Courant Research Centre
'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

Georg-August-Universität Göttingen
(founded in 1737)

Discussion Papers

No. 54

Price Dynamics in Tanzanian Maize Markets:
Insights from a Semiparametric Cointegration Model

Rico Ihle, Stephan von Cramon-Taubadel

December 2010

Price Dynamics in Tanzanian Maize Markets:

Insights from a Semiparametric Cointegration Model

Rico Ihle and Stephan von Cramon-Taubadel

Department of Agricultural Economics and Rural Development
Georg-August-Universität Göttingen, Germany
rihle@gwdg.de

Abstract: Maize is a major staple food in Sub-Saharan Africa. Monthly maize prices in Tanzania are analyzed since the country is an important maize producer and exporter in East Africa. We analyze price transmission between the five most important urban regions of Tanzania between 2000 and 2008 which correspond to major maize production or consumption areas. We propose a novel method for the analysis. The semiparametric vector error-correction model allows the partial impact of the past deviations from price equilibria on current price changes to be potentially nonlinear. The nonparametric estimates of these partial influences suggest that they can be adequately modeled by linear functions.

Keywords: cointegration, maize, nonlinear time series model, price transmission, semiparametric model, Tanzania.

JEL: C32, Q11, Q13

* Rico Ihle is research associate (corresponding author, rihle@gwdg.de) and Stephan von Cramon-Taubadel is professor at the Department for Agricultural Economics and Rural Development of the Georg-August-Universität Göttingen.

Price Dynamics in Tanzanian Maize Markets:
Insights from a Semiparametric Cointegration Model

Introduction

In Tanzania, as in most of Sub-Saharan Africa, maize is one of the major staple foods. It is by far the most important source of carbohydrates for the vast majority of the population. The availability of this staple food is of outstanding importance with respect to a number of socio-economic and political issues which range from political instability in the form of riots due to maize price increases, to famines as experienced in 2009 in East Africa. Food security represents one of the central Millennium Development Goals (MDG) formulated by the United Nations. The MDG Africa Steering Group (2008) notes that per capita food production in Sub-Saharan Africa has decreased during the last three decades.

Besides improvements in food production, well functioning trade relationships which are based on the transmission of price signals between shortage and surplus regions are crucial components in this context. The supply of the product via both domestic trade from production to consumption areas inside a country along with trade between countries in the region or with the world market is important. Since East Africa is exposed to the risk of famines, development cooperation organizations have set up special information services which are aimed at facilitating food trade (see, for example, http://www.ratin.net/).

We focus on maize markets because maize, as a major tradable staple food, plays a crucial role in this respect (FAOSTAT, 2009). Furthermore, we study markets in Tanzania since it is not only the most populated country in East Africa, but also an important regional maize producer (World Bank, 2009). Our analysis assesses the speed at which price signals are transmitted between major maize markets in Tanzania. Moreover, the model which we employ enables us to obtain insights into the nature of the price transmission. It provides

information on how price reactions are related to the magnitude of deviations from long-run price equilibria without assuming a particular functional form *a priori*.

A number of studies analyze price transmission in maize markets in Tanzania. Among the more recent ones are Sarris and Mantzou (2005), van Campenhout (2007), Michigan State University (2008), Rapsomanikis and Karfakis (2008), and World Bank (2009).

Our work contributes to the scientific discussions of three lines of research. First, it sheds further light on the important issue of the magnitude of the transmission of prices of staple foods in East Africa. Second, it addresses the question of the identification of the form of non-linearities which characterizes the data analyzed. The proposed semiparametric model generates insight into the nature of price transmission without assuming a potentially inappropriate model form. The estimates of the magnitude of price transmission are thus more accurate because they reflect the price transmission processes hidden in the data, for example whether they can plausibly be assumed to be linear, to follow threshold behavior or another functional form. Third, our analysis touches upon the issue of whether data-driven or theory-driven model selection approaches are preferable.

The Data and the Market

Apart from Kenya, Tanzania is one of the largest maize producers in Eastern Africa. Since the country produces more than it consumes, maize is often exported to the neighboring countries among which Kenya is the main destination (see, e.g., World Bank, 2008; Michigan State University, 2008). Domestic maize trade in Tanzania is characterized by pronounced disparities in net production so that prominent trade flows between the country's regions occur (FewsNet, 2009). We assess price transmission between the main maize markets of Tanzania with a flexible methodology in order to obtain evidence on the nature, that is, the functional form, of the reaction of price changes in the current period to deviations from the equilibrium price in the previous period.

We consider the five most important markets for maize in Tanzania which are Dar es Salaam (Dar), Arusha (Ar), Iringa (Ir), Mbeya (Mb), and Songea (So). While the capital Dar es Salaam, which is also by far the largest urban agglomeration of the country, lies in the middle of the country on the coast of the Indian Ocean, and Arusha as a major production and trade center for exports to Kenya lies in the North, the latter three cities are all located in the southern part of Tanzania and constitute major net maize production regions. The monthly price series start in January 2000 and last until September 2009 (n=105). The data was obtained from the Regional Agricultural Trade Intelligence Network of the Eastern Africa Grain council (RATIN, 2008). Missing observations were either replaced by monthly averages based on data of the Ministry of Industry and Trade of Tanzania (2008) or imputed using a version of an algorithm proposed by King et al. (2001) and the respective R-package programmed by Honaker et al. (2009). The prices are used in logged form as depicted in **Fehler! Verweisquelle konnte nicht gefunden werden.**.

Methodology

In the context of price transmission analysis, the vector error-correction model (VECM) is the basic model which has been extended into several directions over the last two decades. A crucial assumption of this model is the constancy of its parameters. In particular, this model assumes that the adjustment towards equilibrium is linear, that is, the parameters quantifying the adjustment are assumed to be constant. In most analyses in this field, the mechanism characterizing the nature of price transmission is explicitly chosen by the researcher. In some cases, a traditional linear VECM is used, in others price transmission is allowed to be nonlinear, often using regime-dependent versions of the VECM. A popular regime-dependent model is the threshold vector error correction model (TVECM, proposed by Goodwin and Piggott, 2001) which assumes that the switching between regimes takes place if the indicator variable, which is most often the magnitude of the deviation from the long-run equilibrium,

crosses one of the two threshold values which are estimated together with the other model parameters. Ihle et al. (2009) suggest the Markov-switching VECM for modeling regime-dependence in spatial price relationships in the form of recurring structural breaks.

We propose a novel and very flexible model class recently developed by Gaul and Theissen (2008) for analyzing nonlinearities in the transmission between stock prices. It encompasses several model classes, e.g. the linear VECM and the TVECM, as special cases. Its econometric structure corresponds to the semiparametric partial linear model (PLM). It models the adjustment of the equilibrium errors flexibly via nonparametric techniques. The autoregressive short-run parameters, however, are regarded as constant. Thus, their impact is modeled to be linear, i.e., not regime-dependent. Since one part of the model is modeled nonparametrically while the rest is assumed to be parametric, it is called the semiparametric vector error correction model (SPVECM).

To our knowledge, the SPVECM has only been applied by Gaul (2008) and Gaul and Theissen (2008). However, several studies of price transmission analysis have used similar approaches based on non- and semiparametrics. Mancuso et al. (2003) analyze interest rates using a local linear regression estimator without regarding short-run dynamics alternatively to a threshold autoregressive model. Serra et al. (2006) follow a similar methodological approach to analyze European pork markets. Goodwin and Vavra (2009) apply a Nadaraya-Watson estimator to a VECM which they assume does not have short-run dynamics.

The VECM has the following structure:

(1) $\Delta p_t = \alpha \beta' p_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta p_{t-i} + \varepsilon_t = \text{error corr}^{VECM} + \sum_{i=1}^{k} \Gamma_i \Delta p_{t-i} + \varepsilon_t$

where p_t denotes a v-dimensional vector of logged prices and ε_t a v-dimensional Gaussian error term. This means that the current price change Δp_t is a linear function of the past equilibrium errors $\beta' p_{t-1}$ and past price changes Δp_{t-i} in the sense that the parameters α and

4

Γ_i are constant.[1] Hence, the error-correction term of the VECM is quantified by the following matrix

$$(2) \text{ error corr}^{VECM} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & \ddots & \vdots \\ \alpha_{v1} & \cdots & \alpha_{vr} \end{pmatrix} \beta' p_{t-1}.$$

β is the $(v \times r)$ matrix of the r cointegration relationships, the $(v \times r)$ loading matrix α contains the adjustment speeds at which the past equilibrium errors are corrected from period to period (error-correction rates). Γ_i are matrices of dimension $(v \times v)$ that quantify the impact of past price changes on the current price change (short-run dynamics) . This model can be more compactly rewritten as:

$$(3) \ Z_{0t} = \alpha \beta' Z_{1t} + \Gamma Z_{2t} + \varepsilon_t$$

where $Z_{0t} = \Delta p_t, Z_{1t} = p_{t-1}, \Gamma = (\Gamma_1, \cdots, \Gamma_k)$ and $Z_{2t} = (\Delta p'_{t-1} \cdots \Delta p'_{t-k})'$ have dimensions $(v \times 1), (v \times 1), (v \times vk)$ and $(vk \times 1)$, respectively.

The semiparametric VECM generalizes this model structure by allowing the loading matrix α to be a smooth function instead of a constant, that is:

$$(4) \ \Delta p_t = m(\beta' p_{t-1}) + \sum_{i=1}^{k} \Gamma_i \Delta p_{t-i} + \varepsilon_t = \text{error corr}^{SPVECM} + \sum_{i=1}^{k} \Gamma_i \Delta p_{t-i} + \varepsilon_t.$$

Hence, the linear impact of past price changes (captured by Γ_i) is maintained, but error-correction behavior is captured by an arbitrary smooth function m so that:

$$(5) \text{ error corr}^{SPVECM} = \begin{pmatrix} m_{11}(\beta' p_{t-1}) & \cdots & m_{1r}(\beta' p_{t-1}) \\ \vdots & \ddots & \vdots \\ m_{v1}(\beta' p_{t-1}) & \cdots & m_{vr}(\beta' p_{t-1}) \end{pmatrix}.$$

The current price change is thus the sum of the potentially nonlinear error-correction and of a linear function of past price changes. The term m can be any complicated function which is

---

[1] α and β are sometimes summarized into $\Pi = \alpha \beta'$.

flexibly estimated from the data by modeling the partial impact of the past equilibrium errors nonparametrically. The resulting model is called a semiparametric VECM since some partial influences are modeled parametrically, while others are modeled nonparametrically. Its structure corresponds to the so-called semiparametric partial linear model (PLM) outlined, e.g., in Härdle et al. (2004, ch. 7).

The SPVECM can also be compactly written as:

(6) $Z_{0t} = m(\beta' Z_{1t}) + \Gamma Z_{2t} + \varepsilon_t = m(ect_{t-1}) + \Gamma Z_{2t} + \varepsilon_t.$

Estimation of the model proceeds in three steps as outlined in Gaul (2008, sec. 3.3). First, the cointegration matrix β is estimated via the Johansen approach (Juselius, 2007). Alternatively one may predetermine this matrix as in Gaul (2008). Based on either approach, the estimates of the equilibrium errors $\widehat{ect}_{t-1}$ are obtained.

Second, the short-run dynamics Γ are estimated. The following model which concentrates out the possibly nonlinear impact of the equilibrium errors is obtained:

(7) $Z_{0t} - E(Z_{0t}|ect_{t-1}) = \Gamma(Z_{2t} - E(Z_{2t}|ect_{t-1})) + \varepsilon_t,$

that is:

(8) $Z_{0t}^* = \Gamma Z_{2t}^* + \varepsilon_t.$

If the two conditional means were known, $\widehat{\Gamma}$ could be obtained via OLS. Gaul (2008) suggests using the Nadaraya-Watson estimator as follows:

(9) $\widehat{E}(Z_{0t}|ect_{t-1}) = \frac{1}{Th}\sum_{j=1}^{T} \frac{K\left(\frac{ect_{t-1}-ect_{j-1}}{h}\right)}{\widehat{f}(ect_{t-1})} Z_{0j}$

(10) $\widehat{E}(Z_{2t}|ect_{t-1}) = \frac{1}{Th}\sum_{j=1}^{T} \frac{K\left(\frac{ect_{t-1}-ect_{j-1}}{h}\right)}{\widehat{f}(ect_{t-1})} Z_{2j}$

where $K(*)$ is a kernel function, h is the bandwidth, and $\hat{f}(ect_{t-1}) = \frac{1}{Th}\sum_{j=1}^{T} K\left(\frac{ect_{t-1}-ect_{j-1}}{h}\right)$ is the kernel density estimator of the equilibriums errors. In order to avoid numerical problems, equation (8) is multiplied by $f(ect_{t-1})$ so that it becomes:

(11) $f(ect_{t-1})Z_{0t}^* = \Gamma f(ect_{t-1})Z_{2t}^* + f(ect_{t-1})\varepsilon_t$.

If the quantities in (11) are replaced by their respective estimates, the resulting estimator of Γ is:

(12) $\hat{\Gamma} = \left[\sum_{t=1}^{T} \hat{Z}_{0t}^* \hat{Z}_{2t}^{*\prime} \left(\hat{f}(\widehat{ect}_{t-1})\right)^2\right]\left[\sum_{t=1}^{T} \hat{Z}_{2t}^* \hat{Z}_{2t}^{*\prime} \left(\hat{f}(\widehat{ect}_{t-1})\right)^2\right]^{-1}$.

where $\hat{\Gamma}$ is a $(v \times vk)$ matrix, that is a product of a $(v \times vk)$and the inverse of a $(vk \times vk)$ matrix. The third step involves the estimation of $m(\beta' p_{t-1})$. Based on (12), equation (6) can be reformulated[2] as:

(13) $Z_{0t} - \hat{\Gamma}Z_{2t} = \tilde{Z}_{0t} = m(ect_{t-1}) + \varepsilon_t$.

The estimator of m is then the Nadaraya-Watson estimator applied to (13):

(14) $\hat{m}(z) = \frac{\frac{1}{Th}\sum_{t=1}^{T} \tilde{Z}_{0t} K\left(\frac{z-ect_{t-1}}{h}\right)}{\hat{f}(z)}$.

The nonparametrically estimated function $\hat{m}(ect_{t-1})$ may have any smooth functional form. It quantifies the potentially nonlinear partial impact of the lagged equilibrium error ect_{t-1} on the current price changes Δp_t which is reflected in the estimation by cleaning the current price changes from the (potentially nonlinear) partial impact of lagged price changes as depicted in

---

[2] Note that this estimator is in principle identical to the above mentioned approaches of Mancuso et al. (2003), Serra et al. (2006), and Goodwin and Vavra (2009) if the model is found to have no short-run dynamics, that is, if $\Gamma = 0$. However, some of these approaches use a local linear estimator instead of a local constant estimator as the SPVECM does.

equation (13). Because m is modeled nonparametrically, no parameters are estimated and the resulting estimates $\hat{m}$ can only be plotted.

Results

Time Series Properties

Table 1 shows the results of a number of unit root tests. It includes the Augmented Dickey-Fuller test (ADF), the test of Schmidt and Phillips (1992), Z(rho), and nonparametric range unit root (RUR) tests of Aparicio et al. (2006) (the RUR test and FB-RUR, the Forward-Backward range unit root test). The last two test statistics are robust to phenomena such as structural breaks or outliers. The results of the ADF test are mixed suggesting that four of the five series do not have a unit root. The Schmidt-Phillips test indicates at the 5% level of significance that the Dar es Salaam, Arusha, and the Iringa series have a unit root. The two nonparametric range unit root tests are unambiguous that all series have a unit root. Since they are more robust than the ADF and the Schmidt-Phillips tests, we give stronger weight to their results and regard the series as nonstationary.

As shown in Table 2, the Johansen trace cointegration test suggests one long-run relationship at the 5% level. Thus, we estimate a model consisting of five price series and one long-run equilibrium relationship whose results are addressed in the following subsection.

Estimation

Gonzalo (1994) compares various estimation methods for the long-run equilibrium. He finds that the Johansen maximum likelihood method is superior to all other considered methods in almost all considered cases. Hence, we use this method for the estimation of the cointegrating matrix β. We show the results of the adjustment parameter estimates of model (1) as well in order to compare it with the nonparametric estimates of model (6).

Three of four model selection criteria (Akaike Information Criterion (AIC), Final Prediction Error, and the Hannan-Quinn criterion) suggest one lag, while the Schwarz criterion suggests zero lags. Hence, we regard one lag in the estimation. Additionally, a trend and a constant are regarded as restricted to the cointegration space. Based on a preliminary estimation, we identify seven residual outliers according to the criterion $\hat{\varepsilon} > 3.3\hat{\sigma}_{\hat{\varepsilon}}$ as recommended by Hendry and Juselius (2001, p. 104). These outliers are accounted for by impulse dummies and included in the final estimation of the linear VECM (1). The estimated long-run relationship, that is, the long-run price equilibrium is:

$$(15) \qquad \hat{\beta}' p_{t-1} = 0.84 + p_{t-1}^{Dar} - 1.58^{***}p_{t-1}^{Ar} - 0.40p_{t-1}^{Ir} - 3.74^{***}p_{t-1}^{Mb}$$

$$+4.37^{***}p_{t-1}^{So} + 0.22^{***} \text{ trend.}$$

The adjustment coefficients of this (linear) model are shown in Table 3. Surprisingly, only the Dar es Salaam and the Songea prices react to deviations from equilibrium. The Songea price reacts strongest. The coefficients of the Arusha, the Iringa and the Mbeya prices are not significant.

Following the estimation of the linear model, the semiparametric model is estimated. As recommended by Gaul (2008), the Johansen procedure estimates of the long-run equilibrium and thus of the deviations from equilibrium are used in the first step of the estimation approach. In the second step, the short-run parameters Γ are estimated. Since one lag is indicated by the information criteria, Γ is (5×5).[3] Figure 2 plots the observed actual price changes $Z_{0t} = \Delta p_t$ and the transformed actual price changes $\tilde{Z}_{0t}$ of equation (13) which are used to estimate $m(ect_{t-1})$ in the third step.

---

[3] Results are available from the authors.

Figure 3 depicts the nonparametrically estimated and potentially nonlinear partial impacts $\hat{m}(ect_{t-1})$ of the deviations from equilibrium in the previous period on the price changes in the current period.[4] Two criteria (least squares cross-validation as well as the Kullback-Leibler cross-validation) were employed to assess the sensitivity of the results to the chosen bandwidth. Both cross-validation (CV) methods select the bandwidth which minimizes the respective criterion. Although the chosen bandwidths differ (Table 4), the estimated functions $\hat{m}(ect_{t-1})$ are almost identical.

For comparison, Figure 3 also presents plots of the (parametrically) estimated adjustment speeds $\hat{\alpha}$. The linear character of these coefficients is illustrated by the fact that the dashed line has a constant slope, that is, the magnitude of the coefficient is independent of the magnitude of the past deviation from equilibrium. In contrast, the slope of $\hat{m}(ect_{t-1})$ is allowed to change over the range of estimated equilibrium errors ect_{t-1}, that is, it may be constant but is not restricted to be constant.

Interestingly, the nonparametrically estimated functions $\hat{m}(ect_{t-1})$ are almost identical to the estimated $\hat{\alpha}$ of the linear model for almost all ranges of the equilibrium errors of all price series. Deviations from linearity only appear at the edges of the range where only a few observations are found and hence a few extreme observations can have a strong impact on the shape of the function. These results suggest that the partial impact of the equilibrium errors on current price changes is indeed linear.

Conclusions

This analysis suggests that monthly maize prices of the five large centers of Tanzania share a long-run equilibrium. Furthermore, the correction of short-run deviations from the long-run equilibrium takes place in a linear manner, that is, current changes in maize prices react by a

---

[4] These estimation results were obtained by using the *R*-package *np* (Hayfield and Racine, 2008).

constant rate to previous deviations from equilibrium. The semiparametric estimation results suggest that a linear model in form of the vector error-correction model can adequately model the spatial transmission of maize prices in Tanzania.

The estimation of confidence bands can shed further light on the potential shapes of functional forms of the price adjustment to disequilibrium and yield indications about the significance of the estimated functions. Extreme observations can exert a substantial impact on the shape of the nonparametrically estimated functional forms.. Hence, increasing the number of observations would yield more stable results by reducing the influence of such observations and might also lead to less ambiguous results of the unit root tests in the given context. We suspect that each time series should have at least 80 to 100 observations so that the semiparametric model can yield somewhat reliable results. Thus, the usual limitations of data availability in agricultural economics research might thus limit the applicability of the method. As illustrated above, the semiparametric vector error-correction model is a useful methodology for explorative analysis. It can provide evidence on whether parametric functions might be sensible modeling approaches and can furthermore suggest an appropriate parametric functional form. The approach is data-driven in the sense that the data is allowed to "speak" to the researcher since the approach allows for enormous flexibility in the estimation of functional forms. It is able to uncover the form of error-correction which characterizes the data at hand by modeling this partial in a nonparametric fashion. The model also accounts for the partial impact of past price changes modeled in a linear way.

This approach might thus interact with and support the development of theory in price transmission analysis. In contrast to this point of view, one may opt for theory-based model selection. However, since a number of competing models and underlying theories coexist, one may subjectively argue for either one or the other approach. The semiparametric vector error-correction model encompasses a number of possible model classes and may thus provide

evidence for one and against another. Given the very wide range of products and market structures that are analyzed in price transmission studies, as well as differing socio-economic and political market conditions and the heterogeneous behaviors of economic agents, theory cannot provide detailed prescriptions for each specific case. An approach such as the semiparametric vector error-correction model might provide help for the applied researcher to understand the specific structures of the market analyzed.

References

Aparicio, F., A. Escribano, and A. Sipols. 2006. "Range Unit-root (RUR) Tests: Robust Against Nonlinearities, Error Distributions, Structural Breaks and Outliers." *Journal of Time Series Analysis* 27: 545-576.

Campenhout, B. van. 2007. "Modelling Trends in Food Market Integration: Method and an Application to Tanzanian Markets." *Food Policy* 32: 112-127.

Fackler, P.L., and B.K. Goodwin. 2001. "Spatial Price Analysis." In B.L. Gardner, and G.C. Rausser (eds), *Handbook of Agricultural Economics*. Amsterdam: Elsevier Science, 971-1024.

Famine Early Warning Systems Network (FewsNet). 2009. *Maize Production and Market Flows in Eastern Africa.* Washington: USAID.

FAOSTAT. 2009. *International Agricultural Statistics. Food Balance Sheet.* Rome: U.N. Food and Agriculture Organization.

Gaul, J. 2008. "Three Essays on Unit Roots and Nonlinear Co-Integrated Processes." PhD dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, Germany.

Gaul, J., and E. Theissen. 2008. "A Partially Linear Approach to Modelling the Dynamics of Spot and Futures Prices." Working paper, Rheinische Friedrich-Wilhelms-Universitaet Bonn, Germany.

Gonzalo, J. 1994. "Five Alternative Methods of Estimating Long-run Equilibrium Relation-ships." *Journal of Econometrics* 60: 203-233.

Goodwin, B.K., and N. Piggott. 2001. "Spatial Market Integration in the Presence of Threshold Effects." *American Journal of Agricultural Economics* 83: 302-317.

Goodwin, B.K., and P. Vavra. 2009. "What Can We Learn from Spatial and Vertical Price Transmission Studies? Empirical Examples from U.S. Meat Markets." Paper prepared

for the Courant Research Centre "Poverty, Equity and Growth" Inaugural Conference at the University of Göttingen, Germany, July 1-3.

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

Hayfield, T., and J.S. Racine. 2008. "Nonparametric Econometrics: The np Package." *Journal of Statistical Software* 27(5). Accessed: October 2009. Available at: http://www.jstatsoft.org/v27/i05/.

Hendry, D.F., and K. Juselius. 2001. "Explaining cointegration analysis: Part II." *The Energy Journal* 22(1): 75-120.

Honaker, J., G. King, and M. Blackwell. 2009. "Amelia: Amelia II: A Program for Missing Data." Accessed: October 2009. Available at: http://CRAN.R-project.org/package=Amelia.

Ihle, R., S. von Cramon-Taubadel, and S. Zorya. 2009. "Markov-switching Estimation of Spatial Maize Price Transmission Processes Between Tanzania and Kenya". *American Journal of Agricultural Economics* 91(5): 1432-1439.

Juselius, K. 2008. *The Cointegrated VAR Model: Methodology and Applications*. New York: Oxford University Press.

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

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

Mancuso, A.J., B.K. Goodwin, and T.J. Grennes. 2003. "Nonlinear Aspects of Capital Market Integration and Real Interest Rate Equalization." *International Review of Economics and Finance* 12: 283-303.

MDG Africa Steering Group. 2008: "Achieving the Millennium Development Goals in Africa. Recommendations by the MDG Africa Steering Group." Accessed: December 2009. Available at: http://www.mdgafrica.org/media_kit/mdg%20africa%20summary-English.pdf.

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, Washington, USA.

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

R Development Core Team. 2009. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Accessed: February 2009. Available at: www.r-project.org.

Rapsomanikis, G., and P. Karfakis. 2008. "Margins Across Time and Space: Threshold Cointegration and Spatial Pricing Applications to Commodity Markets in Tanzania." Paper presented on the conference "Markets Integration and Food Security in Developing Countries", CERDI, CNRS-Université d'Auvergne Clermont-Ferrand, France, November 3- 4.

Regional Agricultural Trade Intelligence Network of the Eastern Africa Grain Council (RATIN). 2008. Accessed: November 2008. Available at: http://www.ratin.net/.

Saikkonen, P., and H. Lütkepohl. 2000. „Testing for the Cointegration Rank of a VAR Process with Structural Shifts." *Journal of Business & Economic Statistics* 18: 451-464.

Sarris, A., and E. Mantzou. 2005. "Linkages Between Domestic and International Maize Markets, and Market based Strategies for Hedging Maize Import Price Risks in Tanzania." FAO Commodity and Trade Policy Research Working Paper No. 14, Commodities and Trade Division, June, Rome, Italy.

Schmidt, P., and P.C.B. Phillips. 1992. "LM Tests for a Unit Root in the Presence of Deterministic Trends." *Oxford Bulletin of Economics and Statistics* 54: 257-287.

Serra, T., J.M. Gil, and B.K. Goodwin. 2006. "Local Polynomial Fitting of Spatial Price Relationships: Price Transmission in EU Pork Markets." *European Review of Agricultural Economics* 33: 415-436.

World Bank. 2008. "Regional Trade in Food Staples: Prospects for Stimulating Agricultural Growth and Moderating Short-Term Food Security Crises in Eastern and Southern Africa." Agriculture and Rural Development (AFTAR), Sustainable Development Department Africa Region, Report No. 46929 – AFR, November 27, Washington, USA.

World Bank. 2009. "Eastern Africa: A Study of the Regional Maize Market and Marketing Costs." Agriculture and Rural Development Unit (AFTAR), Sustainable Development Department, Country Department 1, Tanzania, Uganda and Burundi, Africa Region. Report No. 49831 - AFR, December 31, Washington, USA.
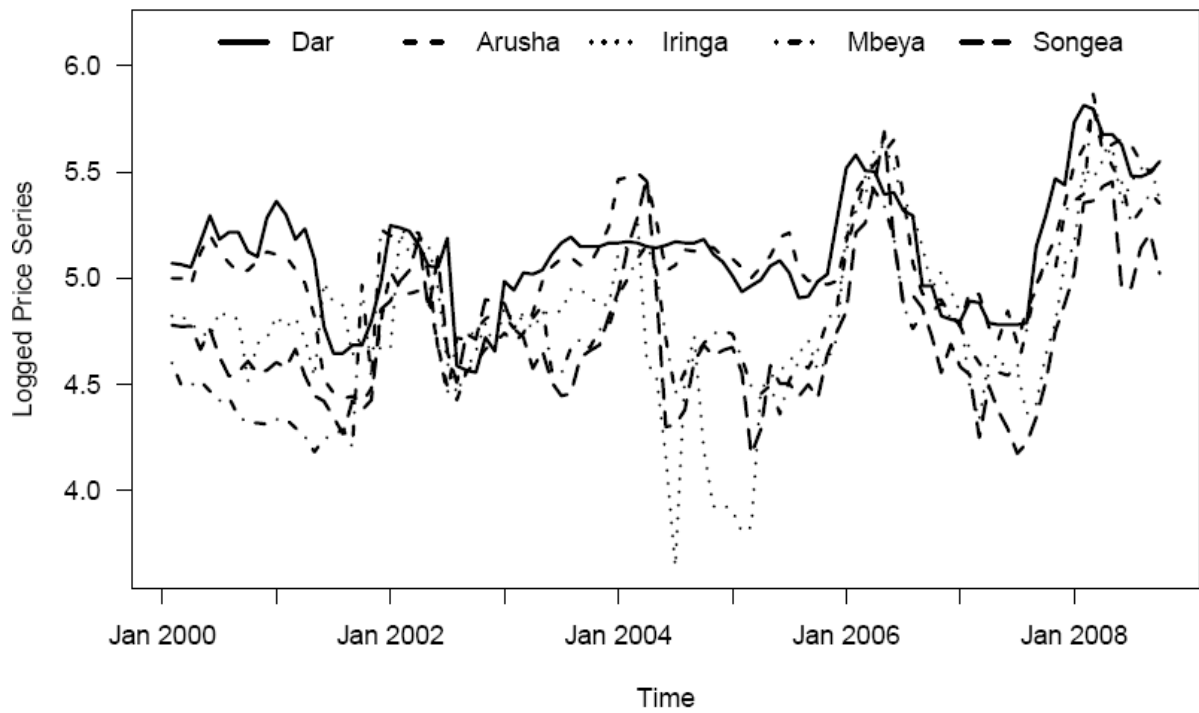
Source: RATIN (2008) and authors' depiction.

Figure 1: The price data

Table 1: Results of Unit Root Tests

Statistic	Data[a]	Dar	Arusha	Iringa	Mbeya	Songea	CV[b] 5%	CV[b] 1%
ADF[c]	L	-4.53***	-3.44**	-2.78	-3.59**	-3.70**	-3.41	-3.96
ADF[c]	D	-5.30***	-3.74***	-9.86***	-5.91***	-9.05***	-1.94	-2.56
Z(rho)[d]	L	-16.3	-13.4	-17.6	-19.9**	-20.7**	-18.1	-25.2
Z(rho)[d]	D	-78.0***	-38.1***	-101.7***	-119.6***	-86.3***	-18.1	-25.2
RUR[e]	L	1.37	1.66	1.46	1.46	1.66	1.17	0.98
FB-RUR[e]	L	1.86	2.28	2.00	2.21	2.00	1.79	1.52

Source: Authors' calculations.

[a] Form of the time series tested: L denotes levels of logged prices; D denotes first differences of logged prices.

[b] CV denotes critical value.

[c] The lag length is selected according to the Hannan-Quinn criterion. All ADF tests include a trend.

[d] For the calculation of the long-run variance of the Schmidt-Phillips test, a lag length of 8 is used.

[e] The null hypothesis of a unit root is rejected if the test statistic is smaller than the critical value.

Note: Two and three asterisks denote significance at the 5% and 1% levels, respectively.

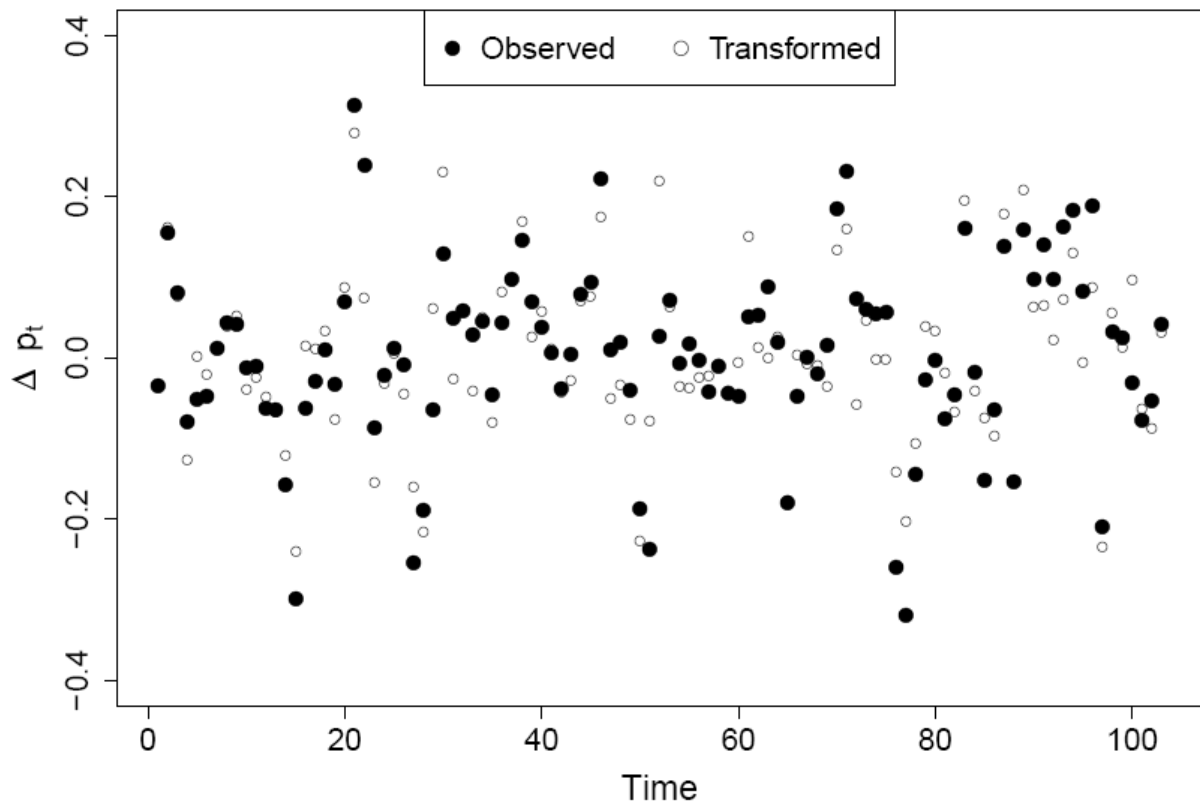**Table 2: P-values of the Johansen Trace Cointegration Test**

H_0	$\text{rank}(\Pi) \leq 0$	$\text{rank}(\Pi) \leq 1$	$\text{rank}(\Pi) \leq 2$	$\text{rank}(\Pi) \leq 3$	$\text{rank}(\Pi) \leq 3$
P-value	0.001	0.076	0.246	0.301	0.235

Source: Authors' calculations.

Table 3: Estimated Adjustment Speeds of the linear VECM (Equation (1))

| Coefficient	α^{Dar}	α^{Ar}	α^{Ir}	α^{Mb}	α^{So}
Estimate	-0.056	-0.023	-0.036	0.030	-0.144
P-value	0.007	0.264	0.337	0.236	<0.001

Source: Authors' calculations.

Source: Authors' calculations.

Figure 2: Example of observed vs. transformed changes of the Arusha price

Table 4: Optimal Bandwidths

Price series	Least Squares Cross-validation	Kullback-Leibler Cross-validation
Dar	0.231	0.196
Arusha	0.636	0.634
Iringa	0.509	0.595
Mbeya	0.304	0.256
Songea	0.330	0.143

Source: Authors' calculations.
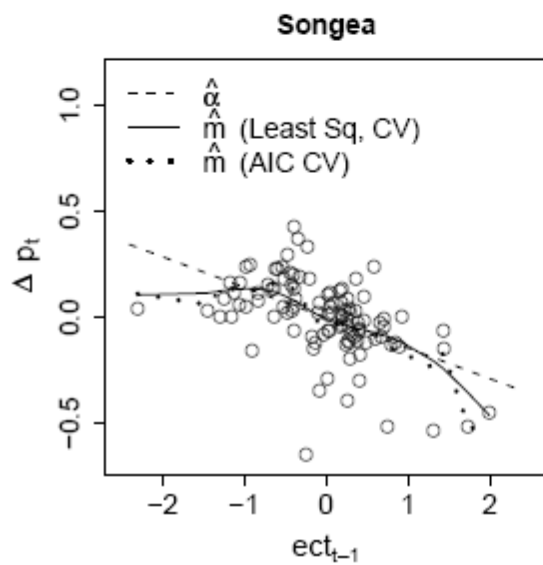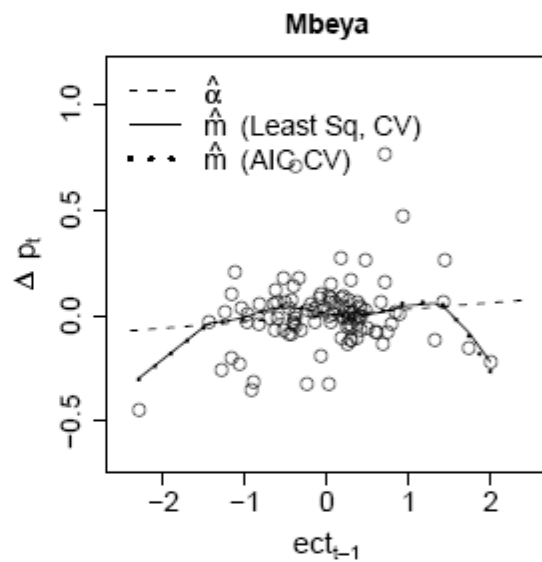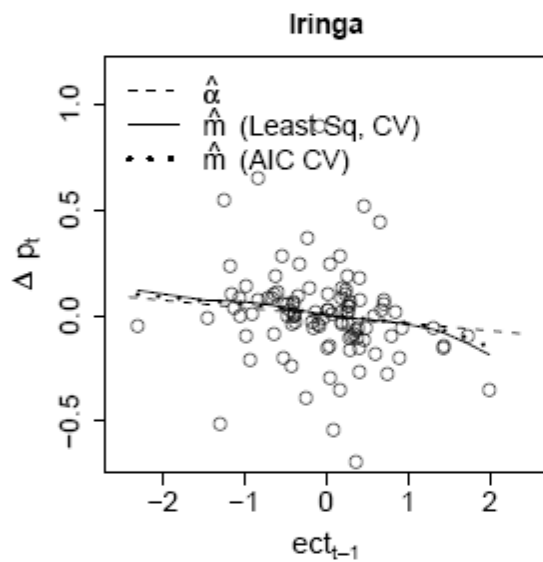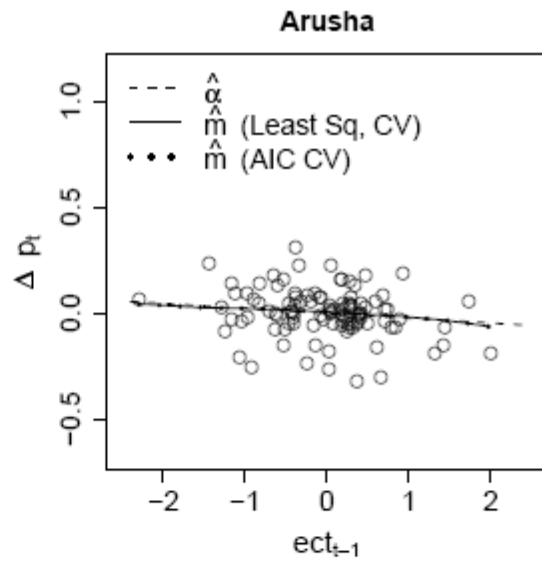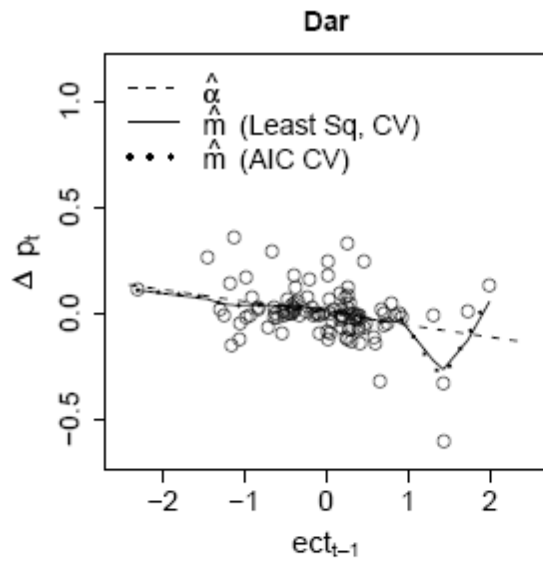
Source: Authors' calculations.

Note: $\hat{\alpha}$ denotes the parametric estimate according to model (2). $\hat{m}$ denotes the nonparametric estimates of model (2) for two different optimal bandwidths according to least squares cross-validation (Least Sq. CV) and Kullback-Leibler cross-validation (AIC CV).

Figure 3: Estimated parametric and nonparametric functions