the trade policy and the larger size of the country and it suggests that inside Tanzania and at its borders considerable trade frictions exist. Holding all other factors equal, this general Tanzania effect reduces the expected pairs-wise adjustment speed by about 25 percentage points, which is a very strong effect. The isolated effect of crossing Tanzania's border, which might be caused by the frequent export bans, is numerically stronger than the isolated "within-Tanzania-effect". The latter effect is mainly caused by underdeveloped infrastructure and domestic non-tariff trade barriers, however, both effects do not significantly differ from each other. Hence the substantial border effect is absorbed by a "general-Tanzania-effect" of approximately the same magnitude. This result indicates that all market pairs which involve at least one Tanzanian market are significantly less integrated than the remaining pairs in these three countries. Based on the ideas developed in section 2, this points to the fact that TC inside the country and for crossing its borders are significantly higher than for spatial arbitrage inside or between Kenya and Uganda, even though transport costs per ton-km are very similar in all three countries (Zorya 2009). The major causes of these effects are the national export bans and the poor transport infrastructure combined with the strong drive to ensure nation-wide food-security resulting in strong price and trade distortions on the local up to the national level. The identification of further Tanzania-specific factors lies beyond the scope of this analysis. However, this issue is of enormous interest from a political point of view since detailed insights into the causes can lead to policy-oriented conclusions on how to decrease barriers to trade and improve the trade infrastructure in and at Tanzania's borders. The country's maize markets appear to be integrated to a much lower extent in comparison with the other two EAC countries, TC are considerably increased.

A general border effect could not be found. However, the distinction between the Kenyan-Tanzanian and Kenyan-Ugandan borders appears to be important. This appears to be plausible in light of the strongly differing trade policies pursued by Tanzania and Kenya. While the crossing of the former border strongly reduces market integration, no effect is found for the latter. This provides some indication that a border effect might not (always) be thought of as a general effect no matter which border is crossed but rather as a heterogeneous effect which might even differ for different parts of the border of one country, as here in the case of Kenya[15]. Again, further research on the reasons for this heterogeneity is very relevant from a political point of view but lies beyond the scope of this paper. Heterogeneity of the border effect, to our knowledge, has attracted almost no attention in the border effects literature which it might a further interesting area of research[16].

Besides the strong negative Tanzania effect, a strongly positive Nairobi effect is identified. The price transmission between market pairs, which involve Nairobi, is on average markedly higher by about 15 percentage points. Interestingly, it does not matter whether Nairobi's partner market is located in Kenya or abroad. Such a result seems very plausible due to the size and economic role of the city in Eastern Africa and in the light of Kenya‘s trade policy. Since it is the capital of Kenya and by far its largest city, it is well-connected by transport infrastructure not only to the rest of the country and especially its agricultural production regions but also to most of the production regions in the neighbouring countries, which lies in the interest of both consumers in the city and producers in the rural regions and abroad. Another factor might lie in its attractiveness for the sale of staple food due to the number and wealth of consumers. In Nairobi, traders can expect to sell larger quantities, maybe even for higher prices around the year in contrast to many small and middle-sized towns in

[15] It seems plausible that such heterogeneity might be more relevant in the case of developing countries.

[16] Chen (2004), for example, considers country-specific border effects which are, however, thought to be homogenous for each country.

the region so that they might specialize in supplying the city. Transporters can expect frequent movements of goods and people to and from the city being faced with much less frequent empty backhaul which reduces costs and increases competition, in whose consequence scale effects and spill-overs are likely to be realized. Moreover, business infrastructure, as shown by Helble (2007) for Europe, might play a crucial role in the context of Nairobi. Mainly due to its economic importance, the condition of infrastructure relevant for trade and business can be expected to be above the region's average so that the city can be reached with less effort and costs than other towns located equally distant from a production region. Such factors, for example, are likely to reduce the costs of spatial arbitrage with Nairobi so that even slight price differences might induce trade with very little delay yielding highly integrated market pairs.

A further explanation seems reasonable before the given context of Nairobi and border effects. In a case such as Nairobi as the heavyweight in the EAC, borders defined by administrative units might be much less relevant than borders defined by economic markets. Although a region located in one of the neighbouring countries and trade to Nairobi has to cross a border, the pair-wise adjustment speed is expected to be on average 15 percentage points higher than for equally distant pairs without Nairobi. This might be the case if the former pair belongs to one "economic entity" in contrast to the latter pair where spatial arbitrage has to cross the border(s) between two (or more) economic entities. Such a view is supported by the strongly negative effect for PI inside Tanzania because the country, while belonging to an administrative entity, might be split into several economic entities, given its large size, poor infrastructure, and barriers to trade at the local and national levels.[17]

Fackler & Goodwin (1997: 979) stress that "it is important, however, to distinguish between the term market integration and other forms of integration." They mention as an example markets which are "'economically integrated' in the sense that there are no border restrictions restricting the flow of goods". The aim of the customs union of the EAC is clearly to establish both economic and market

[17] Various authors, such as Gardner & Brooks (1993) and Berkowitz & DeJong (1999, 2000) find evidence for internal border effects in the context of East European transition countries.

integration. Due to the Tanzanian border measures, however, the union might not even be economically integrated and the country appears to be a rather isolated and internally fragmented island within the EAC regarding both economic and market integration for maize which is obviously the price the Tanzanian government has to pay for its political objectives. This, of course, needs not to be seen negatively insofar as the objectives of ensuring nation-wide food security have been met.

Both the parametric and semi-parametric model give very similar results regarding the partial impacts of the two above discussed effects. Nevertheless, the evidence obtained concerning the role of distance differs strongly between the approaches. While distance is not significant neither in linear, quadratic or logarithmic form, and thus not to be included into the optimal parametric model, it does, by including it in a nonparametric way, improve the quality of the model markedly. How do these seemingly contradicting results go along with each other? As shown above, distance has a significant linear impact on the pair-wise adjustment speed if it is assumed to be only explanatory variable. However, such a specification is far from being the optimal model. In this, a strongly negative Tanzania and a strongly positive Nairobi effect turn out to have the relevant explanatory power in whose consequence the partial impact of distances fades away. This becomes plausible when the information contained in the two effects is considered. Both dummies contain, among others, a certain amount of the partial influence of distance on pair-wise adjustment speeds. Due to its central location in the region, distances between market pairs including Nairobi are rather short while the general Tanzania-effect includes a number of market pairs which are located more than a thousand kilometres away (the pairs consisting of markets in Kenya and Tanzania, see Figure 5 in the appendix). The partial effect of distance which appears to be nonlinear (Figure 4 and Figure 5) turns out to be not significant in the parametric model since its functional form is obviously inadequately approximated by parametric functional forms usually used in the border effects literature.

7 Conclusion

This study examines maize trade in the three largest member countries of the East African Community, Kenya, Tanzania and Uganda, by analyzing 85 market pairs. Although they are in many respects very similar to each other and are located close to each other in space, the policy attitudes towards the agricultural sector and consequently to agricultural trade differ considerably. The study focuses on the factors determining the magnitudes of price transmission, i.e., the magnitude of price reaction in response to deviations from their equilibrium, and reaches results of strong political interest. A range of parametric and semi-parametric models are estimated. The model selection approach followed in this paper greatly facilitates and supports the correct identification of the relevant effects and identifies the semi-parametric model as the most appropriate description of the data.

A significantly negative general effect for Tanzania, which includes a substantial border effect, and a significantly positive effect for Nairobi are found. Distance turns out to have a nonlinear partial impact on the price reaction which could only be revealed by the semi-parametric model. The large negative effect for market pairs inside Tanzania is consistent with its inward-oriented policies towards agriculture and agricultural trade and poor infrastructure. Similarly, the strongly positive Nairobi effect appears to be very plausible in light of the size and economic role of the city in Eastern Africa, as well as the structural maize deficit in Kenya. The absence of an effect for the Kenyan-Ugandan border appears to be consistent with the outward-oriented trade policies of both countries.

The results of the analysis are strongly consistent with the country-specific agricultural and trade policies and the state of the infrastructure. Tanzania appears to be a rather isolated and internally fragmented island within the customs union of the East African Community, a situation which is markedly different for Kenya and Uganda which are both internally and across the border well integrated with high rates of price transmission. The results are thus of large political importance since they give indications regarding the structure of the integration of maize markets in the East African Community. However, further research has to be carried out in order to identify and disentangle the influences of particular determinants causing the identified effects.

References

Anderson JE & E van Wincoop, 2004. Trade Costs. *Journal of Economic Literature* 42(3): 691–751.

Awuor T, 2007. Review of Trade and Markets Relevant to Food Security in the Greater Horn of Africa. A special report by the Famine Early Warning Systems Network (FEWS NET) for USAID.

Barrett CB, 2001. Measuring Integration and Efficiency in International Agricultural Markets. *Review of Agricultural Economics* 23(1): 19-32.

Berkowitz D & DN DeJong, 1999. Russia's Internal Border. *Regional Science and Urban Economics* 29(5): 633 - 649.

Chen N, 2004. Intra-national versus international trade in the European Union: why do national borders matter? *Journal of International Economics* 63: 93–118.

Delgado C & Minot N, 2000. Agriculture in Tanzania since 1986: Leader or Follower of Growth. World Bank Country Study, Washington DC, USA.

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

Evans C, 2003. The Economic Significance of National Border Effects. *American Economic Review* 93(4): 1291-1312.

FAOSTAT, 2009. International Agricultural Statistics: Food Balance Sheet. FAO, Rome, Italy.

Feenstra R, 2002. Border effects and gravity equation: consistent methods for estimation. *Scottish Journal of Political Economy* 49(5): 491 - 506.

FEWSNET, 2009. Maize Production and Market Flows in Eastern Africa. Famine Early Warning System Network, USAID, Washington DC, USA.

Furtan WH & BM van Melle, 2004. Canada agricultural trade in North America: Do national borders matter? *Review of Agricultural Economics* 26(3): 317–331.

Gardner BL & KM Brooks, 1994. Food prices and market integration in Russia. *American Journal of Agricultural Economics* 76: 641-646.

Härdle W, Müller M, Sperlich S & Werwatz A, 2004. Nonparametric and Semiparametric Models. Heidelberg: Springer Verlag.

Helble M, 2007. Border Effect Estimates for France and Germany Combining International Trade and Intranational Transport Flows. *Review of World Economics* 143(3): 433-463.

InfoTradeUganda, 2008. http://www.infotradeuganda.com. Accessed November 2008.

Jayne T, Myers R & Nyoro J, 2008. The Effects of NCPB Marketing Policies on Maize Market Prices in Kenya. *Agricultural Economics* 38: 313-325.

King G, Honaker J, Joseph A & Scheve K, 2001. Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review* 95(1): 49-69.

Lütkepohl H & Krätzig M, 2004. Applied Time Series Econometrics. Cambridge: Cambridge University Press.

McCallum J, 1995. National Borders Matter: Canada-U.S. Regional Trade Patterns. *American Economic Review* 85(3): 615-623.

Meijerink G, Roza P & S van Berkum, 2009. East African Governments' Responses to High Cereal Prices. Report 2009-102. LEI Wageningen, The Hague, The Netherlands.

Michigan State University, 2008. Strengthening Food Security in Sub-Saharan Africa through Trade Liberalization and Regional Integration. Country Annexes. Maize Market Sheds in Eastern and Southern Africa. Report prepared for the World Bank.

Ministry of Industry and Trade of Tanzania, 2008. http://www.mitm.go.tz/ Accessed November 2008.

Morshed AKMM, 2003. What can we learn from a large border effect in developing countries? *Journal of Development Economics* 72: 353-369.

Olper A & Raimondi V, 2008a. Agricultural market integration in the OECD: A gravity-border effect approach. *Food Policy* 33: 165–175.

Olper A & Raimondi V, 2008b. Explaining National Border Effects in the QUAD Food Trade. *Journal of Agricultural Economics* 59(3): 436–462.

Olper A & Raimondi V, 2009. Patterns and Determinants of International Trade Costs in the Food Industry. *Journal of Agricultural Economics* 60(2): 273–297.

Poncet P, 2007. modeest: Mode Estimation -- Chernoff Distribution. R package version 1.06.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.

RATIN, 2008. Regional agricultural trade intelligence network of the Eastern Africa Grain Council, http://www.ratin.net/. Accessed November 2008.

Speckman PE, 1988. Regression analysis for partially linear models. *Journal of Royal Statistical Society* Ser B 50: 413-436.

Teravaninthorn S and G Raballand, 2009. Transport Prices and Costs in Africa: A Review of the Main Trade Corridors. The World Bank, Washington DC, USA.

Temu A, Manyama A & Temu A, 2007. Alternative Staple Food Trade and Marketing Policy Interventions for Tanzania: A Country Level Assessment of the Maize Sub-Sector. Draft Report prepared for the FAO, Dar es Salaam, Tanzania.

United Nations, 2009a. Cartographic Section, Department of Field Support. http://www.un.org/Depts/Cartographic/english/htmain.htm. Accessed January 2009.

United Nations, 2009b. United Nations Commodity Trade Statistics Database, Statistics Division, http://comtrade.un.org/ Accessed January 2009.

University of Texas, 2009. Perry-Castañeda Library Map Collection/ Africa Maps, http://www.lib.utexas.edu/maps/africa.html Accessed January 2009.

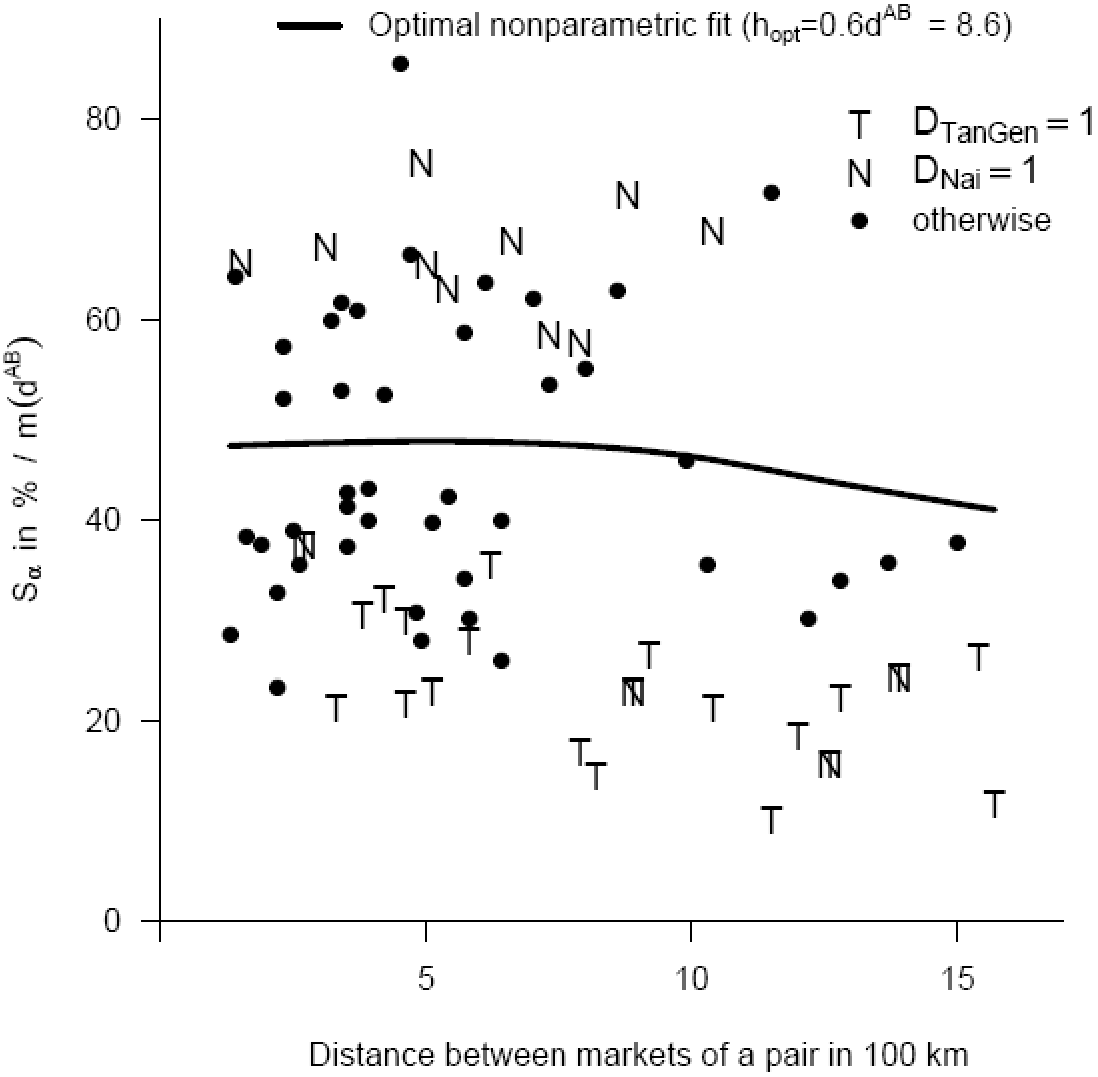
World Bank, 2008a. Regional Trade in Food Staples: Prospects for Stimulating Agricultural Growth and Moderating Short-Term Food Security Crises in Eastern and Southern Africa. AFTAR Report, Draft Version on 11/27/2008, Washington DC, USA.

World Bank, 2008b. Non-Tariff Measures on Goods Trade in the East African Community. World Bank Report No. 45708-AFR, PRM2-AFCRI. Washington DC, USA.

Zorya S, 2009. Eastern Africa: A Study of the Regional Maize Market and Marketing Costs. World Bank Report No. 49831 – AFR. Washington DC, USA.

Appendix

Figure 5: The estimated nonlinear partial impact of distance on S_{α}^{AB} and the data



Source: Authors' calculations.

Figure 6: Map of Kenya

Source: United Nations (2009a) and authors.

Figure 7: Map of Tanzania

Source: United Nations (2009a) and authors.

Figure 8: Map of Uganda

Source: University of Texas (2009) and authors.